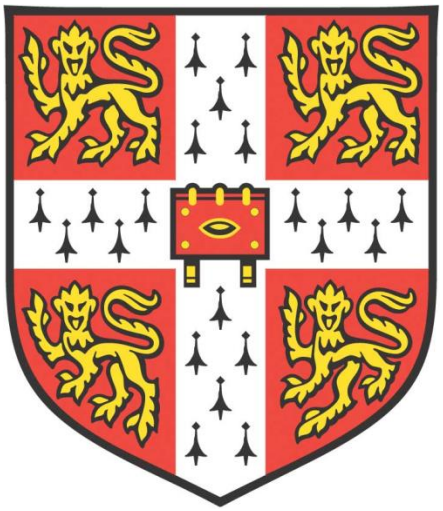
Biosynthesis and function of glucuronic acid substitution patterns on softwood xylan



Jan Jakub Łyczakowski

St Edmund's College

Department of Biochemistry

Graduate School of Life Sciences

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

September 2018

to my wife

Summary

Wood from coniferous trees is an important source of renewable biomass. It can contribute to provision of carbon neutral energy, biomaterials and housing for a growing population. Softwood is mainly composed of cellulose, galactoglucomannan, xylan and lignin. This thesis focuses on the biosynthesis and function of glucuronic acid (GlcA) decorations on softwood xylan. Results demonstrate that this GUX (GlucUronic acid substitution of Xylan)-dependent xylan branching is critical for the maintenance of biomass recalcitrance in a model vascular plant *Arabidopsis thaliana*. Experiments employing *in vitro* and *in planta* activity assays show that conifer transcriptomes encode at least two distinct GUX enzymes which are active glucuronosyltransferases. Interestingly, these enzymes have different specific activities, with one adding evenly spaced GlcA branches and the other one being able to add consecutive decorations. It is possible that these different patterns of xylan branching may have an impact on ability of xylan to interact with cellulose fibrils. To investigate the role for xylan binding to cellulose, Arabidopsis mutant plants in which this interaction is lost were evaluated alongside transgenic mutant lines in which the interaction may be restored. Results of this analysis indicate that the presence of cellulose-bound xylan might have an influence on plant vasculature integrity and thus it may have an effect on plant growth and biomass properties. Moreover, further results indicate that some xylan cellulose interaction is likely to occur in cell wall macrofibrils which can be detected in softwood. Taken together, this thesis provides insights into the process of conifer xylan glucuronidation and the possible role these branches may be playing in the maintenance of softwood recalcitrance and mechanical properties. In addition to identifying potential mutagenesis targets for improving softwood processing, this work is a proof of concept for the use of GUX enzymes for *in vivo* and *in vitro* biosynthesis of novel xylan structures with potential industrial application.

Declarations

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution.

In accordance with the School of Biological Sciences guidelines, this thesis does not exceed the prescribed word limit of 60,000 words.

Jan Jakub Łyczakowski

September 2018

Cambridge

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Professor Paul Dupree. Over the four years of my PhD project Prof. Dupree has helped me to grow as a scientist and offered me great support in a friendly and professional manner. I would like to especially thank him for encouraging me to get involved in collaborative projects what helped me to get the most out of my PhD studies.

Heartfelt thanks to Dr Marta Busse-Wicher who is my mentor and a true friend and was an amazing day-to-day supervisor from the moment I started my PhD project. I would also like to thank Prof. Ray Dupree, whose working ethic and commitment to research excellence is truly inspirational, for devoting his time and energy to help me develop some understanding of solid state NMR.

I was very fortunate to work with an amazing group of driven and passionate researchers in the Dupree Lab at the University of Cambridge who over the four years of my PhD became my very good friends. Thank you Louis, Joel, An, Fede, Nick and Caro for making the time spent working in the lab a really pleasant experience. Thanks to Oliver Terrett and Dr Henry Temple for all their kindness and friendly support. Thanks to Dr Li Yu for all her help and positive energy. Thank you to Dr Theodora Tryfona, Dr Nadine Anders, Dr Tom Simmons and Xiaolan Yu for all the friendly scientific guidance.

Thank you to my excellent collaborators from the Centre for Natural Material Innovation, the Open Plant Initiative and to Dr Raymond Wightman, Dr Weibing Yang and Dr Matthieu Bourdon from the Sainsbury Lab who greatly supported my work. Thank you to the Biotechnology and Biological Sciences Research Council which funded my project through the Doctoral Training Partnership programme. I wanted to express my sincere gratitude to my second supervisor Prof. Chris Howe and to the members of my Graduate Thesis Panel: Dr Nietlispach and Dr Braybrook.

My parents always encouraged and supported my scientific pursuits for what I am incredibly grateful. Thank you to all my family and friends in Poland who always believed in me. Last but not least, I wanted to thank so very much my amazing wife, Aleksandra Lewicka, for all the love, smile, support and understanding she always showed me. You are my star and this work in its entirety is dedicated to you.

Table of contents

List of abbreviations

2-AA	Anthranilic acid
AdhB	Alcohol Dehydrogenase B
AIR	Alcohol Insoluble Residues
ANOVA	Analysis of variances
ANTS	8-aminonapthalene-1,3,6-trisulphonic acid
AP	Alkaline phosphatase
APTS	9-aminopyrene-1, 4, 6-trisulfonate
Ara	Arabinose
At	*Arabidopsis thaliana*
AXY	Altered xyloglucan
BLAST	Basic Local Alignment Search Tool
BS	Brittle leaf sheath
C4H	Cinnamate 4-Hydroxylase
CAD	Cinnamyl Alcohol Dehydrogenase
CCR	Cinnamoyl-CoA Reductase
CoA	Coenzyme A
CP	cross-polarisation
CSLC	Cellulose Synthase Like C
Da	unified atomic mass unit or Dalton
DIG	Digoxigenin
DUF	Domain of Unknown Function
ESK	Eskimo
Gal	Galactose
GalA	Galacturonic acid
GAX	Glucuronoarabinoxylan
GH	Glycosyl hydrolase
GlcA	Glucuronic acid
GlcAT	Glucuronic acid transfer
Glu	Glucose
GT	Glycosyl transferase
GUX	GlucUronic acid substitution of Xylan

GX	Glucuronoxylan
GXM	GlucuronoXylan Methyltransferase
H	hour
HG	Homogalacturonan
Hz	Hertz
INAD	INADEQUATE
IRX	Irregular xylem
Kak	Kaktus
LB	Lysogeny broth
m/z	mass-to-charge ratio
MAGT	Mannan Alpha Galactosyl Transferase
Man	Mannose
MAS	Magic Angle Spinning
MES	4-Morpholineethanesulfonic acid
mRNA	messenger Ribonucleic acid
MS	Murashige and Skoog
NEB	New England Biolabs
NMR	Nuclear magnetic resonance
Pa	*Picea abies*
PACE	Polysaccharide Analysis by Carbohydrate gel Electrophoresis
Pdc	Pyruvate decarboxylase
PDSD	Proton driven spin diffusion
Pg	*Picea glauca*
Pt	*Pinus taeda*
RCF	Relative Centrifugal Field
RE	Reducing end
RGI	Rhamnogalacturonan-I
RGII	Rhamnogalacturonan-II
Rha	Rhamnose
RWA	Reduced wall acetylation
SSF	Simultaneous saccharification and co-fermentation

TBL	Trichome Birefringence-Like
UAX	Glucuronoarabinoxylooligosaccharide
UDP	Uridine diphosphate
UTP	Uridine triphosphate
UX	Glucuronoxylooligosaccharide
WB	Western Blot
XGA	Xylogalacturonan
XXT	Xyloglucan xylosyltransferase
XyG	Xyloglucan
Xyl	Xylose

Chapter 1: Introduction

Current estimates suggest that forests cover a third of global land area (Nabuurs et al., 2007) and as such they play an important role in the global carbon cycle (Whitehead, 2011). Indeed, terrestrial biomass may contain between 600 and 850 Pg of carbon which is equivalent to between two-thirds and nearly all of the atmospheric carbon (McKinley et al., 2011, Pan et al., 2011). The majority of this terrestrial carbon is stored in forests as wood (Ramage et al., 2017, Pan et al., 2011). The current classification system distinguishes two types of timber. Wood from Angiosperm trees is known as hardwood and the wood made by gymnosperm species is described as softwood (Ramage et al., 2017). Despite differences in microscopic structure both these types of timber are almost entirely formed from plant cell walls – an extracellular matrix made primarily from polysaccharides and phenolic compounds (Schweingruber, 2007).

Formation of a rigid cell wall is believed to be one of the critical adaptations required for the colonisation of terrestrial ecosystems by plants, which happened between 480 and 360 million years ago (Kenrick and Crane, 1997). While polysaccharide based cell walls are present in some charophyte algal relatives of land plants (Karol et al., 2001, Mikkelsen et al., 2014), the matrix has changed significantly since the divergence of land plants to allow for colonisation of different terrestrial environments (Popper, 2008). Land specific challenges which cell walls may be involved in tackling include mechanical support for upright growth and formation of water conducting tissues which are required in terrestrial environments (Mikkelsen et al., 2014, Sorensen et al., 2010). This vascular tissue is a defining feature of tracheophyte plants which include all hardwoods and softwoods.

Fossil evidence suggests that gymnosperm plants diverged between 310 and 290 million years ago (Mapes and Rothwell, 1991). Currently the gymnosperm clade consists of four distinct taxa: Pinales, Gnetophytes, Ginkgoales and Cycadales (Figure 1.1). From these groups the Pinales have the largest number of living species with the greatest economic importance (Ramage et al., 2017, Christenhusz and Byng, 2016). These include spruce trees which are the main source of timber in the UK (Moore, 2011). The Ginkgoales are another noteworthy gymnosperm clade. The Ginkgoales include a single species, *Ginkgo biloba*, which is considered to be unchanged since

its divergence more than 200 million years ago (Zhou, 2009). Thus, this living fossil is often an important part of phylogenetic studies looking at the evolution of trees (Guan et al., 2016). Fossil evidence indicates that Angiosperms diverged in the early Cretaceous, about 125 million years ago (Sun et al., 2011). Angiosperms include the hardwood producing taxon of eudicots alongside monocots and magnoliales. Despite contributing more than 55% of global plant species, eudicots account for only 25% of trees planted in timber-oriented managed forests (Ramage et al., 2017).

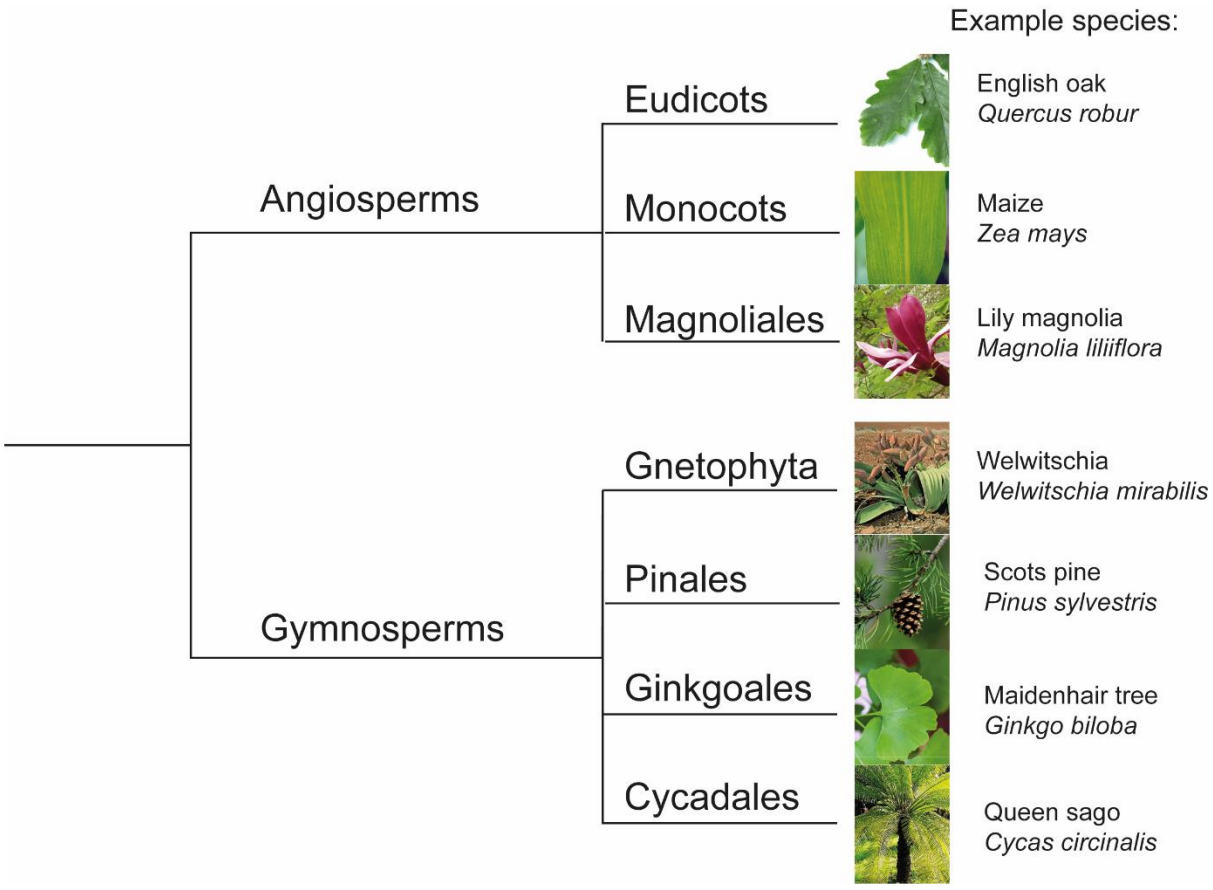


Figure 1.1 Schematic representation of phylogenetic relationships between angiosperm and gymnosperm taxa. Figure modified from (Busse-Wicher et al., 2016b). This representation does not reflect true phylogenetic distances or divergence order.

1.1 Cellular organisation of hardwood and softwood

In addition to the upward growth, maintained by the shoot apical meristem, all trees have the capacity to undergo secondary growth leading to thickening of the stem. For both hardwoods and softwoods this secondary growth is driven by the activity of stem cells localised in a lateral meristem known as the vascular cambium, which deposit phloem and xylem cells (Jones et al., 2013). The main function of the phloem, which is situated on the outside of the cambium, is to transport nutrients produced by photosynthesis. The xylem, localised on the inside of the cambium, is responsible for water transport from the roots of the plant. The fate of differentiating cells, following cambial cell division, is therefore dictated by their positioning either towards the stem periphery or interior following cytokinesis (Zhang et al., 2014). In addition to water transport, xylem is also partly responsible for the maintenance of stem strength in both hardwoods and softwoods (Ramage et al., 2017).

The morphology of xylem is one of the distinct differences between hardwood and softwood species (Figure 1.2) (Ramage et al., 2017). Softwood xylem is composed primarily from water conducting tracheid cells and parenchyma cells forming resin filled rays. Tracheids are between 2 and 4 mm long and approx. 30 μm wide and can form more than 90% of softwood xylem. Hardwood xylem is composed primarily from two types of cells: vessels and fibres. Hardwood vessels are responsible for water transport and are usually less than 1 mm long and have a range of diameters varying from 0.05 to 0.8 mm. Fibre cells provide mechanical support to eudicot trees and are normally up to 2 mm long and 15 μm wide. Fibres can form 50% of mature hardwood xylem. In addition to fibres and vessels hardwood also contains ray cells which play a similar function as in softwood (Myburg et al., 2013). For both hardwood and softwood distinct differences in the xylem deposited in different parts of the vegetative season lead to the formation of earlywood and latewood (Ramage et al., 2017). Earlywood usually has a larger total lumen volume allowing for greater water transport in early spring. Latewood consists of more densely packed smaller cells and its deposition stops in winter (Ramage et al., 2017).

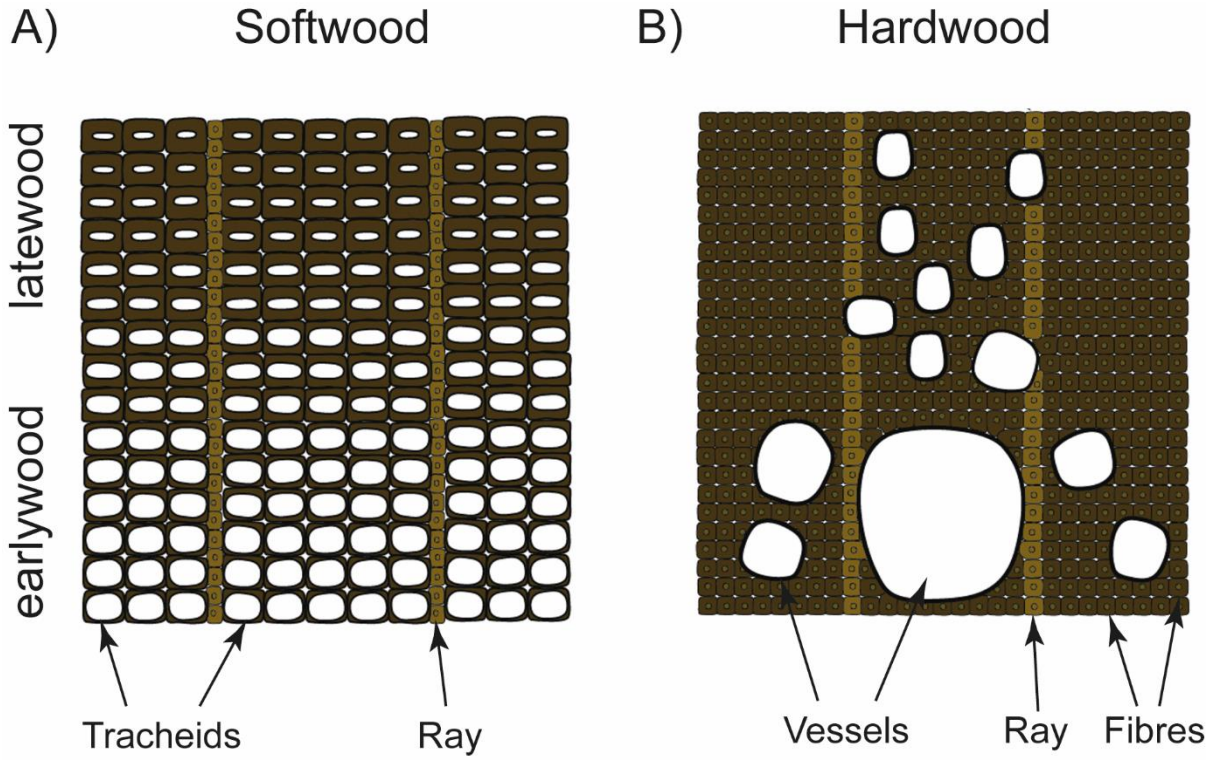


Figure 1.2. Morphology of wood. A) Softwood xylem consists of tracheids and rays. B) Hardwood xylem is made from vessels of varying size, rays and fiber cells. The presence of earlywood and latewood can be detected in most hardwood and softwood species and results in formation of annual growth rings. Figure modified from (Ramage et al., 2017).

1.2 Cell walls of hardwood and softwood xylem.

Plant cells are surrounded by a polysaccharide layer known as the primary cell wall. This extracellular matrix allows for the cell expansion needed during development but provides insufficient mechanical strength to support upright growth of the stem (Jones et al., 2013). Therefore, in addition to primary cell walls, some cell types which include for example, tracheids of softwood and fibre cells of hardwood also deposit secondary cell walls which have three distinct layers, known as S_1, S_2 and S_3 (Figure 1.3) (Myburg et al., 2013, Huang et al., 2003). This secondary cell wall is deposited on the inner surface of the primary cell wall and thus pushes it away from the cell centre. The deposition of secondary cell walls in the xylem is a terminal differentiation pathway and is often associated with apoptosis of the cell upon completion of the process (Meents et al., 2018).

The distinction of three layers in the secondary cell wall is based primarily on the differences in the angle between cellulose microfibrils and the longitudinal cell axis, a main microscopic feature of cell walls described in the next section (Figure 1.3). This varying microfibril angle is believed to cause a mechanical locking effect which dramatically increases mechanical strength of tracheary elements (Ramage et al., 2017, Cave and Walker, 1994). The S_2 layer contributes as much as 90% of the secondary cell wall mass and its microfibril angle is known to influence the tensile strength of wood (Ramage et al., 2017). The transverse strength is believed to be determined by S_1 and S_3 layers (Dinwoodie, 2000) . The cell walls of neighbouring wood cells are joined by a polysaccharide and lignin rich structure known as the middle lamella.

Figure 1.3. Cell walls of a wood tracheary element. Presented from (Myburg et al., 2013).

1.3 Structure and biosynthesis of cell wall components

This section will describe the chemical structure, assembly and biosynthesis of different cell wall components. These can be largely grouped into five categories: cellulose, pectin, hemicelluloses, lignin and cell wall proteins.

1.4 Structure and biosynthesis of cellulose.

Cellulose is the main constituent of the majority of plant cell wall material (Pauly and Keegstra, 2008). Due to the vast amount of plant material worldwide cellulose is often considered to be the most abundant bio-polymer on Earth (Somerville, 2006). At the chemical level cellulose has a simple structure of a homo-polymer of β-1,4-linked glucopyranose units. In hard- and softwood a single glucan polymer in cellulose can reach more than 5000 monomers in length (Hallac and Ragauskas, 2011). Glucan chains associate into a microfibril what causes the formation of a crystalline structure which is the main loadbearing feature in plant cell walls as described in section 1.2 of this Introduction. In the process of cellulose microfibril formation, glucan chains adopt a twofold helical conformation and assemble into sheets through edge-to-edge hydrogen bonding (Nishiyama et al., 2002). Native plant cellulose can form two distinct allomorphs known as cellulose I_α and I_β which can be described in terms of their crystallinity as monoclinic and triclinic respectively (Baker et al., 1997). While both allomorphs co-exist in native samples, the cellulose I_β allomorph (Figure 1.4) is dominant in both hardwood and softwood (Baker et al., 1997, Hallac and Ragauskas, 2011, Jarvis, 2018).

Figure 1.4 The hydrogen bonding scheme for cellulose I_β crystalline allomorph. Shaded bands and arrows denote intramolecular hydrogen bonds. Intermolecular bonding is denoted with green open arrows. Presented from (Jarvis, 2018)

Hydrogen bonding between glucan chains leads to the formation of cellulose microfibrils which vary in the size and number of individual glucan chains between different organisms (Jarvis, 2018). Some of the largest land plant cellulose microfibrils, with a cross section exceeding 8 x 8 nm, were observed in textile fibres, such as flax, and were estimated to contain more than 80 individual glucan chains (Thomas et al., 2013). The exact structure of microfibrils in wood is still a matter of debate. Different models propose the presence of 18 or 24 individual glucan chains in a microfibril (Jarvis, 2013). A particularly precise model has been proposed for the structure of spruce-wood microfibrils following analysis of softwood with a range of spectroscopic methods coupled to small-angle neutron and wide-angle X-ray scattering (Fernandes et al., 2011). In this model softwood microfibrils are suggested to have a cross section of 3.2 x 3.9 nm consistent with the presence of 24 individual glucan chains. The spruce microfibril was proposed to adopt one of two distinct shapes known as diamond and square (Figure 1.5). Further critical evaluation of these results suggested that some of the sizes measured for the spruce microfibrils might have been contributed by cellulose-bound hemicelluloses or water molecules and that the actual number of glucan chains may be 18 (Jarvis, 2018).

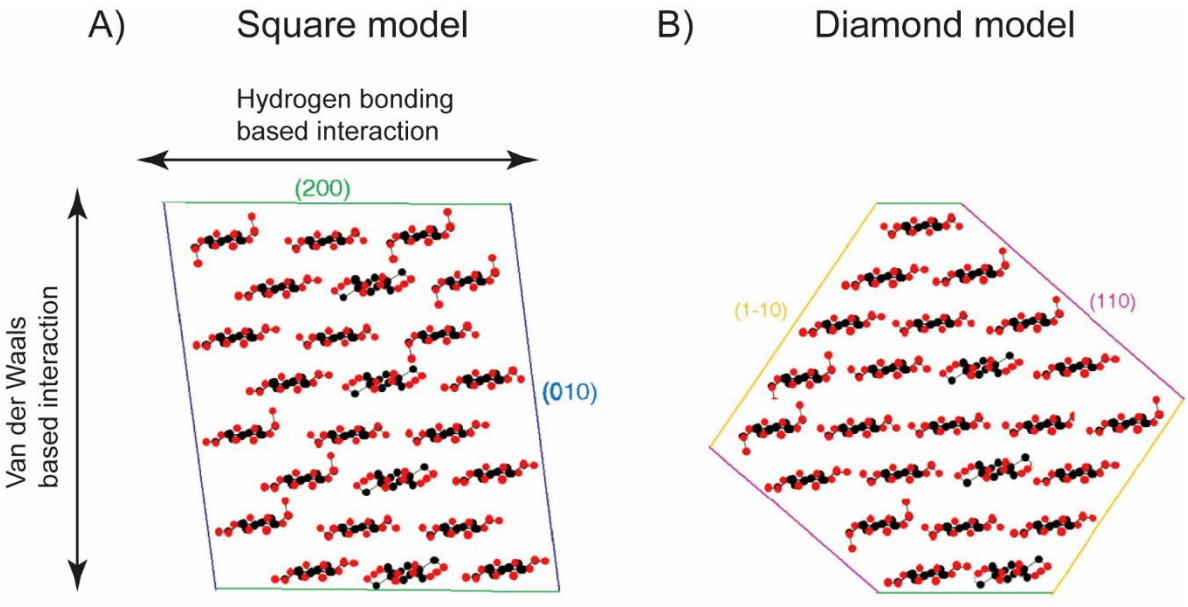


Figure 1.5 Proposed structure of 24 chain cellulose microfibril. A) Square cellulose crystal model with hydrophilic (010) and hydrophobic (200) surfaces marked. Major types of interactions maintaining the crystal structure are indicated. B) Diamond model of cellulose crystal with hydrophilic (1-10 and 110) surfaces marked. Presented models show horizontal cross-sections of cellulose microfibrils.

The presence of distinct faces of either hydrophilic or hydrophobic nature is an important feature of cellulose microfibrils. As indicated on Figure 1.5 the intra- and intermolecular hydrogen bonding occurs between individual glucan chains only in one plane (Gross and Chu, 2010). In the other direction the stacking of glucan sheets is driven by van der Waals interactions. Therefore, each microfibril has both hydrophilic surfaces, in which the hydrogen bonding network can be extended by addition of glucan chains or perhaps hemicelluloses, and hydrophobic surfaces, for which the exposed surface can interact with other components such as hemicelluloses via Van der Waals interaction (Fernandes et al., 2011). These surfaces can have different dimensions depending on the arrangement of glucan chains in the microfibril (Jarvis, 2018). As discussed in other parts of this chapter, the presence of the hydrogen bonding capacity on the hydrophilic surface of the cellulose fibril is proposed to be essential for its interaction with a xylan hemicellulose (Busse-Wicher et al., 2016a).

The biosynthesis of the cellulose glucan chains in plants and their assembly into the microfibril is catalysed by mobile plasma membrane localised cellulose synthesis complexes (CSCs) made from multiple cellulose synthase (CESA) catalytic subunits (Kimura et al., 1999) (Figure 1.6). Distinct CESAs are responsible for the biosynthesis of primary and secondary cell wall cellulose (Holland et al., 2000, Fagard et al., 2000). In *Arabidopsis thaliana* the CESA clade has 10 members (Richmond and Somerville, 2000, Kumar and Turner, 2015) and in xylem the secondary cell wall cellulose biosynthesis is catalysed by the activity of CESA4, 7 and 8 (Doblin et al., 2002). Mutation in genes encoding these secondary cell wall specific CESAs leads to a significant reduction in stem cellulose content and the collapse of xylem vessels. Therefore these CESAs have been also referred to as Irregular Xylem (IRX) 5, 3 and 1 respectively, as they were originally identified through discovery of *irx* mutants (Turner and Somerville, 1997, Taylor et al., 2003). Current models propose that a heterotrimer of CESA4, 7 and 8 would aggregate into a hexamer known as a rosette to use UDP-glucose and produce an 18 chain cellulose microfibril (Gonneau et al., 2014, Hofte and Voxeur, 2017). Auxiliary proteins, such as STELLO and KORRIGAN influence the activity and trafficking of CSCs and thus are important for the process of microfibril formation (Zhang et al., 2016, Vain et al., 2014). The exact mechanism by which these auxiliary proteins act remains unknown and is an active area of investigation.

Figure 1.6 Hypothetical model of cellulose synthase complex from (Jarvis, 2013).

1.5 Pectin – abundance and structure

Pectin is a complex heterogenous polysaccharide or group of polysaccharides which requires a minimum of 67 different transferase activities to be synthesised (Mohnen, 2008). Four principal pectic polysaccharides can be detected in plant cell walls: homogalacturonan (HG), xylogalacturonan (XGA), rhamnogalacturonan I (RGI) and rhamnogalacturonan-II (RGII) (Figure 1.7) (Harholt et al., 2010). Pectin is a major component of plant primary cell walls where its content can reach up to 35% of total biomass (Fry, 1988). In wood, pectins can account for up to 5% of cell wall material (Voragen et al., 2009) and may be less abundant in softwood than in hardwood (Sundberg et al., 1996).

Figure 1.7 Schematic structure of pectic polysaccharides. Presented from (Harholt et al., 2010). Presence of direct covalent linkages between pectic polysaccharides is speculative.

Structurally, homogalacturonan is the simplest and the most abundant pectic polysaccharide. It is a homopolymer of α-1,4-linked D-galacturonic acid (GalA). The synthesis of HG is localised to the Golgi apparatus and catalysed by the activity of GalActUronosylTransferase (GAUT) enzymes. GalA of HG can be acetylated on the O2 and O3 positions and methylesterified on C6 to neutralise the charge present on

the molecule (Ridley et al., 2001). This methylesterification is catalysed by pectin methyltransferase enzymes which use S-adenosyl-L-methionine as a donor (Goubet and Mohnen, 1999). Once in the wall, HG methylesters can be removed by a large family of pectin methylesterase enzymes (Micheli, 2001). This process may be an important part of maintenance of wall rigidity. De-esterified HG molecules can be cross-linked via calcium cations forming egg-box structures which have been suggested to be associated with wall stiffening (Hocq et al., 2017). Xylogalacturonan (XGA) has the same backbone structure as HG and can be both methylesterified and acetylated. In addition to these modifications, in XGA the GalA units of the galacturonan backbone can have a single β-1,3-linked xylosyl substitution added by a member of glycosyl transferase (GT) family 47 (Jensen et al., 2008). XGA has lower abundance than HG but is present in the primary cell walls of most plant tissues (Zandleven et al., 2007).

Rhamnogalacturonanan I and II have a much more complex structure than HG. Similarly to HG and XGA, the RGII backbone is a homopolymer of α-1,4-linked D-GalA. In RGII, the HG backbone is branched with 6 distinct side-chains formed from 13 different sugars linked via 21 types of linkages (Pellerin et al., 1996). This complexity is believed to make RGII the most intricate and recalcitrant polysaccharide structure on the planet (Ndeh et al., 2017). The RGI backbone is built from a repeating disaccharide of α-1,4-GalA-α-1,2-rhamnose (Rha). The rhamnose monomers of the RGI backbone can be branched on positions O3 or O4 with β-1,4-linked galactopyranosyl homopolymers (galactans), branched arabinans with a 1,5-linked α-L-arabinofuranose (Ara) backbone and/or arabinogalactan side chains (Ridley et al., 2001). Rhamnogalacturonan I may play an important role in the maintenance of cell wall molecular architecture with hemicelluloses, namely xylan, and other pectic polysaccharides being linked to RGI (Ralet et al., 2016, Vincken et al., 2003). This model and its importance for the structure of the cell wall will be further discussed in subsequent parts of this Chapter (section 1.12.1).

1.6 Hemicelluloses

Hemicelluloses are a class of non-cellulosic plant cell wall polysaccharides with an equatorial β-(1→4)-linkage forming their backbones (Scheller and Ulvskov, 2010).

Softwood and hardwoods contain different types and amounts of various hemicelluloses (Table 1). These in principle include xyloglucan, mannan and xylan.

Table 1. Summary of hemicellulose classes and their abundance in wood.
Adapted after (Scheller and Ulvskov, 2010)

Hemicellulose	Hardwood secondary cell wall content (% dry mass)	Softwood secondary cell wall content (% dry mass)
Glucomannan	2 to 5	Not detected
Galactoglucomannan	Up to 3	10 to 30
Acetylated Glucuronoxylan	20 to 30	Not detected
Glucuronoarabino-xylan	Not detected	5 to 15
Xyloglucan	Minor	Minor

1.7 Xyloglucan

Xyloglucan (XyG) is made from a β-(1→4)-linked glucan backbone with α-(1→6) xylose branches. The xylose branches of XyG can be substituted further with β-(1→2)-linked Galactopyranose which may carry α-(1→2)-linked L-Fucopyranose (Scheller and Ulvskov, 2010). Xyloglucan is the main hemicellulose in primary cell walls of non-monocot seed plants (Fry, 1989) but it forms only a minor component of mature hardwood (Scheller and Ulvskov, 2010) and softwood (Sjöström and Westermark, 1999), so is not discussed in detail here. Despite its low abundance in xylem, remodelling of xyloglucan structure may play part in the development of vascular cells in poplar (Nishikubo et al., 2011).

A significant body of knowledge exists about the biosynthesis of xyloglucan. The synthesis of the xyloglucan backbone is catalysed by the activity of Cellulose-Synthase Like C (CSLC) proteins (Cocuron et al., 2007). The addition of xylose branches is catalysed by GT34 family members xyloglucan xylosyl transferases

(XXTs) and the galactose branch on the xylose decoration is added by GT47 family enzymes MUR3 and XLT2 (Pauly and Keegstra, 2016). Fucosylation of the galactose branch is performed by xyloglucan fucosyltransferases, the FUTs (Perrin et al., 1999).

1.8 Mannan

Mannose based polysaccharides are some of the main hemicelluloses present in vascular plants (Moreira and Filho, 2008). The backbone of mannan is composed from β-1,4-linked residues of mannose (Man). Polymers with a backbone formed from both β-1,4-linked Man and glucose (Glc) residues are known as glucomannans (Scheller and Ulvskov, 2010). The mannosyl residues of both mannan and glucomannan can be substituted by an α-1,6-linked galactose (Gal) branch. Galactosylated glucomannan is known as galactoglucomannan (GGM). In addition to galactose branches mannans and glucomannans can be acetylated on carbons 2 or 3 of the mannose monomer. The exact composition of the GGM backbone and the frequency of galactose and acetyl branches remains largely unknown. For acetylated softwood GGM, studies report that the Man:Glu:Gal ratio may be FGTbetween 4:1:0.1 and 3:1:1 (Sjostrom, 2013). In angiosperms the structure of seed mucilage GGM has been recently described (Voiniciuc et al., 2015b, Yu et al., 2018). This specific polymer is not acetylated and the backbone is composed from a repeating disaccharide of Glc-Man with frequent Gal branches. The ratio of Man:Glu:Gal monosaccharides in this structure is equal to 3:3:2.2 and its exact structure may be important for the maintenance of cell wall structure (Yu et al., 2018).

Unlike mannan, which is mostly present as a seed storage carbohydrate (Meier and Reid, 1982), glucomannan and GGM are common cell wall structural polysaccharides (Moreira and Filho, 2008). Glucomannans are abundant in cell walls of early land plants such as Bryophyte lineages of mosses and liverworts (Moller et al., 2007). GGM and especially acetylated GGM is the main hemicellulose in softwood where it accounts for up to 30% of wood material (Scheller and Ulvskov, 2010, Willfor et al., 2005a). In addition to being the most abundant hemicellulose in industrially important softwood species, acetylated GGM is easily extractible from conifer wood and thus has been proposed as a component of sustainable biomaterials (Willfor et al., 2008). GGM has been detected in angiosperm species, such as tomato (Seymour et al., 1990) and kiwi (Schroder et al., 2001), where it may be involved in the process of fruit

ripening (Seymour et al., 1990). GGM has also been detected in hardwood where it may be more abundant in xylem cell walls than in phloem cell walls (Willfor et al., 2005b). The possible function of GGM in the maintenance of wall architecture is discussed in subsequent sections of this introduction Chapter.

The synthesis of the mannan and glucomannan backbones is a Golgi localised process catalysed by the activity of the Cellulose Synthase-Like A (CSLA) family of enzymes (Liepman et al., 2005, Goubet et al., 2009). In the *A. thaliana* model plant the *csla2/3/9* triple mutant has close to no detectable mannan in basal stem cell walls (Goubet et al., 2009). Specific CSLA enzymes are involved in the formation of different GGM structures. For example, in *A. thaliana* seed mucilage CSLA2 is responsible for the synthesis of the patterned GGM with a repeating Glc-Man disaccharide structure (Yu et al., 2014, Yu et al., 2018). The addition of acetyl branches to GGM may be catalysed by members of the Trichome Birefringence-Like (TBL) family. Specifically in *A. thaliana,* TBL25 and TBL26 have been proposed as mannan acetyltransferases but experimental demonstration is still lacking (Gille et al., 2011). The addition of galactose branches to glucomannan is catalysed by the activity of GT34 family enzymes (Edwards et al., 1999, Lombard et al., 2014, Voiniciuc et al., 2015b). These galactosyltransferases are likely to have specific structural requirements for the glucomannan they decorate. For example, the A. thaliana Mannan α-Galactosyl Transferase1 (MAGT1/GTL6/MUCI10) can, despite being tested also on mannohexaose and spruce mannan oligosaccharides, decorate only CSLA2-made patterned glucomannan with galactose *in vitro* (Voiniciuc et al., 2015b, Yu et al., 2018).

1.9 Xylan

As analysis of softwood xylan biosynthesis is a main interest of this thesis significant focus will be placed on the description of this hemicellulose in the Introduction chapter. In addition to describing the properties of the xylan polymer this section will describe some approaches used to study its structure which will be applied throughout this project.

1.9.1 Structure of xylan

Xylan is a polymer of β-1,4-linked xylopyranose (Xyl) residues. It is not known what the total degree of polymerisation (DP) is for xylan but most estimates range from 80

to 200 xylose monomers (Faik, 2010, Pena et al., 2007, Sjostrom, 2013). The xyloses of the xylan backbone can be decorated with various modifications. Most xylans for which the structure has been studied, carry variable amounts of α-1–2 linked glucuronic acid (GlcA) branches (Scheller and Ulvskov, 2010). These GlcA branches can be further modified. The most common modification of GlcA includes methylation on carbon 4 leading to formation of 4-O-methyl-glucuronic acid (MeGlcA) (Scheller and Ulvskov, 2010). The degree of GlcA methylation varies between different plants (Ebringerova et al., 2005, Pena et al., 2007, Kulkarni et al., 2012). For xylans which carry both methylated and non-methylated GlcA the branching is denoted as [Me]GlcA to represent both forms (Mortimer et al., 2010). In addition to methylation, GlcA can be further branched by neutral sugar decorations. These include α-1,2-linked arabinofuranose, in the primary cell wall of *A. thaliana* (Mortimer et al., 2015) and monocots (Pena et al., 2016) and α-D-Gal linked to O-2 of GlcA, reported in Eucalyptus (Shatalov et al., 1999).

In addition to [Me]GlcA decorations, xylose monomers of xylan can be acetylated or directly branched with neutral sugars. Acetylation of xylan is present on carbon 2, carbon 3 or both carbons of the monomer (Scheller and Ulvskov, 2010). Direct modifications of the xylan backbone with neutral sugars include α-1,2–linked and α-1,3–linked arabinofuranosylation (Anders et al., 2012) and the possible presence of β-1,2-linked xylose branches in Arabidopsis (Voiniciuc et al., 2015a) and psyllium (Jensen et al., 2013) mucilage. The arabinose branches of xylan can be further modified by xylose on position O-2 (Chiniquy et al., 2012) or ferulic and coumaric acids linked to position O-5 (Scheller and Ulvskov, 2010). Variation in the degree of modification with these xylan branches across the plant kingdom is significant and will be discussed further in other sections of this chapter. Figure 1.8 summarises the main types of xylan branching observed in plant cell walls.

Figure 1.8 Summary of main types of xylan branching in angiosperms and gymnosperms. Figure from (Dodd and Cann, 2009). Xylan backbone is denoted in black with red MeGlcA branch, green arabinose branch and purple acetylation on carbons 2 and 3. Ferulation of arabinose is marked in blue.

The reducing end (RE) oligosaccharide is another important feature of the xylan molecule. This β-D-xylose-(1-3)-α-L-rhamnose-(1-2)-α-D-galacturonic acid-(1-4)-D-xylose structure is present on the reducing ends of hardwood and softwood xylan (Andersson et al., 1983).

1.9.2 Biosynthesis of xylan

Xylan biosynthesis is a multi-step process localised to the Golgi apparatus of plant cells (Scheller and Ulvskov, 2010). The biosynthesis of the xylan backbone is catalysed by members of GT families 43 and 47. In *A. thaliana* this process is catalysed mostly by IRX9, IRX10 and IRX14 proteins in secondary cell walls and by IRX9-Like (IRX9L), IRX10L and IRX14 in primary cell walls (Mortimer et al., 2015, Wu et al., 2010). Members of GT family 43 (IRX9 and 14) are believed to associate with a member of GT47 family, IRX10, to form a Xylan Synthase Complex (XSC) (Zeng et al., 2016). Despite the fact that the presence and exact stoichiometry of this complex remains to be confirmed, it is believed that IRX10 is the catalytic subunit involved in elongation of the xylan backbone while other members of the XSC may be involved in maintenance of its stability or activity (Urbanowicz et al., 2014, Ren et al., 2014). Mutations in XSC members lead to a reduction in the xylan content, vessel collapse and plant dwarfing (Wu et al., 2010, Brown et al., 2009). In addition to *IRX9*, *IRX10* and *IRX14*, mutations of the *IRX15* and *IRX15L* genes also reduce the xylan content in basal stem and decrease its DP (Jensen et al., 2011, Brown et al., 2011). Despite

the presence of this phenotype in *irx15/15l A. thaliana* plants it is unlikely that the protein products of these mutated genes are glycosyltransferases acting in the XSC and their role in xylan biosynthesis remains unknown (Brown et al., 2011). The presence of the RE oligosaccharide of xylan is also believed to be important for the backbone synthesis. Arabidopsis plants mutated in genes encoding IRX7, IRX8 and PARVUS proteins have no detectable RE oligosaccharide structure (Brown et al., 2007). In addition to the loss of the RE oligosaccharide mutations of *IRX7, 8* and *PARVUS* leads to a reduction in xylan content, vessel collapse and plant dwarfing suggesting that the RE biosynthesis is linked to backbone formation as either its primer or terminator (Brown et al., 2007).

The branching of xylan is catalysed by multiple enzymes and similarly to backbone synthesis is a Golgi localised process. The acetylation of the xylan backbone is catalysed by a plant specific Domain of Unknown Function (DUF) 231 family, TBL (Bischoff et al., 2010, Xiong et al., 2013). The TBL family has a total of 46 members in *A. thaliana* and a protein known as both ESKIMO1 (ESK1) and TBL29 is the main xylan O-acetyltransferase in this model species (Xiong et al., 2013). The *esk1/tbl29 A. thaliana* plants have a ~60% reduction in xylan acetylation levels. This reduction results in collapse of xylem vessels and plant dwarfing (Lefebvre et al., 2011, Xiong et al., 2013). Despite the reduction in both O-2 and O-3 monoacetylation in *esk1* plants (Yuan et al., 2013), the *in vitro* analysis suggests that ESK1 is a O-2 specific xylan acetyltransferase using acetyl-CoA as a donor (Urbanowicz et al., 2014). The O-3 acetylation may thus be a result of spontaneous migration of acetyl groups from the O-2 position (Urbanowicz et al., 2014). In addition to ESK1 other TBL enzymes were also demonstrated to be involved in xylan O-acetylation. Most notably, *A. thaliana* TBL33 and TBL32 are likely to be involved in O-3 specific acetylation of xylose residues carrying the GlcA branch on carbon 2 (Yuan et al., 2016). Mutation in both *TBL33* and *TBL32* does not result in the dwarfing phenotypes observed for *esk1* plants. In addition to TBLs, other proteins are also involved in the process of xylan acetylation. The recently described rice GDSL esterase Brittle Leaf Sheath1 (BS1) may be involved in removal of some acetyl branches from the backbone (Zhang et al., 2017). Reduced Wall Acetylation (RWA) proteins (Manabe et al., 2013) and Altered Xyloglucan 9 (AXY9) (Schultink et al., 2015) are believed to be involved in the process of supplying the acetyl-CoA acceptor to TBLs. It is proposed that RWAs are

transporters of the acetyl-CoA (Manabe et al., 2013) and the function of AXY9 is not well understood. In *rwa* triple and quadruple mutants and in *axy9* single mutant plants the acetylation is significantly decreased for multiple polysaccharides suggesting that both proteins act upstream of the TBL activity and are not xylan specific.

The addition of neutral sugar branches to xylan is believed to be catalysed by members of the GT61 family of proteins (Anders et al., 2012, Zhong et al., 2018, Voiniciuc et al., 2015a). Grass xylan is extensively modified with arabinofuranose and the GT61 from wheat responsible for the addition of α-(1,3)–linked arabinose was characterised using mutagenesis and gain of function studies (Anders et al., 2012). Cytosolic enzymes from the BAHD acyltransferase family are likely to be responsible for the addition of ferulic acid to arabinose prior to its positioning on the xylan backbone (de Souza et al., 2018). In Arabidopsis, MUCI21 is a GT61 family member proposed to be responsible for the addition of xylose to the xylan backbone in mucilage cell walls (Voiniciuc et al., 2015a). Rice GT61s capable of addition of 2-O-linked xylosyl branches onto the xylan backbone has also been characterised (Zhong et al., 2018).

Glucuronic acid is a decoration common to all xylan structures studied to date (Scheller and Ulvskov, 2010). This acidic decoration is added by GT8 family members named GlucUronic acid substitution of Xylan (GUX). GUX enzymes are type-II Golgi localised proteins with an N terminal transmembrane region and a globular catalytic domain localised towards the C terminus of the protein (Rennie et al., 2012). *Arabidopsis thaliana* (At) genes coding for GUX1 and GUX2 were initially identified using co-expression analysis with other enzymes involved in the process of xylan synthesis (Mortimer et al., 2010). Arabidopsis plants lacking active copies of *GUX1* and *GUX2* have no detectable [Me]GlcA branches on secondary cell wall xylan (Mortimer et al., 2010). The absence of [Me]GlcA decorations in *gux1/2* plants does not result in plant dwarfing (Lyczakowski et al., 2017). The GUX clade in the GT8 family consists of 5 members in *A. thaliana* and *in vitro* experiments have confirmed that *A. thaliana* GUX1, GUX2 and GUX4 are active glucuronosyltransferases and can transfer GlcA from UDP-GlcA onto a xylohexaose acceptor (Rennie et al., 2012). In addition to these three enzymes, AtGUX3 has been reported to be responsible for glucuronidation of primary cell wall xylan *in vivo* (Mortimer et al., 2015). Despite the fact that the Arabidopsis *gux1/2/3* triple mutant has no detectable glucuronic acid on stem xylan they still grow like WT plants (Mortimer et al., 2015). The primary cell wall

GlcA added by GUX3 can be further modified by the addition of an arabinopyranose as discussed in section 1.9.1. The enzyme responsible for addition of this pentose remains unknown. The process of xylan glucuronidation and acetylation are likely to be linked in the model hardwood species *A. thaliana.* Specifically, reduction in xylan acetylation in *esk1* mutant plants results in an increase in xylan glucuronidation, likely as a direct result of altered GUX1 activity (Grantham et al., 2017).

Methylation of GlcA branches is catalysed by the activity of DUF579 containing proteins known as glucuronoxylan methyltransferases (GXM) (Urbanowicz et al., 2012). Activity experiments indicate that GXM proteins use S-adenosyl-L-methionine to catalyse addition of a methyl group onto the position O-4 of GlcA. In the *A. thaliana* model, *gxm1/2/3* triple mutant plants have greatly reduced levels of GlcA methylation (Li et al., 2013, Cornuault et al., 2015). Similarly to the removal of GlcA itself in *gux1/2/3* plants these modifications in its methylation state do not result in plant dwarfing (Lyczakowski et al., 2017).

No enzymes in xylan synthesis have been characterised from gymnosperms and searching for these is a main topic of this thesis.

1.9.3 Diversity of xylan structure in hardwood and softwood

Despite sharing a backbone structure of β-1,4-linked xylose residues with a specific reducing end tetrasaccharide the exact type and pattern of branches present on the backbone differs between hardwood and softwood (Busse-Wicher et al., 2016b). Hardwood cell walls contain acetylated glucuronoxylan (GX) while in softwood cell walls the polymer is a glucuronoarabinoxylan (GAX) (Scheller and Ulvskov, 2010). These differences in the structure and pattern of xylan branches between hardwood and softwood (summarised in Figure 1.9) are likely to be important for the xylan-cellulose interaction which will be discussed in the next sections of this Chapter.

Xylan is the main hemicellulose in hardwood and can account for up to 30% of cell wall material (Scheller and Ulvskov, 2010). Acetylation is the main type of branching observed on hardwood xylan. On average, 50% of xylose monomers carry either O-2 or O-3 linked monoacetylation or diacetylation (Xiong et al., 2013). The distribution of acetyl branches on hardwood xylan is not random, with predominantly every other monomer being acetylated and very few adjacent acetylated xyloses present in Arabidopsis xylan (Busse-Wicher et al., 2014). In addition to acetylation, the hardwood

xylan is glucuronidated with an average molar ratio of xylose and GlcA being equal to 8:1 (Mortimer et al., 2010). Similarly to acetylation, the distribution of [Me]GlcA branches on hardwood xylan is not random, with two discrete domains, known as compatible and incompatible, being present on each xylan chain (Bromley et al., 2013, Busse-Wicher et al., 2016a). The compatible domain is made from xylan regions with evenly spaced [Me]GlcA decorations present predominantly on every 6th, 8th and 10th xylose monomer. In addition to this short distance patterning, over longer distances, exceeding 20 xylose monomers, the preference for even spacing is also maintained in the compatible xylan regions. In *A. thaliana* the compatible pattern of [Me]GlcA branches accounts for ~70% of the acidic branching and is added to the secondary cell wall xylan backbone by the GUX1 enzyme. In the incompatible domain, the [Me]GlcA branches are more densely packed with main distances between branched xyloses being equal to 5, 6 and 7 monomers. In the *A. thaliana* model the incompatible xylan regions are synthesised by the GUX2 enzyme and account for ~30% of the secondary cell wall xylan. Compatible and incompatible domains are present on a single xylan polymer but their relative positioning is unknown (Bromley et al., 2013).

In softwood, the glucuronoarabinoxylan accounts for up to 15% of wood (Scheller and Ulvskov, 2010). In comparison to the hardwood hemicellulose, softwood xylan is not acetylated and has frequent arabinose decorations (Scheller and Ulvskov, 2010). The degree of xylan glucuronidation in softwood is higher than in hardwood and the Xyl to MeGlcA ratio is close to 6:1 (Willfor et al., 2005a). Interestingly, over the majority of softwood xylan the MeGlcA is present on every 6th xylose monomer (Busse-Wicher et al., 2016b). In addition to this regular, even, pattern a small proportion of xylan with consecutive monomers carrying the MeGlcA branch has also been detected in spruce, cedar and larch (Martinez-Abad et al., 2017, Yamasaki et al., 2011, Shimizu et al., 1978). It is not known if in softwood xylan these differently patterned regions exist on a single polysaccharide. It is also not known if they are synthesised in the same cell types or if different GUX enzymes are responsible for their biosynthesis. In contrast, to the *A. thaliana* but in common with true hardwoods, virtually all GlcA branches on softwood xylan are methylated on the O-4 position (Busse-Wicher et al., 2016b). On softwood GAX the majority of O-3-linked arabinose branches are present two xyloses towards the non-reducing end of the polysaccharide from the monomer carrying the MeGlcA branch (Busse-Wicher et al., 2016b). In addition to that, a spacing of 4 xylose

monomers between glucuronidated and arabinosylated xyloses has also been reported (Martinez-Abad et al., 2017).

It is unknown how the different patterns of [Me]GlcA decoration are formed in softwood and hardwood. Using Arabidopsis as a model hardwood species it was established that AtGUX1 and AtGUX2 play distinct functions generating compatible and incompatible domains of GlcA patterning respectively (Bromley et al., 2013). Interestingly, overexpression of AtGUX1 can only marginally increase the extent of compatible domain and the over-expression of AtGUX2 has no effect on the structure of WT Arabidopsis xylan (Bromley et al., 2013). This observation suggests that other factors may influence formation of xylan domains. Indeed, disruption of xylan acetylation in *tbl29* Arabidopsis mutant was observed to lead to AtGUX1 dysfunction and loss of compatible domain patterning (Grantham et al., 2017). Therefore, xylan acetylation may influence AtGUX1 specificity towards generating the compatible pattern. This specific guidance mechanism cannot exist in softwood which lacks xylan acetylation therefore other xylan branches, namely arabinose, may influence conifer GUX activity. The exact determinants of GUX activity remain therefore an unexplored area of research which may have significance towards generation of compatible and incompatible xylan regions. These in turn are likely to influence xylan cellulose interaction and may impact cell wall properties (discussed in section 1.12.2). The lack of available protein structures is a significant hurdle in identifying determinants of GUX specificity. The only GT8 family member with the structure solved is a cytosolic human glycogenin (Chaikuad et al., 2011). Structural studies on Golgi localised plant GTs are challenging but a recent advance reporting the structure of XXT may contribute to the advance in this field (Culbertson et al., 2018). Despite that, identification of structural features which enable GUX enzymes to generate different GlcA patterns, such as the consecutive patterning on softwood xylan, remains an interesting research area.

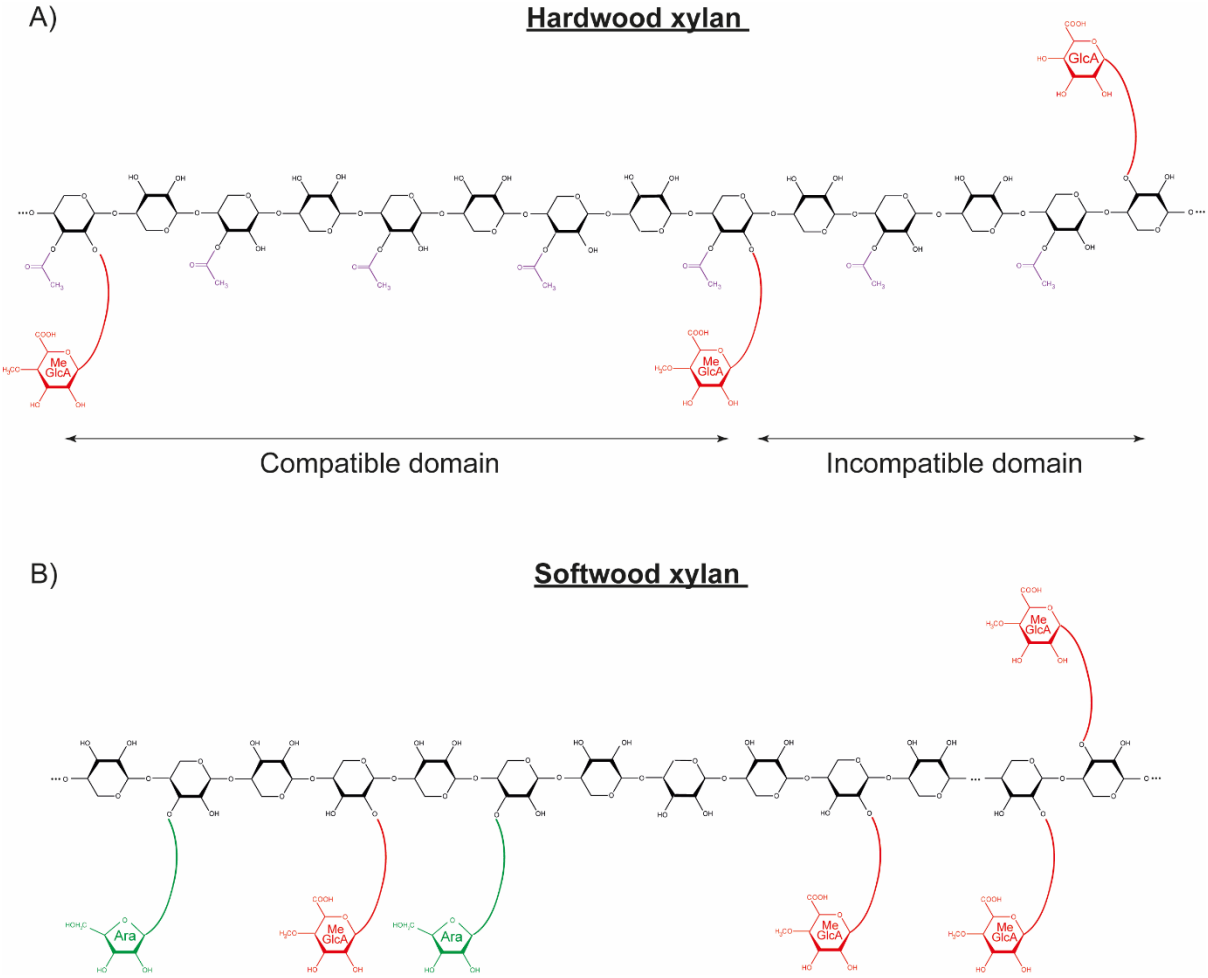


Figure 1.9 Structural models of hardwood and softwood xylan. A) Compatible and incompatible [Me]GlcA domains of hardwood xylan. This image is not an accurate representation of proportions of different domains on one xylan chain B) Patterning of softwood xylan with evenly spaced Ara and MeGlcA decorations marked alongside consecutive MeGlcA pattern. It is unknown if the different types of softwood xylan patterning exist on a single molecule. Again, this image is not an accurate representation of proportions of different xylan structures in softwood. Xylan branches were colour coded as described in Figure 1.8.

1.9.4 Use of glycosyl hydrolases and polysaccharide electrophoresis to probe the structure of xylan.

The intricate differences in the structure of the xylan described in the previous section were primarily discovered by digesting the polysaccharide with specific glycosyl hydrolases (GHs) and profiling the structure of the released oligosaccharides. Similarly to the techniques applied in this project, the main xylanases used to study the xylan structure belong to three GH families: GH10, GH11 and GH30. Both GH10 and GH11 can be used to estimate the degree of xylan branching as they can effectively degrade unbranched regions and are inhibited by the presence of neutral and acidic sugar decorations (Mortimer et al., 2010, Brown et al., 2009). GH30 is an enzyme which requires the presence of [Me]GlcA decorations to catalyse the digestion of the xylan backbone at position +2 towards the reducing end from the branched monomer (Urbanikova et al., 2011). Therefore, GH30 digestion can be used to profile the distance between the [Me]GlcA branches on the xylan backbone (Bromley et al., 2013). In addition to xylanases, enzymes capable of removing the xylan branches are also essential to evaluate the structure of the polysaccharide. These include α-glucuronidases from family GH115 capable of cleaving the [Me]GlcA branches from the backbone (Rogowski et al., 2014) and GH62 arabinofuranosidases which can remove the arabinose decorations (Beylot et al., 2001).

Polysaccharide Analysis by Carbohydrate gel Electrophoresis (PACE) is one of the techniques which can be used to profile the oligosaccharides released by digesting polysaccharides with GHs (Goubet et al., 2002). In this technique fluorophore labelled oligosaccharides are resolved by a polyacrylamide gel based on their size and charge (Goubet et al., 2011). Due to its negative charge GlcA contributes to oligosaccharide migration in this gel system. Therefore, xylotetraose with a [Me]GlcA branch will migrate quicker in a PACE gel than an unbranched xylotetraose (Bromley et al., 2013). For clarity a specific example of GH11 and GH30 digestions and subsequent PACE of oligosaccharides released is demonstrated on Figure 1.10

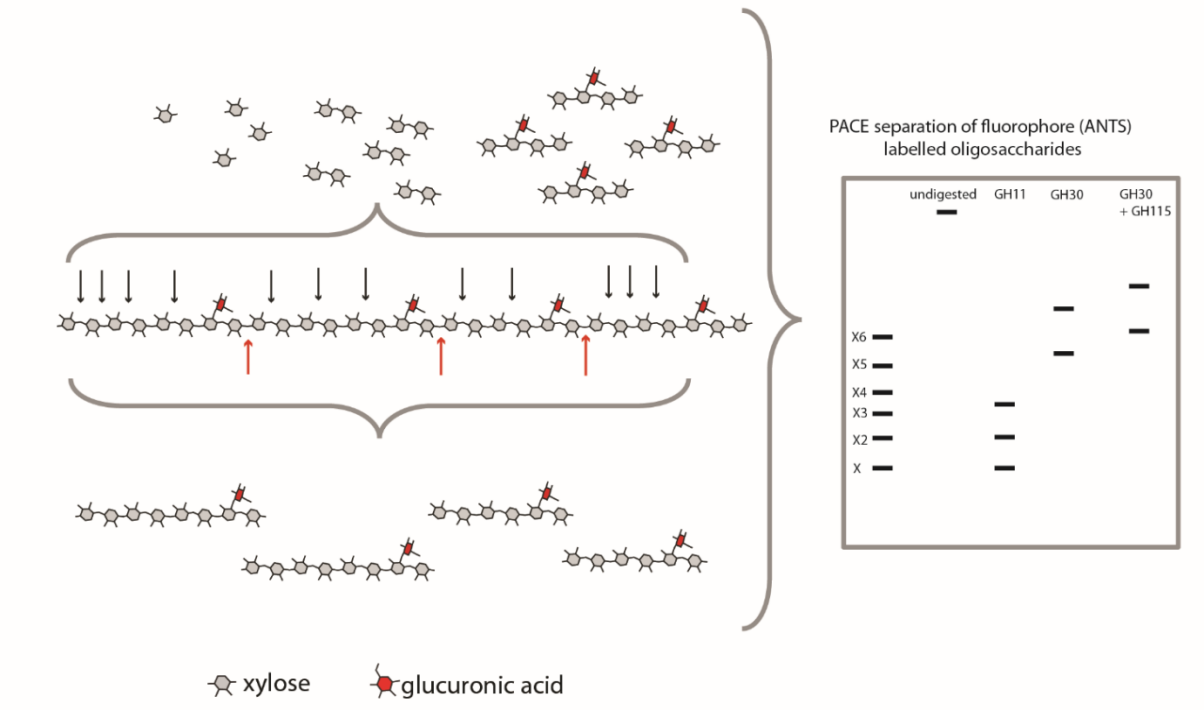


Figure 1.10 Analysis of polysaccharide structure with GH enzymes and PACE. The xylan backbone can be digested with xylanases GH11 (black arrows) or GH30 (red arrows). Digestion products can then be labelled with a fluorophore and separated on a PACE gel. Digestion products can be identified by comparing their migration to the xylooligosaccharide standard. The presence of GlcA enhances the migration rate of longer xylo-oligosaccharides and the branch can be removed with glucuronidase GH115 treatment.

1.10 Lignin

Lignin is the main non-polysaccharide component of both hardwood and softwood. Lignin is a phenolic polymer made from three main building blocks: p-coumaryl, coniferyl, and sinapyl alcohols. The lignin monomers are all derived from phenylalanine and once secreted to the cell wall they are incorporated into the polymer as guaiacyl (G), syringyl (S), and p-hydroxyphenyl (H) units in a random manner based on radical coupling. Lignin is a heterogenous polymer with cell wall localised laccases and peroxidases being responsible for the formation of monolignol radicals which then are likely to polymerise in a random fashion and forming numerous different linkages (Boerjan et al., 2003). As the process of depositing lignin in secondary cell walls, known as lignification, leads to impregnation of the xylem cell wall with a hydrophobic polymer it is often associated with apoptosis. Thus parenchyma cells, neighbouring the differentiating tracheids or vessels often supply monolignols to the cell wall

undergoing lignification (Meents et al., 2018). Monolignol composition is one notable difference between lignin present in different plant species (Figure 1.11). While hardwood contains G, S and H monolignols, softwood lignin is made nearly only from G monolignols (Vanholme et al., 2010). The process of lignification is important in the maintenance of plant biomass properties. Arabidopsis mutant plants with reduced lignin content or altered monolignol composition often have collapsed xylem vessels and can be severely dwarfed (Bonawitz and Chapple, 2010).

Figure 1.11 Diversity in lignin composition. A) Lignin composition differs across the tree of life. (*) Results for Gingko were obtained from cultured callus cells, (**) lignin-like polymers are detected in some algae and bryophytes. B) Structure of different alcohols forming the H, S and G subunits of lignin. Average proportion of each subunit is indicated for hardwood and softwood. Figure adapted after (Vanholme et al., 2010).

1.11 Cell wall proteins

In addition to polysaccharides and phenolic compounds plant cell walls contain a small but functionally important protein fraction. Three main classes of proteins can be found in plant cell walls: catalytic enzymes, glycine rich proteins (GRPs) and hydroxyproline rich glycoproteins (HRGPs) (Cosgrove, 2005, Braidwood et al., 2014).

Catalytic enzymes play an important role in cell wall remodelling (Cosgrove, 2005). Enzymes such as pectinmethylesterases or xyloglucan endotransglycosylases are involved in modification of primary cell wall polysaccharide structure (Cosgrove, 2005) while expansins may catalyse non-covalent polysaccharide interactions (Cosgrove, 2000). Laccases and peroxidases are another class of enzymes critical for wood formation. Their activity is necessary for the formation of monolignol radicals required in the process of lignification (Boerjan et al., 2003). Interestingly, the two enzyme classes play non-redundant functions in lignification of hardwood cell walls (Zhao et

al., 2013) with laccases likely being bound in the secondary cell wall layers and peroxidases localising to the middle lamella (Chou et al., 2018).

GRPs are characterised by the presence of multiple repetitive glycine rich regions in the protein sequence (Mangeon et al., 2010). It has been proposed that GRPs may be involved in the formation of cell wall by acting as a scaffold on which the polysaccharide matrix is deposited (Keller and Baumgartner, 1991). Interestingly GRPs are also localised between different layers of secondary cell walls and were proposed to play a role in connecting these structures in xylem cells of hardwoods (Ryser et al., 2004).

Three different types of proteins are classified as HRGPs: extensins, proline rich proteins and arabinogalactan proteins (AGPs). Extensins are proposed to act as a scaffold for pectic polysaccharides in primary cell walls (Lamport et al., 2011) and proline rich proteins may be involved in development of cell wall cross-links in root hairs (Bernhardt and Tierney, 2000). Arabinogalactan proteins have frequent hydroxyproline (Hyp) residues preceded by a regular amino-acid motif: Ser, Thr, Ala. The Hyp residues of AGPs are frequently glycosylated with type-II arabinogalactans (AG-II) and up to 90% of AGP mass can be derived from its glycan component (Kieliszewski, 2001). AGPs are present in softwood, particularly in timber of Larch trees, where AG-II can constitute up to 25% of compounds isolated in the process of hot water extraction (Timell and Syracuse, 1967). The backbone of the AG-II component of AGPs is made from a β-1,3-linked D-galactose units. The galactan backbone can be branched with β-1,6-linked galactan decorations which can also have α-1,3-linked arabinose side-chains. Galactan branches of AG-II may also carry a terminal β-1,6-linked GlcA (Tryfona et al., 2012). This acidic decoration can be absent or have very low abundance in larch AG-II (Timell and Syracuse, 1967, Goellner et al., 2011). One important function of AGPs in the cell wall can be to cross link its polysaccharide constituents. Indeed, Arabinoxylan Pectin Arabinogalactan Protein 1 (APAP1) is an AGP that was found to have its AG-II glycan component covalently linked to arabinoxylan and RG-I (Tan et al., 2013). The importance of this and other intermolecular interactions in the maintenance of cell wall molecular architecture is discussed in the next section of this Chapter.

1.12 Intermolecular linkages in plant cell walls

The strength and mechanical resistance of cell walls and their recalcitrance to enzymatic digestion are thought to be a result of the extensive cross-linking of cell wall components (Simmons et al., 2016). The exact chemical nature of these linkages remains largely unknown. Only recently studies utilising mostly the *A. thaliana* model have enabled identification of some of these linkages which include mostly pectin-hemicellulose, hemicellulose-cellulose and hemicellulose-lignin interactions. Loss of these interactions frequently results in significant changes in plant growth, development and biomass properties, highlighting the importance of cell wall cross-linking.

1.12.1 Pectin mediated interactions and linkages

Pectic polysaccharides are highly cross-linked. The main types of pectin-pectin interactions include formation of calcium cation mediated egg-box structures between de-methylesterified HG chains and boron mediated linkages between RG-II polysaccharides (Cosgrove, 2005). In addition to that, pectin is involved in several intermolecular interactions (Tan et al., 2013, Ralet et al., 2016).

A significant body of information on pectin synthesis and its interaction with other cell wall components has been obtained through the study of seed coat epidermis and the mucilage released from it. During its development the seed coat epidermis produces large quantities of polysaccharides which are released from the seed upon imbibition to form a structure known as mucilage (Western, 2012). Mucilage is composed predominantly of RG-I but it does contain some cellulose and hemicelluloses. Importantly, large quantities of mucilage are released from a single cell type, facilitating the study of the structure of its composite polysaccharides (North et al., 2014). Enzymatic (Macquet et al., 2007) or mechanical (Voiniciuc et al., 2013, Zhao et al., 2017) treatments allow isolation of mucilage from the seed, further facilitating study. Furthermore, changes in mucilage structure or composition frequently have no effect on Arabidopsis growth while in other tissues, such perturbations may be lethal to the organism (North et al., 2014).

Use of the mucilage model has enabled identification of direct covalent linkages between RG-I and xylan (Ralet et al., 2016). This linkage is believed to be mediated by the activity of MUCI21/ MUCILAGE MODIFIED5 (MUM5) which is a GT61 family

member suggested to be responsible for the transfer of xylose side chains onto the xylan backbone (Voiniciuc et al., 2015a).Interestingly, mucilage RG-I was observed to carry a xylose on the GalA of the backbone, connecting RG-I and xylan through a mutual side-chain. Thus, MUM5 may be generating a direct covalent linkage between RG-I and xylan (Ralet et al., 2016). The presence of this MUM5 dependant xylose was demonstrated to be essential for RG-I cellulose interaction, suggesting that pectin-cellulose binding is actually mediated by xylan in mucilage (Ralet et al., 2016). This is further supported by the fact that mutations in genes coding for proteins involved in both cellulose and xylan synthesis result, similarly to phenotypes observed in *mum5* plants, in loss of mucilage adherence (Voiniciuc et al., 2015a, Griffiths and North, 2017). In addition to direct pectin-xylan interaction an in-direct binding between these two polysaccharides was also proposed (Tan et al., 2013). As indicated previously, APAP1 AGP may contain both xylan and RG-I bound to its arabinogalactan component. This AGP mediated interaction is therefore likely to be a second possible mode for formation of covalent linkages between RG-I and xylan.

Pectin mediated intermolecular cross-linking may be important for hardwood development and properties. Recent studies using poplar, amongst a range of other model plant species, has indicated that mutations in Galacturonosyltransferase 4 (GAUT4) result in a significant increase in plant growth (Figure 1.12A) (Biswal et al., 2018). The presented experiments indicated that GAUT4 is responsible for HG biosynthesis and that loss of its activity may result in decreased levels of cell wall cross-linking. The work suggested that a reduction in the number of HG dependant egg-box cross-links may result in knock-on effects on the entire cell wall molecular architecture and as a result it may allow for greater plant growth. If the presented hypothesis is correct, it is possible that pectic polysaccharides may play an important function in the maintenance of cell wall molecular architecture. Assuming existence of covalent connections between different types of pectin (Harholt et al., 2010) it is possible that alterations to galacturonan biosynthesis observed in GAUT4 mutant plants can have knock-on effects on different types of pectin. It is proposed that in some tissues RG-I can be covalently cross-linked to xylan (). It is unknown if these or other pectin-xylan interaction exists in wood. However, due to changes in pectin biosynthesis, these putative linkages between xylan and pectin may be disrupted in GAUT4 mutant leading to reduction in xylan mediated pectin-cellulose interaction.

Model presented by Biswal et al., 2018 indicates that these interactions may contribute to the maintenance of wood properties and control of plant growth (Figure 1.12B).

Figure 1.12 Proposed pectin mediated intermolecular interactions in hardwood. A) Mutation of *GAUT4* significantly increases growth of poplar B) Model of pectin mediated cell wall cross-linking: (i) HG forms egg-box structures via calcium cations (+), these can connect homogenous HG molecules with these linked to other pectic polysaccharides (ii). HG may form part of one molecule with RG-I (iii). RG-I may then link to xylan, either via arabino-galactans or direct covalent linkage. Xylan can bind by hydrogen bonding (H) to the hydrophilic surface of the cellulose fibril (iv).HG may also be linked to RG-II which can cross-link via boron anions (v, -). Figure adapted after (Biswal et al., 2018).

1.12.2 Xylan – cellulose interaction

Xylan has been proposed to interact with cellulose (Cosgrove and Jarvis, 2012) but until recently little was known about the exact nature of this binding. *In vitro* experiments demonstrated that arabinoxylans can bind and cross-link isolated bacterial cellulose fibrils (Mikkelsen et al., 2015) but how this was achieved remained unknown. Xylan is a highly branched polysaccharide and in solution it is known to have a three-fold screw conformation (Nieduszynski and Marchessault, 1971). These features, don't seem compatible with binding two-fold glucans in the cellulose fibril, thus it was challenging to explain how this hemicellulose may associate with the microfibrils in a manner different to aggregation with its hydrophobic surface (Kabel et al., 2007, Bosmans et al., 2014).

Early insights into the specific mode of xylan- cellulose interaction came from analysis of the branching patterns on acetylated hardwood xylan (Bromley et al., 2013, Busse-Wicher et al., 2014). As previously discussed, xylan molecules were discovered to be constructed from two types of domains termed compatible and incompatible (Busse-Wicher et al., 2016a). The compatible domain forms the majority of the xylan molecule and has a regular even spacing of [Me]GlcA branches with the dominant distances between the acidic decorations being equal to 8, 10 and 12 xylose units. In the *A. thaliana* model, this even pattern of [Me]GlcA substitutions is established by the activity of the GUX1 enzyme and accounts for ~70% of the total xylan glucuronidation. The remaining 30% of xylan glucuronidation is added, in secondary cell walls of *A. thaliana*, by the GUX2 enzyme and forms what is known as the incompatible domain. In these regions the [Me]GlcA branches are more densely packed and the dominant spacing between the acidic decorations is equal to 5, 6 and 7 xylose monomers. The discovery of xylan domains was followed by the description of the acetylation pattern of *A. thaliana* xylan (Busse-Wicher et al., 2014). The analysis of acetate distribution on the xylan polymer extracted from *gux1/2* plants, lacking secondary cell wall glucuronidation, revealed that acetate branches are present predominantly on every other xylose monomer. It was hypothesised that the even spacing of acetyl branches and [Me]GlcA branches in the compatible domain may generate a free surface on a xylan in the two-fold conformation which may allow it to bind to the hydrophilic surface of the cellulose microfibril (Figure 1.13). Molecular dynamics simulations support the hypothesis that a stable interaction between patterned acetylxylan and the hydrophilic surface of the cellulose microfibril is feasible (Busse-Wicher et al., 2014).

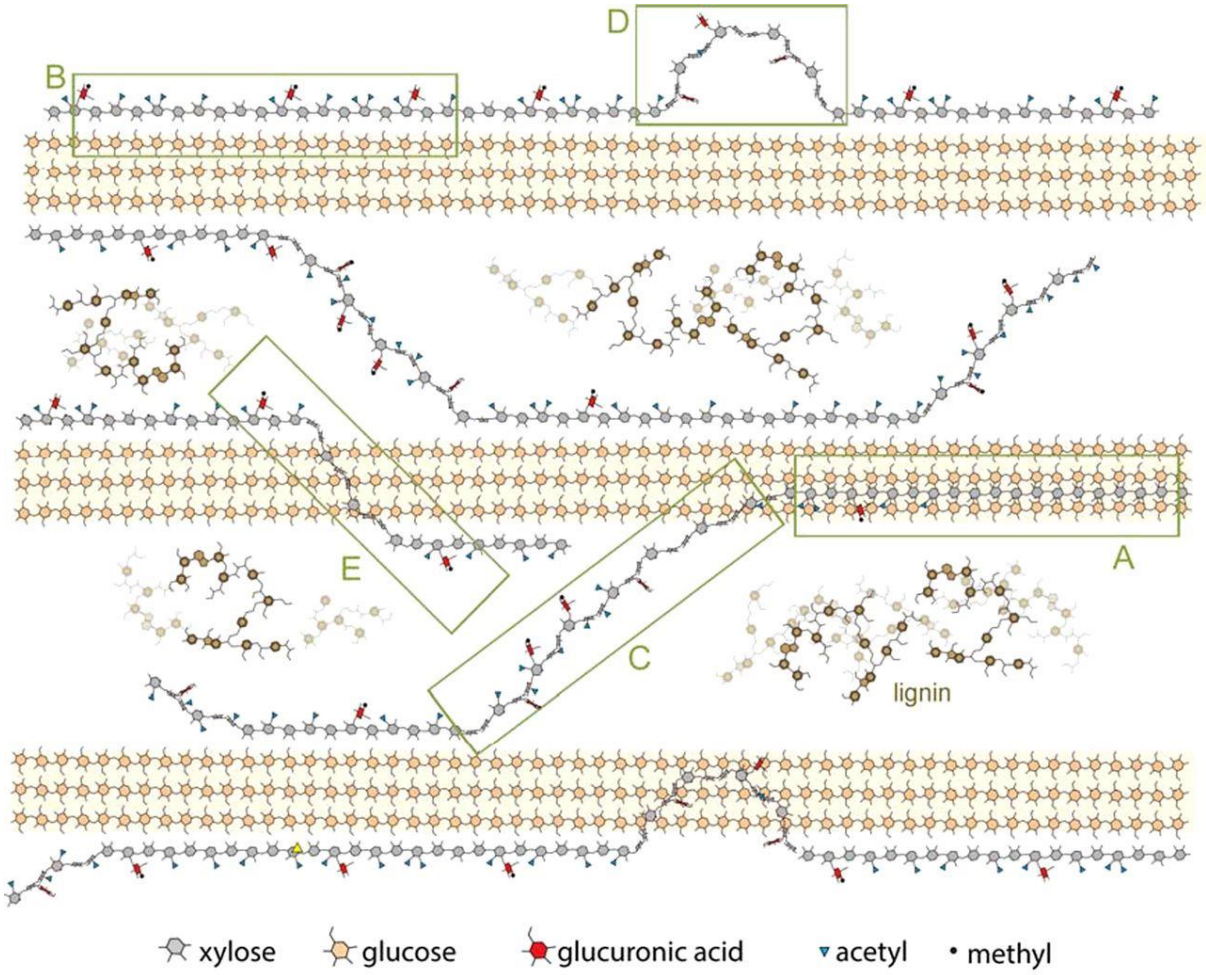


Figure 1.13 Hypothetical model of xylan – cellulose interaction in secondary cell walls of eudicots. A) Top view of xylan binding to the hydrophilic surface of the cellulose microfibril. B) In a two-fold screw conformation the compatible domain of xylan can bind, via hydrogen bonding, the hydrophilic surface of the cellulose microfibril. The incompatible domain may be responsible for cross-linking (C) different cellulose microfibrils. Incompatible xylan may also form loops in the proximity of the fibril (D) which may lead to association of a subsequent compatible domain with a different hydrophilic surface of the same microfibril (E). Figure presented after (Busse-Wicher et al., 2016a).

Further evidence supporting the binding of the two-fold screw xylan to the hydrophilic surface of the cellulose microfibril came from solid state NMR experiments looking at the arrangement of polysaccharides in the cell walls of WT and mutant *A. thaliana* plants. Early experiments using dried plant stems indicated that peaks associated with the xylan molecule exhibit a range of chemical shifts on CP-INADEQUATE spectra, indicating the possible presence of multiple xylan folds in the cell wall (Dupree et al., 2015). To increase the resolution of the spectra, never dried, hydrated material was used in subsequent solid state NMR experiments. The comparative analysis of never dried WT and *irx3 A. thaliana* biomass indicated that the existence of one of the two xylan conformations is dependent on the presence of cellulose in plant secondary cell walls (Simmons et al., 2016). Specifically, the peaks associated with the two-fold screw conformation of the xylan backbone were absent in *irx3* plants, which lack secondary cell wall cellulose. The same two-fold screw xylan peaks formed the majority of the xylan signal in the WT biomass, while signals associated with three-fold xylan conformation formed a smaller proportion of the total xylan signal (Simmons et al., 2016). Importantly analysis of the spatial distribution of polysaccharides indicated that the two-fold screw xylan molecule is in close proximity (<1nm) to the cellulose fibril.

These experiments did not indicate if the xylan molecule, in a two-fold screw conformation, is bound to the hydrophilic or hydrophobic surface of the cellulose microfibril. Analysis of the structure of xylan in the *A. thaliana esk1* mutant indicated that in addition to the reduction in xylan acetylation, caused by the absence of active ESK1 acetyltransferase, the pattern of xylan [Me]GlcA branches is altered on *esk1* xylan (Grantham et al., 2017). As the majority of [Me]GlcA on *esk1* xylan is incompatibly patterned, with oddly spaced [Me]GlcA dominating over evenly spaced ones, the biomass enabled the investigation of the importance of xylan patterning for the xylan-cellulose interaction. Two-fold xylan signals were absent from the *esk1* solid-state NMR spectra. As the even spacing of xylan branches is only required for interaction of xylan with the hydrophilic surface of the cellulose fibril the lack of two-fold signals may indicated that the xylan-cellulose interaction is likely to occur by the association of the compatible domains of xylan with the hydrophilic surface of the cellulose microfibril (Figure 1.14) (Grantham et al., 2017). Alternatively, part of the reason for the loss of two-fold xylan signal may be associated with aggregation of

xylan molecules in *esk1* plants due to potential reduction in the solubility of a polysaccharide with reduced number of branches (Xiong et al., 2015).

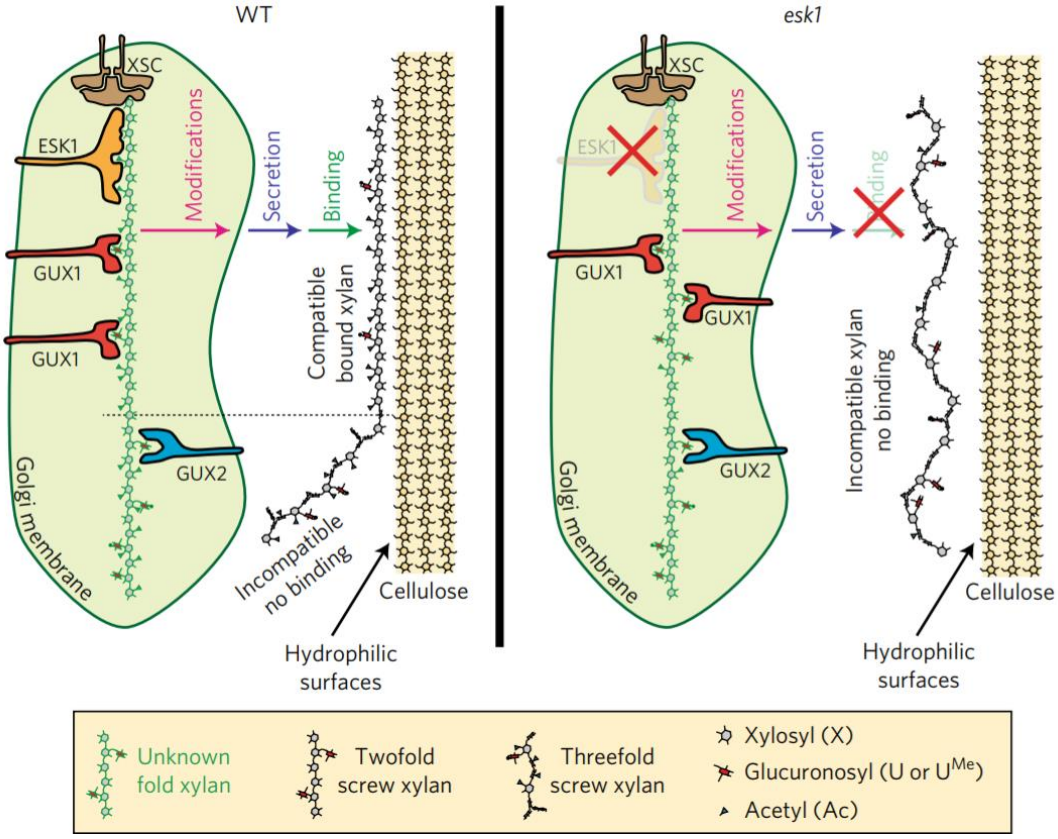


Figure 1.14 Role of ESK1 in the synthesis of the compatible domain of xylan. In WT plants (left) ESK1, GUX1 and GUX2 act together to generate xylan molecules which have compatible and incompatible regions. The even spacing of [Me]GlcA in compatible regions is established by GUX1 and allows for hydrogen bonding between xylan and the cellulose microfibril. Once bound to the cellulose fibril xylan maintains the two-fold screw conformation over the majority of the polymer. In *esk1* plants (right) the absence of acetylation added by the activity of ESK1 leads to deregulation of GUX1 activity. This causes formation of largely incompatible xylan molecules which cannot interact with the cellulose microfibril. The resulting xylan molecule has a three-fold screw conformation in the cell wall. Figure presented after (Grantham et al., 2017).

Importantly the binding of two-fold screw xylan to the hydrophilic surface of the cellulose microfibril is likely to extend beyond the model hardwood plant species of *A. thaliana.* The analysis of softwood xylan structure indicates that the majority of arabinose and MeGlcA branches are placed with an even spacing. This conservation of xylan patterning across gymnosperms and angiosperms suggests that softwood xylan, similarly to the hardwood one, is likely to bind to the hydrophilic surface of the cellulose microfibril as a two-fold screw (Busse-Wicher et al., 2016b). Indeed, analysis of distances between individual cellulose microfibrils in spruce wood indicates that they are divided by a gap which can accommodate a single chain of two-fold screw xylan bound to the hydrophilic surface of one of the microfibrils (Fernandes et al., 2011, Jarvis, 2018).

1.12.3 Xylan – lignin interactions

Interactions between xylan and lignin have been proposed to be mediated by covalent linkages between two different types of xylan branches. The first one, in which ferulic acid present on arabinose decorations of grass xylan forms covalent linkages with lignin monolignols is well described in the literature (Scheller and Ulvskov, 2010). Much less direct evidence supports the presence of the second type of linkage in which the [Me]GlcA branches of xylan are cross-linked to lignin (Giummarella and Lawoko, 2016). Nonetheless, published data together with results presented in this thesis indicate that both these types of linkages may contribute to the formation of lignin-carbohydrate complexes (LCCs) and that these LCCs might have a role in the maintenance of cell wall molecular architecture and recalcitrance.

Ferulic acid is esterified to O-5 of some of the arabinofuranosyl residues of grass xylan (Scheller and Ulvskov, 2010). Once in the cell wall the ferulic acid moieties can form dimers, leading to cross-linking of xylan molecules, or they can be co-polymerised into lignin (Ralph et al., 1995, Grabber, 2005). The presence of these ferulate mediated cross-links is likely to be important for the maintenance of cell wall resistance to enzymatic degradation. In mutant rice (Chiniquy et al., 2012) and Setaria (de Souza et al., 2018) with reduced xylan ferulation levels, the amount of monosaccharides released in the process of saccharification is increased.

GlcA is proposed to form a LCC by a direct ester bond to the hydroxyl of lignin γ carbons (Figure 1.15) (Du et al., 2014). Evidence for the presence of GlcA – lignin

ester linkages was first presented in softwood (Watanabe and Koshijima, 1988) but the treatment used in this work might have led to artefactual LCC formation (Giummarella and Lawoko, 2016). Different extraction protocols followed by NMR analysis have been used to provide indication that the GlcA-lignin LCC might exist in hardwood (Giummarella and Lawoko, 2016) and in softwood (Oinonen et al., 2015) but the NMR evidence supporting its existence is inconclusive due to the complex spectra of the wood extracts. The existence of the lignin– GlcA LCC is further supported by the presence of a specific enzyme class, known as glucuronyl esterases, which belong to carbohydrate esterase family 15 and are predicted to be involved in cleavage of this link during fungal and bacterial wood degradation (Monrad et al., 2018). Purified glucuronyl esterases have activity towards model GlcA-lignin ester compounds *in vitro* (Spanikova and Biely, 2006, Biely et al., 2015) and form part of industrial enzymatic cocktails used in biomass degradation (d'Errico et al., 2016). Results presented in this thesis and published as a direct result of the work described in it (Lyczakowski et al., 2017) indicate that the presence of GlcA branches on xylan is important for the maintenance of hardwood biomass recalcitrance. As discussed here and further in the Thesis text it is likely that this is mediated by formation of GlcA-lignin LCCs.

Figure 1.15 Proposed structure of the GlcA – lignin ester. Figure presented from (Du et al., 2014)

1.12.4 Mannan mediated cell wall cross-linking

The mannan class of polysaccharides may be bound to cellulose fibrils and it also may form linkages to lignin. Mutations in CSLA2 (Yu et al., 2014) and MUCI10/MAGT1 (Voiniciuc et al., 2015b) lead to the reduction in mucilage galactoglucomannan levels and result in loss of mucilage adherence. A reduction in mucilage GGM levels is also associated with changes in cellulose structure (Voiniciuc et al., 2015b) suggesting the hemicellulose may coat the cellulose microfibril. Further structural analysis demonstrated that mucilage GGM has a specific patterned structure made from a repeating glucose-mannose disaccharide with frequent galactose decorations (Yu et al., 2018). Therefore, similarly to the compatible regions of xylan, the mucilage GGM in two fold screw conformation has a branch-free surface which may enable its interaction with the cellulose microfibril (Yu et al., 2018). In addition to the analysis of mucilage cell walls, GGM was demonstrated to bind and cross-link bacterial cellulose fibrils *in vitro* (Whitney et al., 1998). GGM was also proposed to bind the cellulose fibril in softwood (Terashima et al., 2009) but too little is known about the exact structure of softwood GGM to determine the molecular details of any interaction.

In addition to interacting with the cellulose microfibril softwood GGM was also proposed to form covalent linkages with lignin. Indeed, recently published work has used NMR to provided direct evidence for the existence of α ether linkages between lignin and the GGM backbone in pine wood (Nishimura et al., 2018).

1.13 Rationale and research aims

Work presented in this thesis focuses predominantly on identifying the genes involved in addition of MeGlcA branches to softwood xylan and the possible role of MeGlcA branches and their pattern in the maintenance of plant cell wall properties. Coniferous forests have significant global importance. Recent estimates suggests that more than 75% of planted trees are conifers (Ramage et al., 2017). In countries where the majority of forest planting occurs in temperate or boreal climates this proportion is even higher. For example, in the USA as much as 98% of planted trees represent members of the coniferophyta division (Sedjo, 2001). This dominance of coniferous plants in temperate and boreal regions is primarily due to their better resistance to soil drought, a particularly significant stress factor during cold winter months (Brodribb et al., 2014).

As outlined in this introduction chapter, despite significant progress in the understanding of the cell wall biosynthesis process, significant gaps remain, especially in the knowledge of the assembly of cell components to form a biomaterial such as wood. This lack of understanding is particularly significant for commercially important softwoods, where a small amount of accessible genomic information, among other factors, has hampered progress in elucidating cell wall biosynthesis enzymes and their role in formation of the cell wall matrix. To address these challenges this thesis aims to combine various experimental techniques to investigate the process of softwood xylan biosynthesis. It is important to consider that a significant proportion of the work will be performed using model hardwood species *A. thaliana* to assay the activity of softwood glycosyltransferases.

Arabidopsis is an important plant model system used to evaluate gene function and was the first plant to have its genome fully sequenced (Arabidopsis-Genome-Initiative, 2000). The short seed-to-seed rotation period established Arabidopsis as an important model system in plant cell wall research and with the genomic information and seed-banks of mutant germplasm available the plant was used to identify multiple genes responsible for the biosynthesis of cell wall components (Liepman et al., 2010). The particularly important class of phenotypes, the irregular xylem, which is associated with vessel collapse was described for the first time in Arabidopsis cellulose mutants (Turner and Somerville, 1997) and is still a common feature evaluated when analysing plant cell wall biosynthesis. Importantly, the modifications in secondary cell wall structure in Arabidopsis, even these associated with vessel collapse, do not lead to significant transcriptional changes in the plant (Faria-Blanc et al., 2018). This indicates lack of secondary cell wall integrity signalling pathway in Arabidopsis and enables the study of cell wall biosynthesis largely in isolation from other processes. This factor makes Arabidopsis a good test platform to evaluate the activity of heterologously expressed glycosyl transferases generating non-native and sometimes non-functional polysaccharide structures.

To investigate the largely unknown process of cell wall cross-linking the first aim of this PhD project will be to evaluate the importance of xylan branching for wall integrity. This analysis will be performed using saccharification and extractability experiments on WT and mutant plants lacking specific types of xylan branching.

Secondly, the project will focus on evaluating the process of softwood xylan biosynthesis with a particular interest in the GUX enzymes and the type of GlcA patterning they may be responsible for establishing. Reports presented in Introduction section 1.9.3 indicate that softwood xylan has two distinct types of patterning which, according to the proposed cell wall molecular architecture model (Busse-Wicher et al., 2016a), may facilitate or prevent interaction between xylan and the hydrophilic surface of the cellulose fibril. Therefore, the biosynthesis of these structures may be important for the maintenance of polysaccharide interactions in softwood. The study of softwood GUX enzymes will be performed using *in vitro* and *in vivo* assays. These will be greatly facilitated by recent progress in the development of synthetic biology tools enabling heterologous protein expression and generation of genetic constructs which are described in Chapter 2.

Finally, the project will also aim to investigate the role of the xylan patterning and xylan-cellulose interaction in the maintenance of cell wall molecular architecture and function. Xylan cellulose interaction may be controlled by different patterns of wood xylan branching and it is important to understand how this interaction contributes to the maintenance of wood properties. This analysis will be performed using the Arabidopsis model in which both hardwood and softwood GUX enzymes will be used to generate novel xylan structures, some of which may not be able to bind the hydrophilic surface of the cellulose fibril. The impact of these modifications will be evaluated using a range of techniques including solid state NMR which can provide information about conformation of polysaccharides and their proximity in the modified cell walls.

Chapter 2: Materials and Methods

All reagents used in the described protocols were purchased from Sigma-Aldrich unless indicated otherwise.

2.1 Plant material growth and AIR preparation

2.1.1 Plant material used:

Arabidopsis thaliana plants of the Columbia-0 ecotype were grown in a cabinet maintained at 21 °C, with a 16-h light, 8-h dark photoperiod. Mutant insertion lines described in published work were used in this study. Specifically, *gux1/2* (Mortimer et al., 2010), *gxm1/2/3* (Cornuault et al., 2015) and *esk1-5* (Grantham et al., 2017) was used in saccharification experiments. In addition to that, *gux1/2/3* (Mortimer et al., 2015) and *esk1-5:kak-8* (Bensussan et al., 2015, Grantham et al., 2017) plants were used for transformation experiments.

Both *Picea glauca* and *Picea abies* softwood was acquired from 30-50cm tall potted plants purchased from Scotsdale (Great Shelford, Cambridgeshire). *Gnetum montanum* stem material was obtained from the University of Cambridge Botanical Garden.

2.1.2 Preparation of Alcohol Insoluble Residues (AIR):

Plant stems (25 – 50 mg wet weight) were homogenised in ethanol absolute by ball milling with steel balls in a Teflon vessel at 20 Hz for 5 minutes. The milling cycle was repeated four times with 5 minute intervals between each milling step. Homogenised material was washed with ethanol absolute and incubated over-night in a 2:3 (v:v) mixture of methanol and chloroform. The following day, the methanol: chloroform wash was repeated for 1h in a fresh solution. Thereafter, the sample underwent consecutive wash steps in ethanol absolute, 60% ethanol (v/v), 80% ethanol (v/v) and ethanol absolute. Following the final wash, a pellet of alcohol insoluble residues (AIR) was dried in an oven set at 50°C for 48 hours.

2.2 Analysis of xylan structure

2.2.1 AIR digestion with xylanase enzymes

For each enzymatic digestion 0.5 mg of AIR material was used. Prior to application of xylanase enzymes AIR were deacetylated with an alkali treatment (20 μL of 4 M NaOH applied for 1 hour at room temperature). Following alkali treatment the solution was neutralised with 1M HCl and re-suspended in 500 μL of 0.1 M ammonium-acetate buffer with a pH equal to 5.5.

De-acetylated AIR was digested with *Neocallimastix patriciarum* GH11 enzyme overnight at room temperature by amending the suspension with 5 μL of 200 to 500 μM enzyme stock. These conditions achieved complete digestion of accessible xylan to xylose, xylobiose and glucuronidated oligosaccharides. Purified *Np*GH11 was a kind gift from Prof. Harry Gilbert (Newcastle University, United Kingdom). For GH30 digestion 1 μL of 1 mg/mL purified *Erwinia chrysanthemi* enzyme was applied onto de-acetylated AIR for 1 hour at room temperature. *Ec*GH30 used in this work was a kind gift from Dr Kristian Krogh (Novozymes A/S, Denmark). GH115 glucuronidase used in this work was also received from Dr Kristian Krogh. For de-glucuronidation, 1 μL of 1 mg/mL GH115 was added to oligosaccharides re-suspended in 0.1 M ammonium acetate pH = 5.5 buffer and incubated over-night at 55°C. Prior to GH115 treatment, the GH11 enzyme was heat inactivated by incubating the oligosaccharide sample at 120°C for 10 minutes.

2.2.2 Polysaccharide Analysis by Carbohydrate Gel Electrophoresis (PACE)

For the PACE analysis, oligosaccharides released with a xylanase treatment were dried *in vacuo* and derivatised with 8-aminonapthalene-1,3,6-trisulphonic acid (ANTS; Invitrogen). The derivatisation process was performed by reductive amination of oligosaccharides released from 0.5 mg of AIR in acidified DMSO (0.3 μL glacial acetic acid, 3 μL water and 3.3 μL DMSO) in the presence of ANTS (3.3 μL of 0.2 M solution in DMSO) at 37°C over-night. Following ANTS derivatisation the solution was dried *in vacuo* and re-suspended in 100 μL of 6 M urea.

PACE gel casting and running was performed following a customised protocol. A single gel mix was poured into a pre-assembled glass gel mould. The gel mix was composed from water (20 mL), 1M Tris-Borate buffer pH = 8.2 (3 mL), 40% acrylamide solution (29:1 feed ratio, 8.4 mL), 10% ammonium persulphate (0.18 mL) and 1,2-Bis(dimethylamino)ethane (TEMED, 0.04 mL). Following casting the gel was cross-linked at 4°C over-night prior to the run. Gel running was performed in 0.1M Tris-Borate pH = 8.2 buffer at 1000 V for 45 minutes. Sample amount was customised to allow for adequate band intensity and varied between 1 and 5 μL of urea re-suspended ANTS labelled oligosaccharides. Gel visualisation was performed using a G:Box UV transluminator (Syngene). Band intensities were quantified using a volume integration script built into ImageJ software (National Institute of Health).

2.2.3 Mass-spectrometry analysis of xylooligosaccharides

Oligosaccharides resulting from an enzymatic digestion of 0.5 mg of AIR were purified using HyperSep^{TM} Hypercarb^{TM} cartridge (Thermo-Fisher), dried and derivatised with 2-aminobenzoic acid (2-AA). For this reductive amination, 2-AA (6 mg) was re-suspended in 100 μL of 2:7 (v:v) mix of acetic acid and DMSO. This suspension was combined with 6 mg of $NaCNBH_4$ and mixed by pipetting until fully dissolved. A 10 μL aliquot of the resulting mix was used to label lyophilised oligosaccharides at 65 °C for 3 hours. Following labelling, excess 2-AA was removed with a GlycoClean S cartridge (ProZyme) following manufacturer's instructions. Released labelled oligosaccharides were dried *in vacuo* and re-suspended in 20 μl water. 1 μl of the oligosaccharide solution was mixed with an equal volume of 20 mg/mL 2,5-dihydroxybenzoic acid (DHB) in 50% acetonitrile, 0.1% TFA on a SCOUT-MTP 384 target plate (Bruker). The spotted samples were dried in a vacuum desiccator. The samples were analysed by mass spectrometry on an Ultraflex III matrix-assisted laser desorption ionization-time of flight/time of flight (MALDI/TOF-TOF) instrument (Bruker). The data were collected with a 2-kHz smartbeam-II laser and acquired on reflector mode, the mass range was 300-3000 Da. Data acquisition and analysis was performed in FlexControl (Bruker) and FlexAnalysis (Bruker) software respectively. Fragmentation analysis was performed using protocols built into the FlexControl software. On average, about 10,000 shots were used to obtain the spectra. MS experiments were performed with great help from Dr. Li Yu.

2.3 Analysis of *A. thaliana* biomass recalcitrance

2.3.1 AIR saccharification and determination of monosaccharide yield

Cellic® CTec2 (Novozymes A/S) was used for all saccharification experiments. Enzyme stock (35 µL) was diluted to a total volume of 2.5 mL with 0.1M ammonium acetate pH = 5.0 buffer. The enzyme sample was cleared from residual sugars using PD-10 desalting column (GE Healthcare) and eluted using 3.5 mL buffer, generating 1:100 (v/v) Cellic® CTec2 solution. AIR aliquots (1 mg) were homogenised in 1 mL of buffer. Homogenised AIR was amended with 25 µL 1:100 Cellic® CTec2 working solution and incubated for 24 h at 45 °C with 1400 rpm applied for 30 s every 4 min. The reaction was terminated by heat-treating the suspension at 100 °C for 10 min. D-glucose and D-xylose release from the biomass was quantified using commercial enzymatic kits (Megazyme, catalogue codes: K-XYLOSE and K-GLUHK-220A). Sugar concentration for each experiment was standardised with readings obtained from biomass and enzyme only controls.

Glucuronidase digestion was performed by incubating the products of Cellic® CTec2 saccharification with 10 µL of 1 mg/mL *Bacteroides ovatus* GH115d (Bo_03449).

2.3.2 Simultaneous saccharification and co-fermentation (SSF) experiments

50 mg dried stems of WT and *gux1/2* plants were used for each fermentation reaction. Stems were ball milled in 7 mL LB medium as described for AIR preparation. To fully recover the biomass, the vessel was washed with further 2.5 mL LB. The stem suspension was sterilised by heat treatment at 85 °C for 10 min followed by cooling on ice. Each fermentation reaction was amended with 250 µL 1:10 Cellic® CTec2 solution and 250 µL of TOP10 *E. coli* inoculum bearing the BBa_K1122676 BioBrick (OD600 of the inoculum was within 0.55–0.6 range). BBa_K1122676 encodes a Pyruvate decarboxylase and Alcohol dehydrogenase from *Zymonomas mobilis* which allow ethanol production in *E. coli* (Lewicka et al., 2014). Biomass only reactions were supplemented with 250 µL of CTec buffer and the bacterial inoculum. The plasmid was maintained by provision of 25 µg/mL Chloramphenicol (Duchefa Biochemie). The simultaneous saccharification and fermentation reactions were carried out for 96 h. at 37 °C and 200 rpm. Fermentation vessel was kept air-tight throughout the experiment. Ethanol levels were analysed using a commercial kit (Megazyme, catalogue code: K-ETOH).

2.4 Molecular cloning and generation of transgenic *A. thaliana* lines

Purification of bacterial DNA and PCR products was performed using kits manufactured by Qiagen (Hilden, Germany). The DNA sequencing was performed by the sequencing facility at the Department of Biochemistry, University of Cambridge.

2.4.1 Construct preparation – Gateway

DNA construct required for the over-expression of PgGUX1 in *gux1/2/3* background was prepared using Gateway technology (Katzen, 2007). The gene encoding the *P. glauca* enzyme with a 3xMyc C-terminal tag was synthesised by GenScript (Nanjing, China). Gateway cloning was used to insert the gene into the p3KC binary vector (Atanassov et al., 2009). Protein expression was driven by a 1.7 kbp promoter sequence of *A. thaliana IRX3* gene which forms part of the p3KC vector. The construct preparation was done by Dr Marta Busse-Wicher prior to the start of this PhD project.

2.4.2 Construct preparation – GoldenGate

For all other over-expression experiments GoldenGate MoClo technology was used to assemble binary vectors. In this technique L2 binary vectors are assembled from L1 Transcriptional units which are made from L0 parts such as promoters, coding sequences, tags and terminators (Weber et al., 2011, Patron et al., 2015). GoldenGate assembly relies on the use of Type IIS restriction enzymes together with T4 DNA ligase to digest and ligate L0 modules (or L1 transcriptional units) in one pot reaction. The order of assembled units is guided by the fact that specific linkers, designed during part synthesis, are released upon BsaI (L0 modules) or BpiI (L1 transcriptional units) digestion. These form sticky ends which ligate with the backbone donor provided.

Specifically, 100 ng of each L0 module needed (or L1 transcriptional unit) was mixed together with 100 ng of backbone donor, 1 μL of BsaI (or BpiI) enzyme (NEB), 1 μL of T4 DNA ligase (NEB), 0.15 μL BSA protein (NEB), 1.5 μL x10 T4 DNA ligase buffer with ATP (NEB) and diluted up to 15 μL in DNAse free water.

This mix was incubated in a thermocycler in a following program:

25 cycles of: 37 °C for 3 minutes followed by 16 °C for 4 minutes

50 °C for 5 minutes

80 °C for 5 minutes

2 μL of the resulting product were used for NEB 5α *E. coli* transformation.

For sub-cloning and plasmid preparation 50 μL aliquots of NEB 5α competent *E. coli* cells were used. The transformation was performed by incubating the cells with plasmid DNA (100 – 200 ng) for 30 minutes on ice. This was followed by a heat shock at 42 °C carried out for 30 seconds. After 2 minutes of recovery on ice the cells suspension was amended with 450 μL of LB medium. After 1 hour of outgrowth at 37 °C the suspension was spun at 10000 RCF for 2 minutes and the cell pellet was plated on LB agar amended with adequate antibiotic. Table 2.1 summarises the types of antibiotics used for selection of different GoldenGate assembly components.

Table 2.1. Antibiotics used for selection of transformed bacterial cells

Antibiotic used:	**Antibiotic concentration in the medium:**	**Used for selection of:**
Spectinomycin	50 μg / mL	L0 modules
Ampicillin	100 μg / mL	L1 transcriptional units and binary partner for AGL-1
Kanamycin	50 μg / mL	L2 binary vectors and pEAQ
Gentamycin	50 μg / mL	Binary partner for GV3101

Transcriptional units made were composed from promoter modules (p), coding sequences, tags and terminators. All binary vectors used for *A. thaliana* transformation were composed from three distinct L1 transcriptional units:

Position 1: p35S – Kanamycin resistance cassette – 35S terminator

Position 2: p*IRX3* – GUX – 3xMyc – Nos terminator

Position 3: p*Oleosin* – OLE1 – eGFP – *A. thaliana* Actin2 terminator

Unit used at position two varied according to the GUX enzyme over-expressed. Both position one and position three transcriptional units were used to screen for transgenic *A. thaliana* plants as described in section 2.4.3. The p*IRX3* sequence used in these

assembly reactions was the same as one described by (Atanassov et al., 2009). Standard promoter and terminator encoding L0 modules were a kind gift from Andrew Breakspear and Dr Christian Rogers from John Innes Centre (Norwich, UK). Representative L2 binary vector used for *A. thaliana* transformation generated using GoldenGate assembly is presented on Figure 2.1.

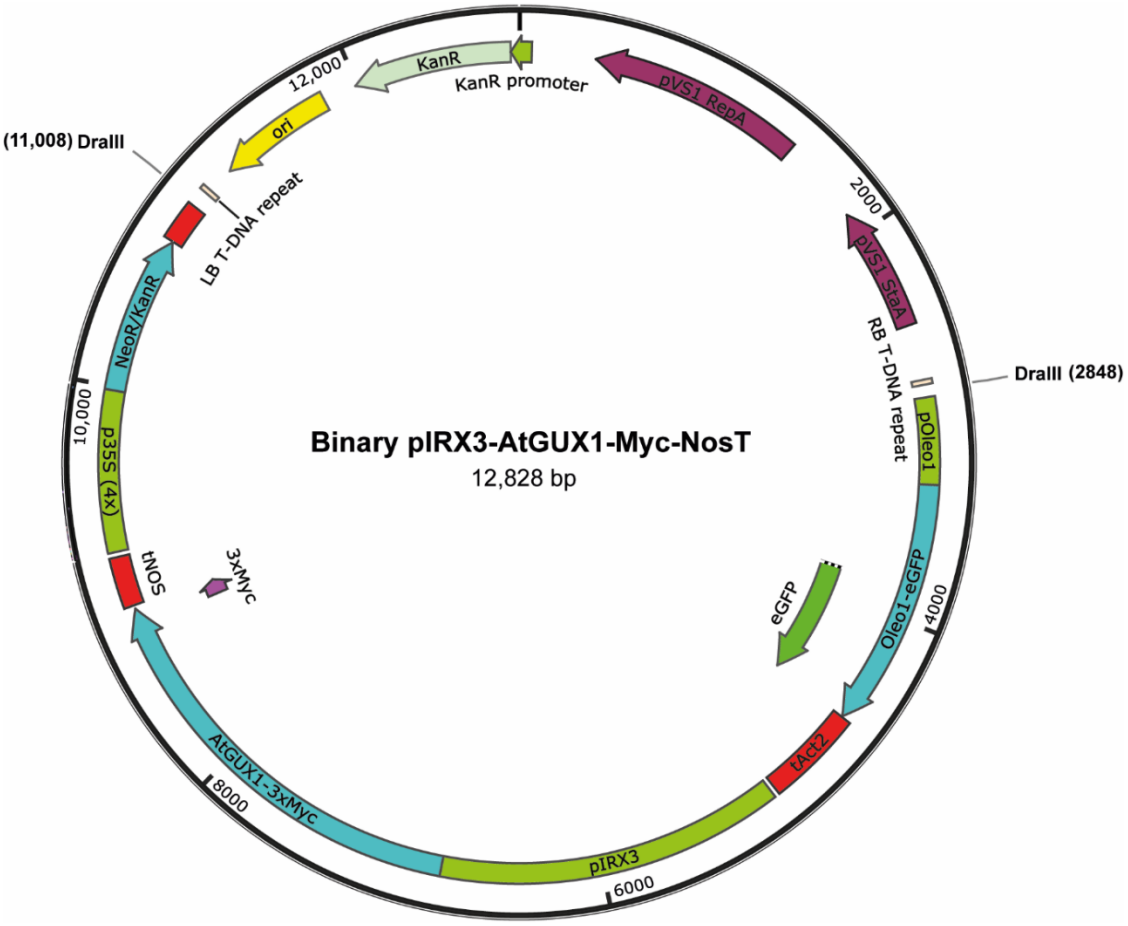


Figure 2.1 Overview of a representative GoldenGate binary vector used for plant transformations. Plant promoter (p) regions were marked in light green, coding sequences in blue and terminator regions (t) in red. This figure illustrates the vector used for over-expression of AtGUX1 under the control of p*IRX3*. DraIII restriction sites marked can be used to release the entire sequence integrated into plant genome.

2.4.3 Agrobacterium tumefaciens and Arabidopsis thaliana transformation

GV3101 strain was used for Arabidopsis transformation and AGL-1 was used for expression in *N. benthamiana*. The same protocol was used for preparation of chemically competent cells for both strains. For this, a single colony of *A. tumefaciens* was picked and grown in LB (5mL) over-night at 30 °C in the presence of an antibiotic required for the maintenance of the binary partner plasmid. The following day this starter culture was used to inoculate 50 mL LB (with antibiotic). This was grown at 30 °C until OD_{600} reached ~0.6. Thereafter, the culture was chilled on ice, spun at 3000 RCF for 5 minutes at 4 °C and the pellet was re-suspended in 1 mL 20 mM $CaCl_2$. This suspension was aliquoted (100 μL / aliquot) and stored at -80 °C until needed.

For *A. tumefaciens* transformation an aliquot of chemically competent cells was thawed on ice. Liquid suspension was amended with 1 μg of plasmid DNA and incubated on ice for 10 minutes. Thereafter, the cells were frozen in liquid nitrogen and heat shocked at 37 °C for 5 minutes. This was followed by the addition of 400 μL of LB and outgrowth at 30 °C shaker for 3 hours. Following that, the cell suspension was spun (3000 RCF, 5 minutes) and the entire pellet was re-suspended in small volume of LB and plated on LB agar amended with antibiotics required for selection of the binary vector and the partner plasmid (see Table 2.1). Colonies of transformed *Agrobacterium tumefaciens* were detected following 48 h. of plate incubation at 30 °C.

For *Arabidopsis thaliana* transformation the floral dip protocol was followed (Clough and Bent, 1998). In brief, an overnight culture (50 mL) of *A. tumefaciens* bearing the binary vector and the partner plasmid was spun (3000 RCF, 5 minutes) and the pellet was re-suspended in 50 mL of the dipping solution (5% sucrose, 10 mM $MgCl_2$, 0.001% Silwett® L-77). Flowers of 4 week old *A. thaliana* were dipped in the suspension and following that the plant was bagged for 48 hours. After 4 more weeks of growth the seeds were collected from dipped *A. thaliana* plants.

For selection of transgenic *A. thaliana* seeds, green fluorescence seed screening system was used adapted after (Shimada et al., 2010). Specifically, the L1 transcriptional unit encoding OLE1-eGFP protein under the control of p*Oleosin* enables seed specific expression of a stable GFP tagged protein. Thus, transformed

seeds appear bright green under the UV light. These green seeds were collected and grown to generate between 10 and 15 T1 plants per construct. Seeds from T1 plants were collected and lines for which 75% of seeds were fluorescent were identified as the mono-insertional ones and used for generation of stable transgenic lines. To do that, T2 plants were grown and their seeds were screened for 100% fluorescence signal which means homozygosity.

2.4.4 DNA extraction and genotyping

In addition to the GFP seed screening each transgenic line was studied with PCR to confirm insertion of the construct. To do that, a small section of a leaf was placed in 100 μL of TNE-SDS genomic DNA extraction buffer (200 mM Tris-HCl, pH 8.0, 250 mM NaCl, 20 mM EDTA, 0.5% (w/v) SDS). Leaf material was ground in the buffer using a sterile burnt 1 mL pipette tip. The solution was spun at 3500 RCF for 5 minutes and the supernatant was mixed with 100 μL isopropanol in a new tube. This mix was incubated at -20 °C for 30 minutes and thereafter spun at 3500 RCF for 5 minutes. Supernatant was combined with 100 μL 70% ethanol (v/v) and spun again using the same conditions. The ethanol precipitation step was repeated and the resulting mix was dried *in vacuo* following centrifugation. The resulting pellet was resuspended in 20 μL sterile MiliQ water and 0.75 μL of it was used as a template for PCR.

Each genotyping PCR was run using RedTaq PCR Master Mix supplied by Sigma-Aldrich and used according to manufacturer's instructions. In brief, for each reaction the template was mixed with 12.5 μL RadTaq master mix, 10 μL water and 1.25 μL of 10 μM forward and reverse primer solutions (see Table 2.2 for primers used). Following initial melting step (2 minutes, 94 °C), the RedTaq reaction was run for 35 cycles with 94 °C melting for 1 minute, 55 °C annealing for 2 minutes and 72 °C extension for 3 minutes. This was followed by final elongation at 72 °C for 5 minutes. Products were analysed on 1% agarose gel prepared with GelRed DNA labelling reagent (Biotium) and visualised on a UV transluminator. PCR product size was estimated by comparison to a DNA standard (HyperLadder™ 1kb, Bioline).

Table 2.2 Primers used for cloning, sequencing and genotyping in this project

Target sequence	Primer sequence (5' -> 3')	Primer application
AtGUX1	ATGGCAAACTCTCCCGCTGCTCCT	Blunt ended cloning of CDS
	CTCTGCCACTCTCTTCACCA	Screening for correct insertion
	ATCAGTCAGCTGCCATAGCC	Screening for correct insertion
AtGUX2	ATGACGATAATGACGATGATAA	Blunt ended cloning of CDS
	GGCTCCTTCCAAATCTGCAC	Screening for correct insertion
	CTGAAGGAGGCTTTGAGCCA	Screening for correct insertion
PgGUX1	ATGAGGCCCTCTTCAGGAGTTC	Blunt ended cloning of CDS
	ACTCCCAGTTGGATCCTGTG	Screening for correct insertion
	TCCATAAGCAGCTGGAAGGT	Screening for correct insertion
PtGUX2	ATGAAATATAAAGGCCAGGC	Blunt ended cloning of CDS
	CAGCAATGGGAAGCTCTAGG	Screening for correct insertion
	TCCTGTGCCACCATGTAAAA	Screening for correct insertion
AtMAGT1	ATGGTCTCGCCTGAGAC	Blunt ended cloning of CDS
	TCGTGACGCATTTTACGGGA	Screening for correct insertion
pEAQ backbone	TTCTTGTCGGTGTGGTCTTG	Screening for correct insertion
	ACATAGAAATGCACACCGAATAACA	Screening for correct insertion

Promoter *IRX3*	CAGAGAGAGAAACTAAAAGC	Sequencing and screening
NOS terminator	TAATCATCGCAAGACCGGCA	Sequencing and screening
Myc tag	TCACAGATCTTCCTCAGAGA	Blunt ended cloning of tagged CDS
GoldenGate binary backbone donor	AAACCTTTTCACGCCCTTTT	Sequencing of final binary vectors made with GoldenGate
	ATCGAGTGGTGATTTTGTGC	

2.5 Heterologous protein expression and GUX activity assays

2.5.1 *Nicothiana benthamiana* expression using the pEAQ-HT system

Blunt ended Myc tagged GUX and AtMAGT1 CDSs were amplified from the synthetic construct using Q5 DNA Polymerase (NEB) using primers detailed in Table 2.2. For each PCR 55°C were used as primer annealing temperature and manufacturer's instructions were followed for the remaining part of the protocol. The PCR product was ligated into NruI (NEB) digested pEAQ-HT *N. benthamiana* overexpression vector (Sainsbury et al., 2009) using T4 DNA ligase (Thermo-Fisher Scientific). The construct was transformed into competent AGL-1 *Agrobacterium tumefaciens* following protocol detailed in section 2.2.3. Bacterial culture (OD_{600} between 0.6 and 0.8) was pelleted by centrifugation (3200 RCF for 10 minutes) and resuspended to the same bacterial cell concentration in infiltration medium (0.5% D-glucose, 50 mM MES and 2 mM Na_3PO_4, 1 mM Acetosyringone, prepared after (Sparkes et al., 2006). This bacterial solution was infiltrated into *N. benthamiana* leaves. Leaves were harvested 3 days following the infiltration and the membranes fraction enriched for GUX was collected. For the membrane collection, protocol was adapted after Rennie et al. 2012. In brief, leaves were homogenised in a microsome buffer by grinding with mortar and pestle in a cold room. Homogenate was strained through a nitrocellulose mesh and a cleared solution was pelleted by centrifugation (3200 RCF for 10 minutes). Cleared supernatant was aliquoted into 2 mL tubes and membranes were collected by two rounds of centrifugation (21000 RCF for 1 h. and 0.5h.). Obtained membranes were frozen in liquid nitrogen and stored at -80°C for further use in the *in vitro* activity assays.

2.5.2 Western blot detection of expressed proteins

Protein concentration in the membranes fraction was quantified using modified Bradford reagent (Expedeon). Each well of SDS-PAGE (10–15% gradient, BioRad) was loaded with 2.5 or 5 µg of GUX enriched *N. benthamiana* leaf membrane protein. Following the run, the gel was transferred onto nitrocellulose membrane using iBlot system (Life Technologies). The membrane was blocked over-night in 5% milk in TBS solution. The following day it was probed with 1:2000 anti-Myc primary antibody (rabbit polyclonal, Santa-Cruz, A14) and with 1:10,000 mouse anti-rabbit HRP linked secondary antibody (Bio-Rad, 170-6515). Amersham ECL prime HRP substrate (GE-Lifesciences) was used to obtain signal from membrane bound antibodies.

2.5.3 Extraction of acetylated heteroxylan

AIR (100 mg) from *gux1/2* plants was used as a starting material for extraction of the acetylated heteroxylan lacking [Me]GlcA decorations. The material was de-pectinated by incubating in ammonium oxalate solution (0.5%; 10 mL) at 85 °C for 2 h. Thereafter, the suspension was washed with water and delignified by incubating in per-acetic acid (11%; 5 mL) for 30 minutes at 85 °C. Per-acetic acid was removed from the suspension with one acetone wash and two water wash steps. Obtained holocellulose pellet was dried in an oven set at 50°C for 48 h.

For xylan extraction 25 mg of holocellulose was used. The extraction was performed by homogenising the pellet in 5 mL DMSO and incubating at 80 °C for 24 h. This was followed by exchange of the DMSO medium and another 24 h extraction period at 80°C. DMSO fractions from both days were combined. Acetylated xylan was washed into water by exchanging the solvent using PD-10 desalting collumns (GE Healthcare). Xylan suspension was dried *in vacuo* and stored at room temperature for further use. From 25 mg of holocellulose 25 aliquots of acetylated heteroxylan were obtained following extraction and PD-10 purification.

For estimation of xylan extraction efficiency fractions were collected for the ammonium-oxalate extracted xylan, post-ammonium oxalate extraction residues and the final DMSO extract. Each fraction was exchanged into 0.1 M ammonium acetate buffer using PD-10 desalting columns and digested with xylanase *Np*GH11.

2.5.4 GUX activity assays

Each reaction mix was prepared as described by Rennie et al., 2012 with omission of UDP-xylose. In brief, dried aliquots of acetylated heteroxylan were resuspended in 30 μL of a reaction master-mix (0.5 mM DTT, 10 mM $MnCl_2$, 10 mM $MgCl_2$, 2% Triton-X100, 10 mM UDP-GlcA). UDP-GlcA concentration was changed for some reactions or the sugar nucleotide was replaced with water in certain reactions to control for non-specific glucuronosylation. The reaction mix was amended with 30 μL of undiluted microsomal proteins extracted from *N. benthamiana* leaves. For most experiments the reaction was performed for 5 h. at room temperature and terminated with a 10 minute long 100 °C heat treatment. In the case of reactions including acetylxylan esterase (Acetylxylan esterase CE4 from *Clostridium thermocellum,* NzyTech CZ00341) 5 μL of enzyme suspension was added to the reaction following 3 h. from its start. Additional UDP-GlcA was introduced into the reaction upon addition of the esterase enzyme leading to doubling of the initial concentration in the reaction tube. Reactions performed in the presence of the esterase enzyme were carried out over-night.

Following termination of the reaction the polysaccharide product was extracted with a Methanol: Chloroform lipid removal protocol adapted after (Bligh and Dyer, 1959). In brief, the reaction mix was amended with 450 μL of 1:2 (v:v) Methanol: Chloroform mix and shaken vigorously. This was followed by addition of 150 μL of Chloroform and 150 μL of water. Each liquid addition step was followed by vigorous shaking. Obtained mixture was phase-separated by centrifugation at 160 RCF for 10 minutes. Top aqueous layer was collected and amended with ethanol absolute to the final concentration of 80%. To precipitate polysaccharides, this mix was incubated at -20 °C for 1 h. and spun at 21000 RCF for 15 minutes. This was followed by two wash steps with 1 mL of ethanol absolute. Obtained polysaccharide pellet was dried *in vacuo* and stored for further analysis at -20 °C. In most cases, the pellet was de-acetylated and digested with xylanase enzymes as described in section 2.2.1.

2.6 Molecular phylogenetics and GUX expression profiling

2.6.1 Molecular phylogenetics

The coding sequences of *A. thaliana* GUX 1, 2 and 3 were used to identify putative GUX encoding transcripts from *Populus trichocarpa* using data available via the NCBI BLAST service. The same Arabidopsis CDSs were used as a query to identify transcripts encoding putative GUX enzymes from Coniferophyta and Gnetophyta transcriptomic data available via 1000 Plant Transcriptomes, Congenie and Gymno Plaza BLAST services (Matasci et al., 2014, Nystedt et al., 2013, Proost et al., 2015). All amino acid sequences were reconstructed from transcripts with ExPASy translate tool and aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE) algorithm. A maximum likelihood phylogenetic tree was constructed using MEGA 6 software (Tamura et al., 2013) including the gamma distribution algorithm in the reconstruction. Table 2.3 summarises codes required to access gymnosperm *GUX* sequences studied in this Thesis.

Table 2.3 Summary of codes required to access sequences of gymnosperm *GUX* transcripts or protein products using publically available databases.

Putative GUX enzyme code	**Database code for sequence access**
Abies lasiocarpa GUX1	VSRH-2058742 (OneKP)
Gnetum montanum GUX1	GTHK-2014663 (OneKP)
Juniperus scopulorum GUX1	XMGP-2003994 (OneKP)
Larix speciosa GUX1	WVWN-2055446 (OneKP)
***Pcea glauca* GUX1 (PgGUX1)**	**GQ03239 (GeneBank)**
Picea englemanii GUX1	AWQB-2005588 (OneKP)
Pinus ponderosa GUX1	JBND-2013068 (OneKP)
Picea abies GUX1	MA_84103g0010 (Congenie)
Taxus baccata GUX1	WWSS-2012177 (OneKP)
Welwitchia mirabillis GUX1	TOXE-2013077 (OneKP)
Pinus taeda GUX1	PTA00041003 (Gymno Plaza)
***Pinus taeda* GUX2 (PtGUX2)**	**PTA00017485 (Gymno Plaza)**
Picea sitchensis GUX2	PSI00008771 (Gymno Plaza)
Picea abies GUX2	MA_35705g0010 (Congenie)

2.6.2 Bioinformatics tools used to map conifer GUX expression

For the computational expression profiling of GUX enzymes Congenie (Nystedt et al., 2013) and Norwood (Jokipii-Lukkari et al., 2017) bioinformatics tools were use. For the Congenie analysis, the exImage tool was used to visualise the expression across the entire organism. For NorWood analysis, the expression profiles of putative conifer *GUX* were compared to these responsible for the biosynthesis of other cell wall polysaccharides and lignin. These putative cellulose, mannan, xylan and lignin biosynthesis genes were identified using Congenie BLAST service using *A. thaliana* protein sequences as a query. Data visualisation was performed by an imaging script built into the NorWood tool.

2.6.3 *In situ* hybridisation experiments

In situ hybridisation probes were designed against the 1000 bases close to the 3' end of the *PgGUX1* and *PaGUX2* mRNA and ordered as synthetic DNA. Remaining parts of the *in situ* hybridisation experiments were performed by Dr Weibing Yang at the Sainsbury Laboratory – University of Cambridge following a published protocol (Yang et al., 2016). In brief, probes were ligated into a pGEM-T Easy® vector (Promega) and *in vitro* transcribed in the presence of Digoxigenin-11-Uridine triphosphate (DIG-UTP). Probes were hybridised to 8 μM sections of wax embedded fixed 1 year old *Picea glauca* stems. Following probes derivatisation the section was probed with AP linked anti-DIG antibody and signal was visualised using a colorimetric reaction.

2.7 Solid state NMR experiments

2.7.1 Enrichment of plant material with ^{13}C

Arabidopsis thaliana plants used for ^{13}C enrichment were grown for 15 days on solid MS medium prior to the transfer into the enrichment chamber. For this, seeds were surface sterilised by first washing for 5 minutes in 70% ethanol (v/v) supplemented with 0.02% Tween-20 and then washing with 96% ethanol (v/v). Sterilised seeds were spread on filter paper soaked in ethanol absolute and air-dried in a sterile hood. Dry sterilised seeds were transferred into a sterile tube and soaked in sterilised tap water for 48 hours at 4 °C prior to plating on solid medium. For seedling growth, the seeds were plated on plates containing Murashige and Skoog (MS) Basal Salt Mixture (2.2 g / L), sucrose (1% w:v) and fine agar (0.9% w:v, Duchefa Biochemie). Seeds

were grown vertically on MS plates and after 15 days of growth seedlings were transferred into 7X7X10 cm pots filled with rockwool (Cultilene, OptiMaxx) soaked in hydroponics solution (2 mM $MgSO_4$, 2 mM $Ca(NO_3)_2$, 50 µM FeEDTA, 5 mM KNO_3, 2.5 mM K_2HPO_4 + KH_2PO_4 buffer pH = 5.5, 70 µM H_3BO_3, 14 µM $MnCl_2$, 0.5 µM $CuCO_4$, 1 µM $ZnSO_4$, 0.2 µM $NaMoO_4$, 10 µM NaCl and 0.1 µM $CoCl_2$).

Enrichment of biomass with ^{13}C was performed by growing the *A. thaliana* in an in-house designed and built sealed $^{13}CO_2$ growth chamber (Figure 2.2). For the growth, *A. thaliana* plants in rockwool medium were supplied with a gas mix in which air, with $^{12}CO_2$ removed, was mixed with $^{13}CO_2$ in a v:v ratio of 2000:1. This gas mix was pumped into a sealed, temperature and humidity controlled, chamber with *A. thaliana* plants. As a part of this PhD the growth process was optimised further over what was published by Simmons et al., 2016 by constructing new LED light box and exchange of hydroponics medium every 48 hours. Plants were grown in the ^{13}C enrichment chamber for 6 weeks after which 2 cm of basal stem was harvested, frozen in liquid nitrogen and stored at -80 °C for further analysis.

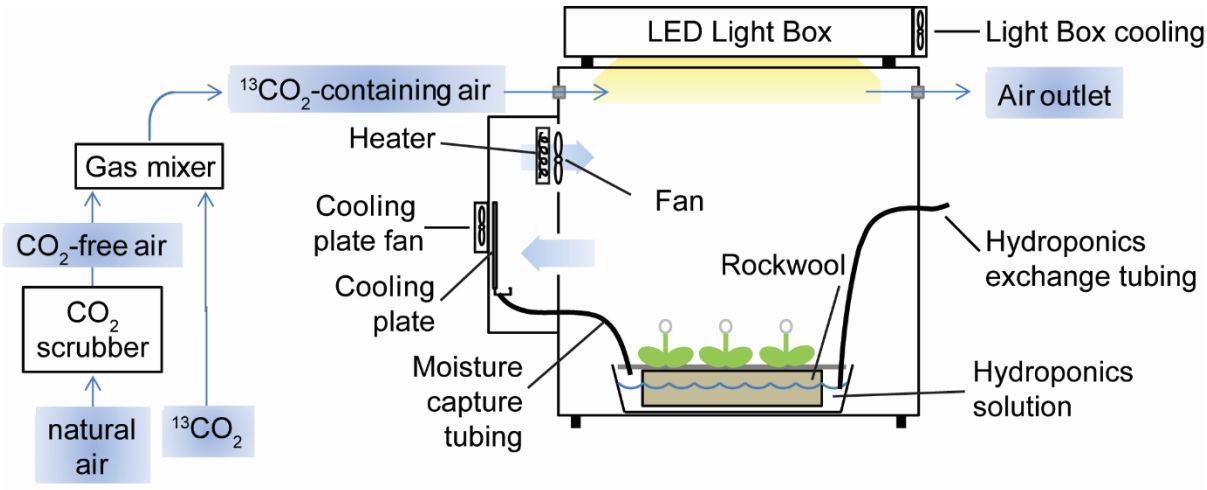


Figure 2.2 Schematic representation of the ^{13}C enrichment chamber. Figure modified from (Simmons et al., 2016), detailing improvements introduced as a part of this project.

2.7.2 Running of CP-INADEQUATE experiments

Solid state NMR experiments were performed as described in Simmons et al., 2016. For the experiments between 35 and 50 mg of hydrated plant material was packed into 3.2 mm MAS probe. The experiments were performed on Bruker (Karlsruhe, Germany) 850 MHz Advance III solid-state NMR spectrometer by Prof Ray Dupree. During the experiments the sample was kept at room temperature and spun at the magic angle with a frequency of 12.5 kHz. Optimisation of experimental conditions and choice of parameters was decided by Prof Dupree and matches what was described by Simmons et al., 2016. Obtained spectra were Fourier transformed and analysed using TopSpin software (Bruker) with the help from Prof. Dupree and Mr Oliver Terrett.

2.8 Cryo-SEM imaging of *A. thaliana* and *P. abies* stems

2.8.1 Sample preparation and imaging

Fresh stems of 7 week old *A. thaliana* plants were prepared for imaging as outlined on Figure 2. Each imaging experiment had three main phases. Firstly, 1 cm sections were cut from the bottom part of the stems and mounted in glue (Parts 1 to 4 on Figure 2.3). Secondly, stem sections were frozen in liquid nitrogen (Part 5), fractured and coated with 3 nm of platinum (Part 6). Finally, sections were imaged with Zeiss EVO HD15 Scanning Electron Microscope (Part 7). Images were acquired using a secondary electron detector. Electrons were accelerated with voltage between 6 and 15 kV.

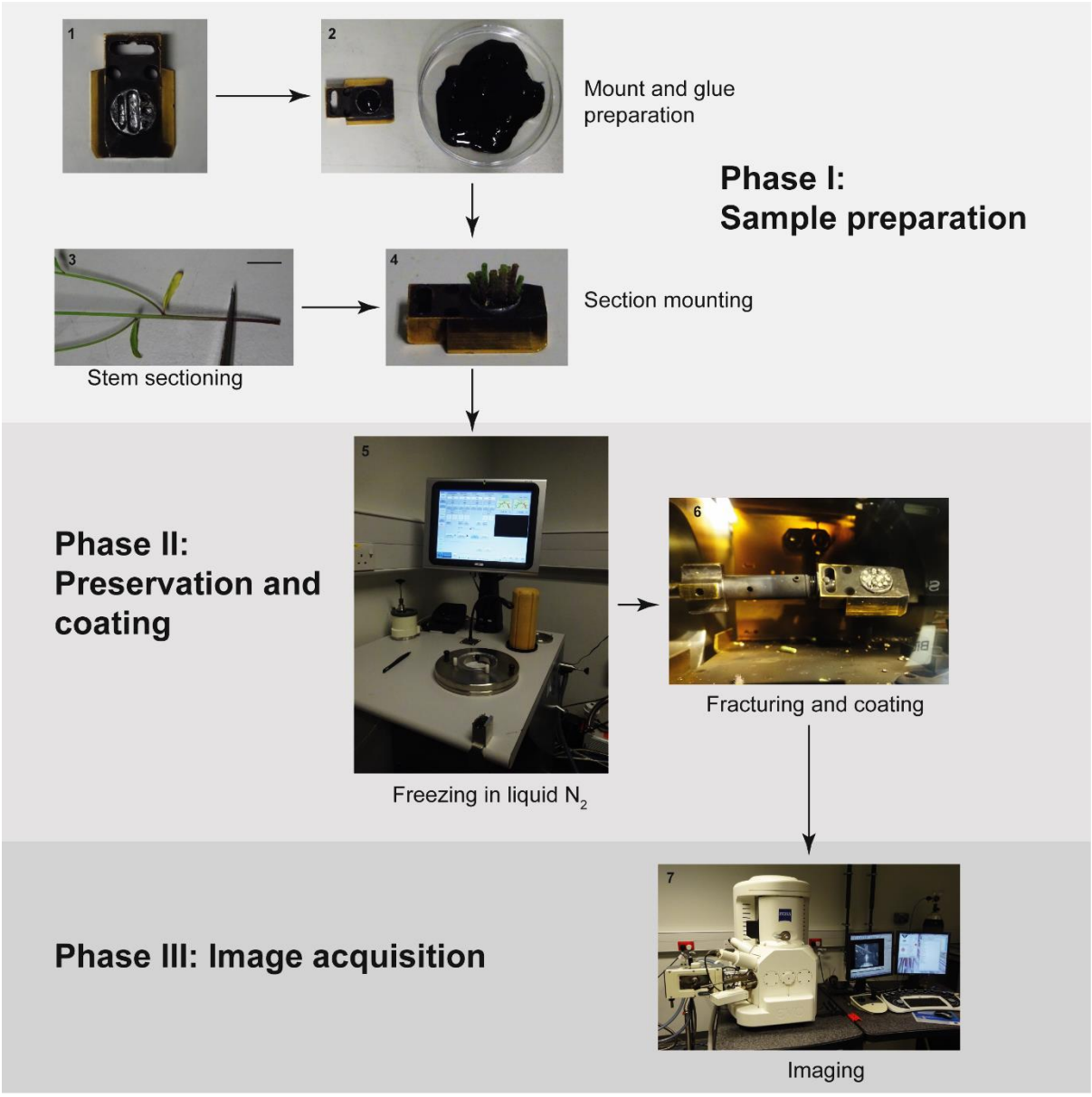


Figure 2.3 Overview of cryo-SEM procedure. Part 2 size bar is 1 cm long.

2.8.2 cryo-SEM image quantitation

Quantitation of cell wall microfibril width was performed using ImageJ software (National Institute for Health). For each vessel analysed a line was drawn across its width and the line length was quantified by comparing its pixel dimension to this of the standard provided when acquiring the image. Each fibril width measurement was standardised for the platinum layer applied during the coating process.

2.9 Statistical analysis, sampling and thesis preparation

For all quantitative experiments, unless otherwise stated, three biological replicates of plant material were grown and analysed. Each biological replicate of AIR consisted of a pooled sample of 36 plants. For each biological replicate, 3 technical replicates were analysed for the quantitation of sugar release and fermentation efficiency.

Statistical analysis was performed using packages available with R software (R Foundation). Statistical tests used to compare average measurements for samples are mentioned in Figure legends and mostly include Student's T test and ANOVA. The variance between each pair compared was estimated to be similar with Levene's test.

This thesis was written using Word software (Microsoft) and figures were prepared using Adobe Illustrator CC 2018 (Adobe).

Chapter 3: The presence of GlcA branches on xylan contributes to the maintenance of biomass recalcitrance.

3.1 Introduction

For a long time plant biomass has been suggested as a possible renewable source of energy and biomaterials. This includes the use of plant secondary cell walls as a source of sugars for biofuel fermentation (Hood, 2016) but also for other applications, such as the use of timber in building construction (Ramage et al., 2017). Progress in the development of biomass derived products has been hindered by the lack of understanding of plant cell wall molecular architecture. This includes our understanding of what structures crosslink different cell wall polysaccharides and phenolic compounds to maintain resistance of the cell wall to enzymatic or mechanical processing, a property known as biomass recalcitrance (McCann and Carpita, 2015).

To achieve a high degree of biomass conversion, industrial processes, including second generation biofuel production, employ multiple harsh chemical or mechanical pre-treatment processes to decrease recalcitrance. These techniques include, but are not limited to, ionic liquid, alkali, acidic or steam explosion pre-treatments (Silveira et al., 2015). All these processes are energy intensive and contribute significantly to the final cost of the biomass derived product. Moreover, some pre-treatments generate toxic chemical effluents. Therefore, research has focused on engineering or breeding plants with reduced recalcitrance. Modifications observed in many of the plant lines with reduced recalcitrance include changes in cell wall biosynthesis pathways.

This chapter focuses on modifications of xylan branching and its impact on biomass recalcitrance and uses *Arabidopsis thaliana* as a model of hardwood xylan. *A. thaliana* secondary cell wall xylan is a polymer of β-1,4- linked xylopyranosyl units which carry O2, O3 or double acetylation and α-1,2-linked glucuronic acid (GlcA) branches which can be 4-O-methylated. The enzymes involved in the biosynthesis of xylan and its exact structure were described in Chapter 1. Importantly, xylan branches may contribute to biomass recalcitrance. As discussed extensively in Chapter 1, [Me]GlcA branches may be esterified to lignin. The presence of xylan acetylation may impede the activity of xylanase enzymes which can result in an increase in biomass recalcitrance (Poutanen et al., 1987). Previously, a reduction of GlcA methylation levels was observed to be associated with a decrease in biomass recalcitrance

(Urbanowicz et al., 2012). Work presented in this chapter will attempt to probe the importance of these branches for the maintenance of biomass recalcitrance using a single saccharification assay. Unlike some of the published research, experiments performed as part of this work do not use any chemical pre-treatment to reduce biomass recalcitrance. This approach is likely to enhance the sugar yield difference observed between mutant and WT plants that arise from cell wall architectural changes, and thus might be more suited to identification of recalcitrance determining factors.

Taking this into consideration the specific aims of this chapter are:

- To establish a saccharification assay which is capable of probing the importance of xylan branching to the maintenance of biomass recalcitrance.

- To compare the release of xylose and glucose from biomass lacking single types of xylan branching with that obtained from WT Arabidopsis material.

- To investigate possible reasons for any discovered saccharification phenotypes and evaluate if they can result in an increase in biofuel yields in a bacterial fermentation system.

Most of the results presented in this chapter were described in a *Biotechnology for Biofuels* journal publication (Lyczakowski et al., 2017). The publication is attached as Appendix 1 to this thesis.

3.2 Plant biomass with xylan lacking glucuronic acid branches has reduced recalcitrance to enzymatic digestion.

In order to evaluate the impact of xylan branching on biomass recalcitrance, Alcohol Insoluble Residues (AIR) were prepared from the stems of three types of 8 week old mutant *A. thaliana* plants: the first one, *esk1* or *tbl29*, which has reduced xylan acetylation, another lacking secondary cell wall xylan glucuronosylation (*gux1/2*), and the last one with reduced levels of GlcA methylation (*gxm1/2/3*). AIR from mutant and matching WT plant biomass was incubated with Novozymes Cellic® CTec2 enzymatic saccharification cocktail for 24 h, following which the D-glucose (Figure 3.1A) and D-xylose (Figure 3.1B) release were quantified.The reaction conditions and the amount of enzymes used aimed to achieve maximal possible biomass digestion and were optimised by Dr. Marta Busse-Wicher before the start of this PhD project. For both

acetylation (*tbl29*) and GlcA methylation (*gxm1/2/3*) mutants, no significant change in monosaccharide release was observed when compared to matching WT controls. Interestingly, AIR from *gux1/2* plants, lacking secondary cell wall xylan glucuronidation, released 30% more D-glucose and as much as 700% more D-xylose. This data suggests that of the three types of xylan modification found in secondary cell walls of hardwood plants, GlcA is likely to be the only one with a significant impact on the resistance of the material to enzymatic saccharification.

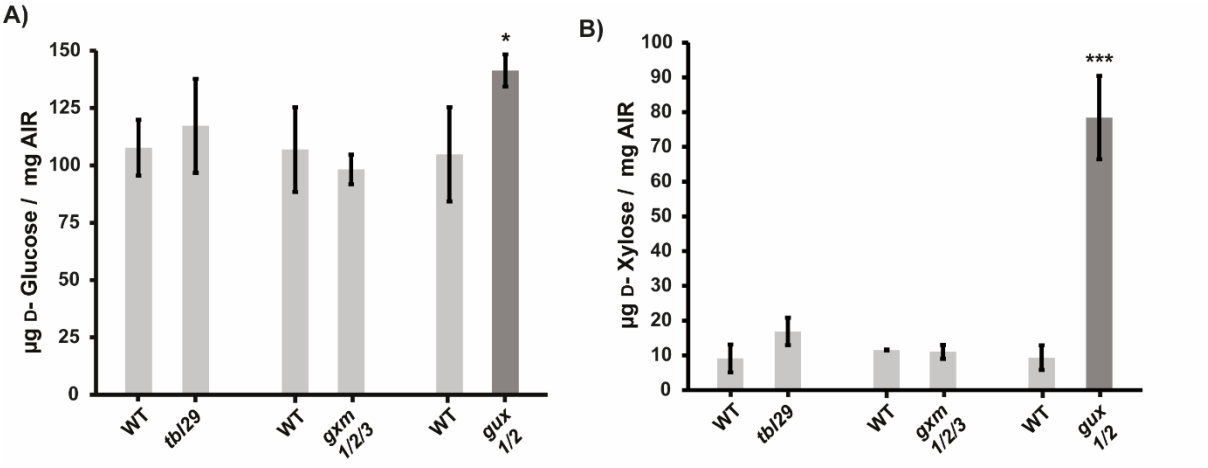


Figure 3.1 Biomass lacking xylan–[Me]GlcA decorations has reduced recalcitrance. Mean release of D-glucose (A) and D-xylose (B) from WT and *tbl29/esk1*, *gxm1/2/3* and *gux1/2* following AIR saccharification. Error bars represent standard deviation of three matching WT and mutant biological replicates of biomass, *p value ≤0.05; **p value ≤0.01; ***p value ≤0.001 in Student's t-test.

3.3 Use of biomass lacking [Me]GlcA branches on xylan as a feedstock improves ethanol production in the bacterial fermentation system.

The ease of biomass conversion and plant growth are both an important consideration for any bioenergy crop. The majority of the increase in monosaccharide release observed in *gux1/2* biomass can be attributed to D-xylose (Figure 3.1). Valorisation of this pentose sugar is challenging, as it is not metabolised by the traditional *Saccharomyces cerevisiae* ethanol fermentation system (Lewicka et al., 2014). To demonstrate if the increase in sugar release from *gux1/2* biomass can successfully contribute to ethanol production, ground stems of WT and mutant plants were incubated with Pyruvate decarboxylase (Pdc) and Alcohol dehydrogenase (AdhB) expressing *E. coli* in the presence of Cellic® CTec2. Expression of *Zymomonas*

mobilis Pdc and AdhB enables *E. coli* to produce ethanol from pyruvate (Lewicka et al., 2014). As pyruvate is generated in bacterial cells from both pentose and hexose sugars, the transgenic *E. coli* strain was hypothesised to be capable of utilising the excess D-xylose released from *gux1/2* biomass. Following 96 h of simultaneous saccharification and fermentation, ethanol yields were measured in vessels containing WT or *gux1/2* feedstock (Figure 3.2). On average, the fermentation reactions utilising *gux 1/2* biomass released twice as much ethanol as the WT feedstock.

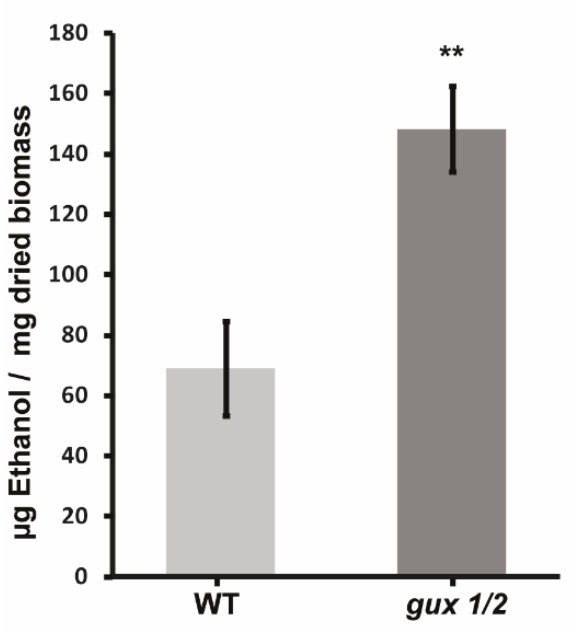


Figure 3.2 Use of biomass lacking xylan–[Me]GlcA decorations improves ethanol yields in the bacterial fermentation system. Average ethanol release from SSF experiments using WT and *gux1/2* stem AIR. Error bars represent standard deviation of three matching WT and mutant biological replicates of biomass **p value ≤0.01 in Student's t-test

Previous reports indicate that the removal of GlcA branches from both primary and secondary cell wall xylan in *gux1/2/3 A. thaliana* plants has no significant impact on plant growth (Mortimer et al., 2015). To further support this notion, total plant height, total plant mass and the basal stem mass were quantified for WT and matching *gux1/2 A. thaliana* (Figure 3.3A to C). For all of the parameters measured, no significant differences between the WT and the mutant were observed. Together with the results obtained from the SSF experiments, this data indicates that genetic removal of GlcA from xylan generates biomass which outperforms the WT as a biofuel feedstock.

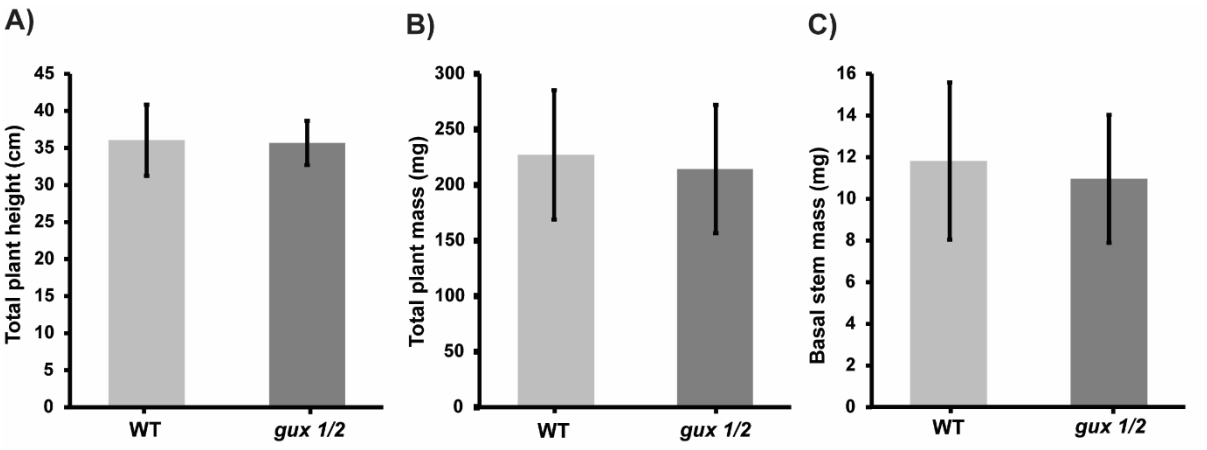


Figure 3.3 Plants lacking xylan-GlcA decorations have no negative growth phenotype. Average height of 7 week old plants (A, n = 36 for WT and 34 for *gux1/2*), average total plant mass (B, n = 36 for WT and 34 for *gux1/2*) and average mass of 5 cm basal stem sections (C, n = 36 for WT and 33 for *gux1/2*). Error bars represent standard deviation. There is no statistically significant difference between the values measured for the WT and the mutant plant (Student's t-test).

3.4 The increased digestibility of *gux1/2* biomass may be caused by changes in the molecular architecture of the cell wall.

It is possible that the observed increase in D-xylose release from *gux1/2* biomass compared to the WT may be caused simply by the composition of the Cellic® CTec2 enzymatic saccharification cocktail. Full degradation of xylan requires the activity of α-glucuronidase enzymes which remove GlcA branches from oligosaccharides released by xylanases. As *gux1/2* biomass lacks GlcA branches on xylan, the polysaccharide may be degraded without the action of α-glucuronidases. Indeed, simple digestion of alkali extracted *gux1/2* xylan with xylanase GH11 and β-xylosidase can result in complete conversion of the polymer to monosaccharides while for WT as much as 50% of the xylose monomers remains in glucuronidated oligosaccharides (Mortimer et al., 2010). To test if Cellic® CTec2 contains sufficient glucuronidase activity, CTec2 digested WT and *gux1/2/3* AIR was treated with GH115 α-glucuronidase and re-digested with the saccharification cocktail (Figure 3.4A). D-xylose release was measured following the first and the second saccharification event (Figure 3.4B). No significant increase in the release of monosaccharides was observed in the WT biomass.

One possible explanation of this observation is that both the GH115 present in Cellic® CTec2 and the one supplemented to the saccharification reaction may be unable to cleave all of the [Me]GlcA from xylooligosaccharides. It is conceivable that the enzymatic cocktail used may lack specific esterase capable of removing O3 acetylation from glucuronidated xyloses, a process which is required for GH115 activity (Razeq et al., 2018). Alternatively, this observation may indicate that Cellic® CTec2 may contain adequate amounts of α-glucuronidase activity to allow for complete digestion of accessible xylan. This would suggest that the increase in sugar release observed in *gux1/2* plants is not merely an indication of saccharification cocktail composition and may be a result of alterations in the secondary cell wall molecular architecture. This may include loss of putative ester linkages to lignin in *gux1/2* plants and their presence on undigested WT xylan or xylooligosaccharides.

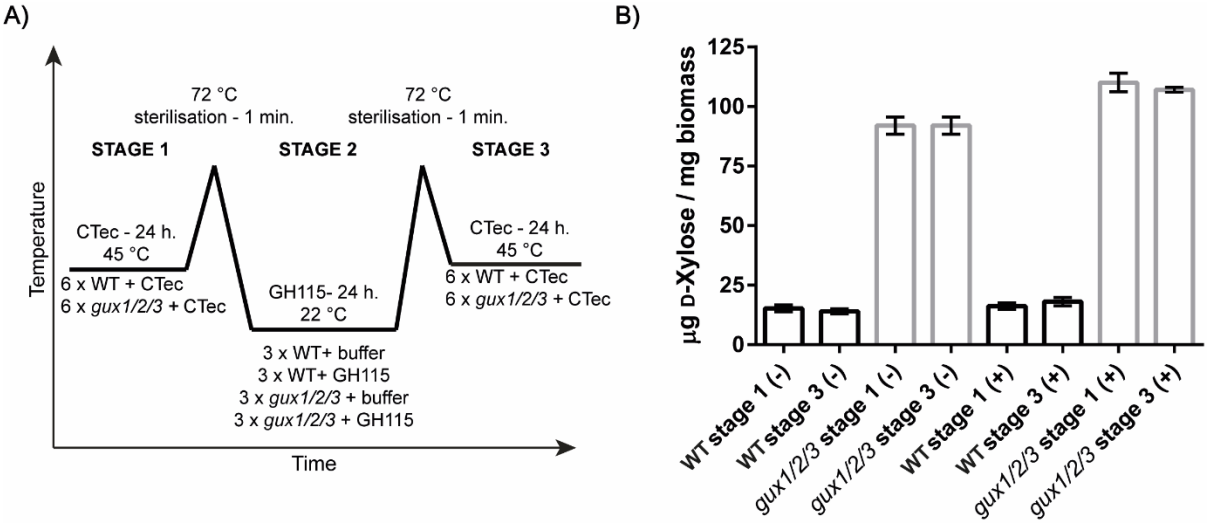


Figure 3.4 Supplementation of enzymatic cocktail with additional α-glucuronidase does not greatly improve sugar release from WT biomass. Reaction scheme for GH115 supplementation (A) and D-xylose release from WT and *gux1/2/3* biomass after supplementation (+) or not (-) with GH115 (B). Sterilisation steps were carried out between different stages of the experiment to avoid microbial growth in the reaction tubes. Reactions were performed in triplicate, and D-xylose release was measured after stage 1 and stage 3 of the experiment. Error bars represent standard deviation.

Xylan extractability is one of the measures of the degree of its interaction with other cell wall components (Mortimer et al., 2010). Changes in hemicellulose-cellulose or hemicellulose-lignin interactions may be reflected in the behaviour of the polysaccharides in an extractability assay. Indeed, *gux1/2* xylan was observed to be more easily extractable than the WT during mild alkali extraction (Mortimer et al., 2010). To investigate the importance of [Me]GlcA for xylan extractability further, an assay developed by Dr. Marta Busse-Wicher to study *gux1/2* biomass was applied to *gux1/2/3* biomass which lacks xylan glucuronidation in both primary and secondary cell walls. In this experiment, extractability is measured without the alkali pre-treatment step by subsequent washes with ammonium oxalate (pectic fraction), peracetic acid and DMSO. The amount of xylan in each fraction is analysed with PACE. Polysaccharides extracted from WT and *gux1/2/3* AIR into the pectic fraction, DMSO fraction, and those remaining in the residues, were deacetylated with an alkali treatment to facilitate full enzyme digestibility and digested using xylanase GH11. Released oligosaccharides were resolved using PACE (Figure 3.5A). As *gux1/2/3* biomass lacks [Me]GlcA branches, GH11 xylanase digestion resulted in the release of Xylose (X) and Xylobiose (X2). Due to the inhibitory effect of GlcA on GH11 activity,

the WT biomass releases X, X2 and an XUXX (UX_4) pentasaccharide. The identity of the detected bands and their fluorescence intensity was used to quantify the relative total xylan backbone quantity extracted into each fraction (Figure 3.5B) following analysis of three biological replicates of the biomass, each on an independent PACE gel. This revealed that the majority of WT xylan remained unextracted, with close to 90% of it being detected in the residues fraction and only ~10% extracted into DMSO. For *gux1/2/3* xylan, 45% of it was released into the DMSO fraction. This close to 5 fold increase in xylan extractability following genetic removal of [Me]GlcA branches supports the conclusion that the absence of this modification may lead to alterations in the molecular architecture of plant biomass.

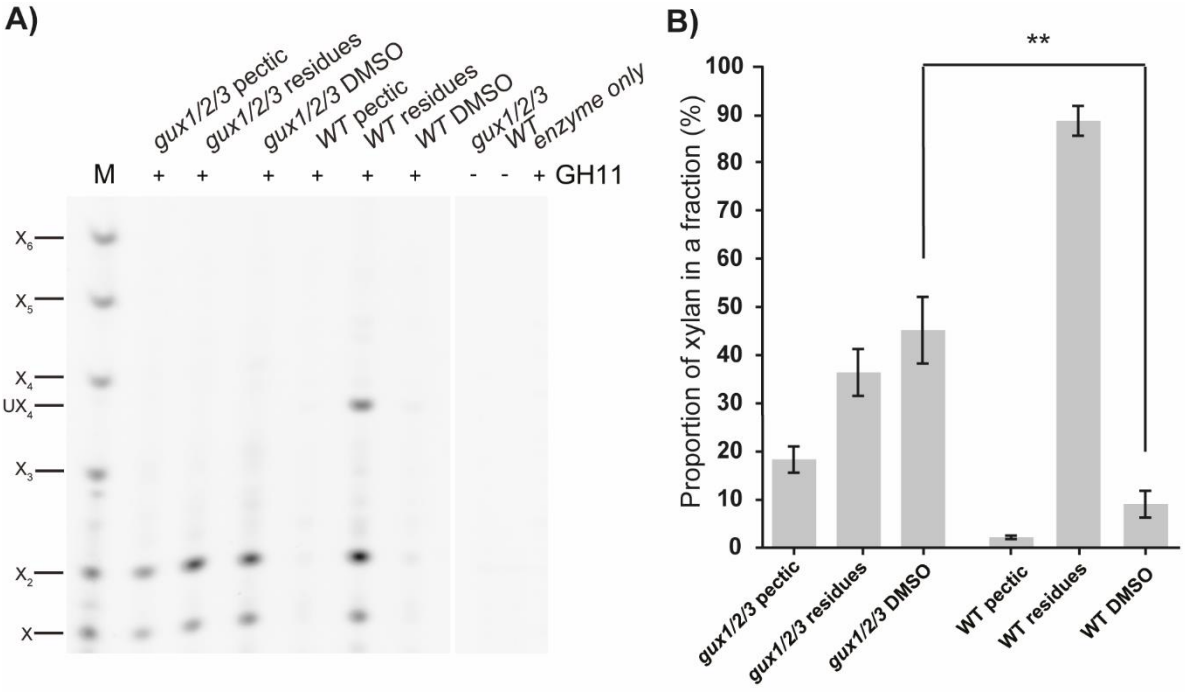


Figure 3.4.2 Genetic removal of [Me]GlcA branches improves xylan extractability. PACE analysis of *gux1/2/3* and WT xylan remaining in solid residues and extracted into pectic and DMSO fractions (A). The GlcA-xylotetraose band (UX4) was observed only in WT. Analysis includes non-digested material extracted into DMSO and an enzyme only control. B) Quantitation of the amount of xylan in pectic, residues and DMSO fractions extracted across three biological replicates of *gux1/2/3* and WT biomass. Error bars represent standard deviation of three matching WT and mutant biological replicates of biomass **p value ≤0.01in Student's t-test

3.5 Discussion

3.5.1 Modifications of xylan structure influence biomass recalcitrance

Numerous studies have attempted to decrease biomass recalcitrance by genetically engineering cell wall biosynthesis (Loque et al., 2015). Nonetheless, these cell wall alterations are often associated with reduced plant growth. Cellulose crystallinity correlates negatively with the recalcitrance of plant biomass (Harris et al., 2009). The efficiency of converting biomass to sugars can be increased by as much as 150% when the crystallinity of cellulose is decreased by a third in plants mutated in a gene encoding a secondary cell wall cellulose synthase complex member IRX5 (Harris et al., 2009). This gain, however, is offset by a severe growth phenotype exhibited by *A. thaliana irx5* plants (Brown et al., 2005). Similarly to cellulose crystallinity modifications, mutations in the lignin biosynthesis pathway can lead to reduced biomass recalcitrance (Van Acker et al., 2013). Mutations in *A. thaliana* genes encoding Cinnamate 4-Hydroxylase (C4H) and Cinnamoyl-CoA Reductase (CCR1) can enable over 4 fold improvement in cellulose conversion to glucose. However, the decreased lignin content and reduced recalcitrance observed in these plants is associated with severe dwarfing with as much as 80% reduction in plant mass observed in *ccr1-3 A. thaliana* plants. Interestingly, there are lignin engineering approaches which allow for reduced recalcitrance without yield penalty (Mottiar et al., 2016). These include downregulation of Cinnamyl Alcohol Dehydrogenase (*CAD*) expression in switchgrass, which allowed for a doubling in the saccharification efficiency with plants growing like wild-type organisms in the greenhouse (Fu et al., 2011).

In addition to cellulose and lignin engineering, hemicellulose, and especially xylan structure engineering, has been used to develop biomass more suited for biofuel production. In particular, similarly to the reports outlined in this chapter and as discussed in subsequent sections of this discussion, previously published research aimed to reduce acetylation, glucuronidation and glucuronic acid methylation of xylan. Reduction of xylan acetylation has, for a long time, been considered to be an excellent target for improvement of biomass properties (Loque et al., 2015). In addition to reducing the need for inclusion of xylan acetyl esterases in the enzymatic cocktails, the removal of xylan acetylation may stimulate the activity of hydrolytic enzymes, such

as xylanases and glucuronidases. Moreover, removal of xylan acetylation may prevent acidification of the saccharification broth. This process is believed to impede yeast growth and thus may reduce biofuel yields obtained from SSF (Pawar et al., 2013).

Two main approaches were used to genetically reduce xylan acetylation to evaluate its importance for the maintenance of biomass recalcitrance. Firstly, genes encoding a broad range acetylation mediator, Reduced Wall Acetylation (RWA), were targeted for mutagenesis and silencing. As mutations in *RWA* lead to a reduction in acetylation of not only xylan but also other non-cellulosic polysaccharides, the enzyme is believed to be responsible for provision of a substrate to the transferase enzymes (Manabe et al., 2013). *A. thaliana* genome encodes four *RWA* genes and a triple mutant *rwa1/3/4* shows ~60% reduction in xylan acetylation (Manabe et al., 2013). Downregulation of *RWA* expression in hybrid aspen resulted in c. 25% reduction of xylan acetylation and was associated with a 14% increase in D-glucose release and 40% increase in D-xylose release (Pawar et al., 2017).

Secondly, *Trichome Birefringence Like 29 (TBL29)* / *Eskimo 1 (ESK1)*, which encodes an active xylan acetyl transferase (Urbanowicz et al., 2014) was mutagenised. T-DNA insertional mutants of the *ESK1* gene show c. 60% reduction in xylan acetylation (Xiong et al., 2013). However, similarly to the results reported in this chapter, when expressed per mg of plant biomass, *esk1* AIR performed like the WT in saccharification assays. As *esk1* plants are dwarfed (Xiong et al., 2013, Lefebvre et al., 2011) a suppressor screen was performed to identify secondary mutants allowing generation of non-dwarfed plants without ESK1 activity (Bensussan et al., 2015). The screen has identified that secondary mutations in the *KAKTUS* (*KAK*) gene partially rescue growth of *esk1* plants, while still maintaining the reduction in xylan acetylation. Saccharification of *esk1kak* double mutant biomass resulted in a 20% increase in D-glucose release. This difference might result in differences in saccharification conditions used by Bensussan et al.. Specifically, their work involved longer saccharification (60h compared to 24h in this work) and the use of a different enzymatic cocktail. Alternatively, it is therefore possible that the absence of a saccharification phenotype of *esk1* plants observed in work reported in this chapter may be caused by a secondary effect of plant dwarfing. Harvesting matching biomass from WT and the dwarfed *esk1* mutant is challenging as the proportion of primary and

secondary cell walls in the final AIR preparation depends on the plant growth. It will be therefore interesting to use the developed saccharification assay to evaluate recalcitrance of both *esk1-5:kak-8* and a recently reported *esk1:max4* plants (Ramirez et al., 2018) which have reduced xylan acetylation levels and close to WT-like growth.

On top of the attempts to modify xylan acetylation, efforts were made to affect the GlcA methylation with an aim of generating a less recalcitrant biomass source. Reduction in 4-O methylation of xylan GlcA in *gxm1* plants was reported to lead to a c. 5% increase in xylose release from acid pre-treated AIR. The *A. thaliana gxm1/2/3* mutant has a significantly greater reduction in GlcA methylation than the one observed in the *gxm1* plants (Cornuault et al., 2015, Urbanowicz et al., 2012). The results indicate that *gxm1/2/3* biomass releases a similar amount of both D-glucose and D-xylose when compared to WT AIR (Figure 3.1). Therefore, the data presented here does not support the hypothesis that GlcA methylation plays a significant role in the maintenance of biomass recalcitrance.

Data presented in this chapter suggests that out of all xylan branches in Arabidopsis, GlcA is the main determinant of *A. thaliana* biomass recalcitrance. Surprisingly, *gux* saccharification results reported previously indicate little to no improvement in sugar release upon genetic removal of GlcA branches (Lee et al., 2012). Lee et al. used cellulases on WT, *gux1/2* and *gux1/2/3* biomass and observed a non-significant, increase in D-glucose release only from *gux1/2* and *gux1/2/3* feedstock. This discrepancy is likely to be caused by the differences in the enzymatic cocktail used for saccharification in this study and the one used by Lee et al. In addition to that, it is unclear if Lee et al. used an acidic pre-treatment on the biomass. If so, this might have obscured the improvement in D-glucose release observed following saccharification of *gux1/2* without pre-treatment. Importantly, no other research evaluates the release of D-xylose from biomass with xylan lacking [Me]GlcA. The 7 fold increase in D-xylose release observed here is thus a novel and significant contribution to the biomass engineering field.

3.5.2 Loss of xylan-lignin but not the loss of xylan-cellulose interactions may be a reason for the reduced recalcitrance of *gux1/2* biomass.

The exact molecular reason for the improvement in sugar release observed in *gux1/2* plants remains unknown. Previously published analysis of *gux1/2* biomass monosaccharide composition has reported an increase in xylan content in *gux* stems (Lee et al., 2012, Chong et al., 2015). It was speculated that the excess UDP-GlcA remaining in the Golgi apparatus of cells lacking GUX activity may be converted to UDP-xylose by the activity of UDP-GlcA decarboxylase enzymes and used for xylan biosynthesis (Chong et al., 2015). Whereas some publications reported increased xylan content in *gux1/2* plants (Lee et al., 2012, Chong et al., 2015), other researchers have not observed any significant increase (Mortimer et al., 2010). This discrepancy might be caused by the differences in the biomass hydrolysis and monosaccharide analysis protocols. In any case, the maximal increase in xylan content reported in the literature amounted to ~60% which cannot be a sole explanation for the 7 fold increase in D-xylose release from *gux1/2* AIR.

The observed increase in the sugar yields in saccharification, together with improved xylan extractability, suggest that the absence of [Me]GlcA on xylan may lead to changes in the molecular architecture of cell walls. Cross-linking between cell wall components is one of the features that may be affected in *gux1/2* plants. In particular, the carboxylic group of GlcA is speculated to form ester linkages with the α or γ carbons of monolignols in β-O-4 linkages, forming structures known as Lignin-Carbohydrate Complexes (LCCs) (Takahashi and Koshijima, 1988, Giummarella and Lawoko, 2016). The abundance of these LCC structures would be greatly reduced in *gux1/2* plants, but their existence requires further verification. Feruloylation, a specific phenolic ester modification present on arabinose branches of grass xylan, has been confirmed to form covalent linkages with lignin. Interestingly, genetic decreases in the levels of xylan feruloylation by down-regulation of BAHD acyltransferases, likely to be responsible for feruloylation of arabinose prior to its transfer onto xylan, results in up to a 60% increase in D-glucose release in saccharification (de Souza et al., 2018). As ferulic acid might be involved in grass xylan-lignin crosslinking this data suggests that LCCs contribute to the maintenance of cell wall recalcitrance. Therefore, it is possible that the improvement in xylan extractability observed in *gux1/2* biomass and

associated increase in sugar release during saccharification are a direct result of the loss of xylan – lignin covalent linkages.

Changes in lignin and cellulose deposition patterns are another possible reason for the improvement in biomass processing observed in *gux1/2* biomass. Glucuronoxylan has been suggested to play a critical role in the helicoidal assembly of cell walls (Reis and Vian, 2004). This property of xylan is speculated to be enabled by the presence of negatively charged GlcA substitutions, which are absent in *gux1/2* mutants. Thus, it is possible that *gux* cell walls have a different mesoscale organisation which could involve alterations in cellulose clustering.

Loss of xylan-cellulose interactions is one of the possible reasons for the improved saccharification performance of *gux1/2* biomass. Although the binding of xylan to cellulose has not yet been investigated in the *GUX* mutants, changes that alter recalcitrance seem unlikely, as in silico, *gux 1/2* patterned xylan binds stably to cellulose (Busse-Wicher et al., 2014). However, the interaction of *gux1/2* xylan with cellulose remains an open question as 1,2-linked xylan substitutions, greatly reduced in *GUX* mutants, were proposed to stabilise the interaction (Pereira et al., 2017). The reduction of xylan acetylation in *esk1* plants affects GUX1 activity. This leads to loss of the compatible xylan [Me]GlcA patterning, which is proposed to be required for the establishment of normal xylan-cellulose interaction. Indeed, using a series of solid state NMR experiments, *esk1kak* xylan was observed to form a 3-fold screw and much of the xylan was not bound to the cellulose fibril (Grantham et al., 2017). As *esk1* biomass shows little to no improvement in sugar release in saccharification, when compared to WT plants, (Figure 3.1) it is highly likely that the normal two-fold xylan-cellulose interaction, which is lost in *esk1* plants, has little significance towards the maintenance of cell wall recalcitrance in the assay conditions used. Thus, it is unlikely to be the factor responsible for the improved performance of *gux* biomass in the saccharification assays.

3.6 Conclusion

Work reported in this Chapter indicates that the absence of [Me]GlcA on xylan results in a reduction in biomass recalcitrance. Thus, mutations in *GUX* genes are a possible strategy for the improvement of biomass properties for conversion into sugars for bioenergy and other industrial applications. Improvements in xylan extractability can contribute to sectors other than biofuel production. Paper manufacturing utilises harsh chemical and thermal treatments to remove xylan and lignin from plant biomass in order to obtain pure cellulose fractions (Bajpai, 2015). This harsh treatment might not be necessary if GUX-deficient biomass were to be used in the pulping process. Alternatively, more environmentally neutral processes, such as enzymatic bleaching, which uses enzymes instead of chemicals to degrade the hemicelluloses, might become more economically viable if *GUX* mutant feedstock is used instead of WT.

To facilitate the application of the discoveries described in this chapter it is critical to characterise GUX enzymes in industrially relevant plants. This will be facilitated by the use of genetic engineering approaches, such as CRISPR-Cas9 technology, to generate *GUX* knock-out organisms. Globally, softwood is the most significant timber biomass source (Ramage et al., 2017). Therefore, any improvements in its industrial processing could have a beneficial economic and environmental impact. Taking this into consideration, identification and characterisation of conifer GUX enzymes is a major focus of this thesis and is described in Chapters 4 and 5. . Despite the lack of growth phenotypes associated with the removal of [Me]GlcA in *A. thaliana* plants it is important to consider that reduction in biomass recalcitrance may influence performance of any modified biomass crops when challenged with pathogens. Influence of GUX mutations on plant pathogen response should therefore be evaluated in field trials.

Chapter 4: Identification and characterisation of conifer GUX enzymes

4.1 Introduction

The economic importance of softwood and coniferous forests was discussed extensively in chapter 1 of this thesis. Softwood contains predominantly cellulose, galactoglucomannan, xylan and lignin (Busse-Wicher et al., 2016b) with GGM being the dominant non-cellulosic polysaccharide and xylan accounting for up to 15% of cell wall material (Scheller and Ulvskov, 2010). As presented in chapter 3 of this thesis, the absence of [Me]GlcA branches on xylan results in significant reduction in recalcitrance of Arabidopsis biomass. The addition of these branches to xylan is catalysed by GUX enzymes. In secondary cell walls of *A. thaliana* GUX1 and GUX2 enzymes contribute to xylan glucuronidation (Mortimer et al., 2010, Bromley et al., 2013). Softwood xylan, similarly to hardwood xylan, does carry MeGlcA branches. However, softwood xylan is not acetylated and the MeGlcA branches are accompanied by arabinose substitutions, giving rise to molecules known as glucuronoarabinoxylan (GAX). Structural differences between hardwood and softwood xylan and its biosynthesis are discussed in section 1.7 of chapter 1.

The majority of softwood GAX has a simple pattern of MeGlcA branches with most decorations present on every 6th xylose residue (Busse-Wicher et al., 2016b). In addition to this regular MeGlcA spacing, some consecutive glucuronidation has also been detected in softwood (Shimizu et al., 1978, Martinez-Abad et al., 2017). Work performed as a part of this thesis has led to the identification and characterisation of a first conifer GUX enzyme from *Picea glauca* (Lyczakowski et al., 2017; Appendix 1). Taking this into consideration the work described in this chapter aimed to:

- Identify putative conifer GUX enzymes encoded in gymnosperm transcriptomes;

- Map the expression of conifer *GUX* genes to evaluate their role in softwood biosynthesis;

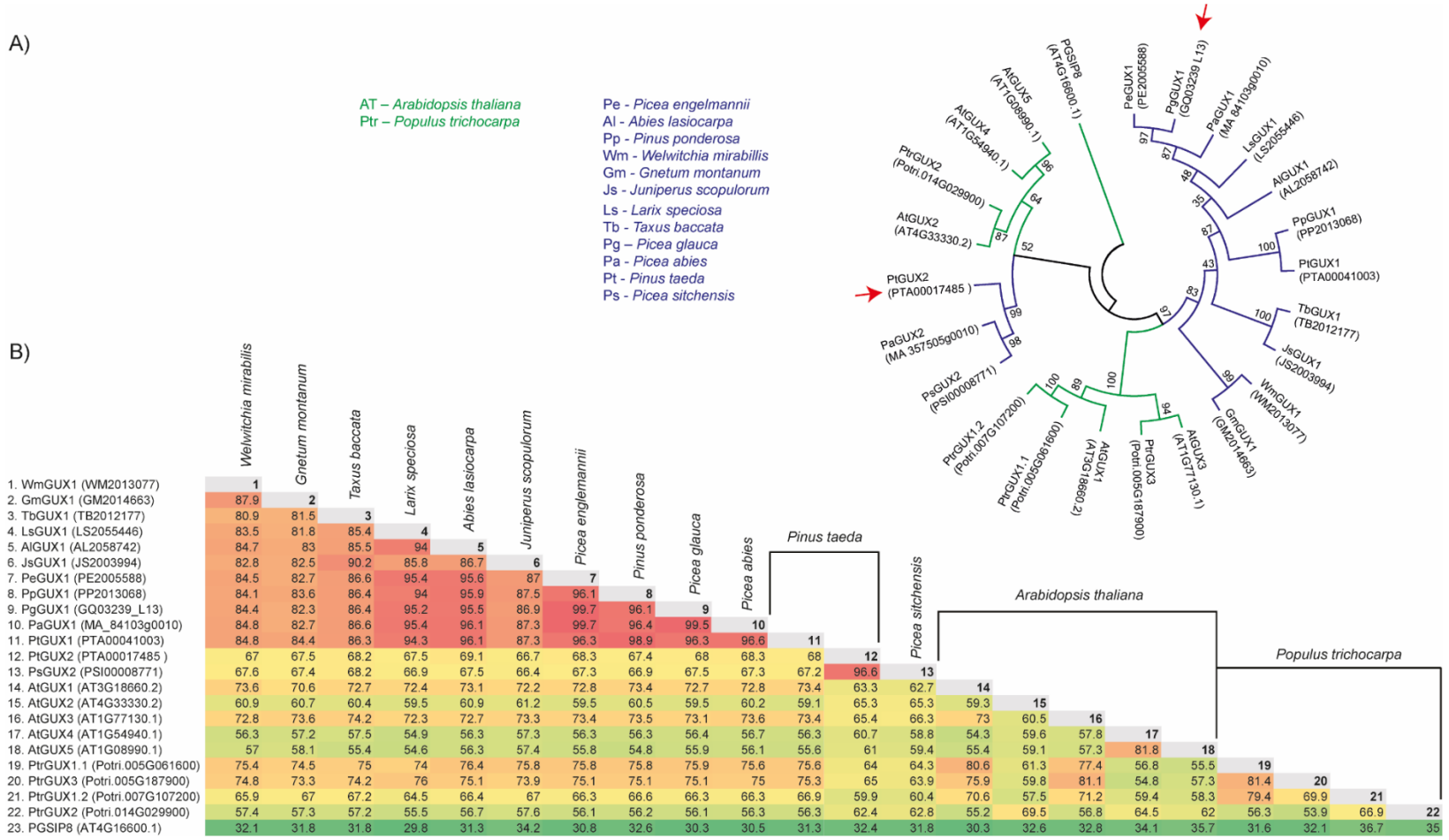
- Develop and apply a set of *in vitro* and *in vivo* assays to investigate if the putative conifer GUX enzymes are indeed active glucuronosyltransferases;

Identification and characterisation of conifer *GUX* genes may contribute to breeding or mutagenesis efforts aimed at the reduction of softwood recalcitrance. This may benefit industries such as paper production or biofuel manufacturing.

4.2 Conifer genomes may encode two distinct GUX enzymes

To identify putative conifer GUX enzymes both transcriptomic and genomic data was analysed. Before the start of this PhD project, Dr Krzysztof Wicher had identified, using Expressed Sequence Tags (ESTs), a putative GUX encoding transcript from White spruce (*Picea glauca,* sequence clone: GQ03239_L13, GeneBank code: BT111578.1). To identify other GUX sequences, the genomic data for Norway spruce (*Picea abies*) was analysed using the CONGENIE genome tool (Sundell et al., 2015). BLASTP analysis using the *A. thaliana* GUX1 sequence allowed identification of two genes: MA_84103g0010 and MA_35705g0010, which encode proteins sharing a high degree of sequence similarity to Arabidopsis GUX1 (72.7% and 69.4% respectively). To identify further putative gymnosperm GUX enzymes, the coding sequences for *P. glauca* and *P. abies GUX* genes were compared to translated gymnosperm transcriptomes available from the OneKP database using the TBLASTX tool (Matasci et al., 2014). OneKP provides data for over 50 different gymnosperm species and as such is one of the largest transcriptomic databases available for this taxon. This facilitates identification of a significant number of putative sequences which improves the quality of reconstructed phylogenies.

The analysis of OneKP data led to identification of multiple GUX enzymes from a range of coniferous and non-coniferous gymnosperm species. Similarly to Arabidopsis GUX enzymes the gymnosperm enzymes have an N-terminal cytosolic sequence, a transmembrane region and a GT domain localised close to the C terminus. The protein sequences for the putative gymnosperm GUX were reconstructed from identified transcripts using EXPASY and their Glycosyl Transferase (GT) domains were aligned together with these for *A. thaliana* and *Populus trichocarpa* GUXs. This alignment was used to reconstruct a maximum likelihood phylogeny (Figure 4.1A). The analysis suggested some of the gymnosperm sequences clustered away from the angiosperm ones. Interestingly, similarly to the hardwood GUX enzymes, the gymnosperm GUX GT sequences may be clustering into two clades. The presence of the two distinct conifer GUX clades is supported by the full length protein sequence similarity analysis (Figure 4.1B), which indicates a high (>80%) sequence similarity within each clade and a lower (<70%) sequence similarity between the clades. To follow the notation previously proposed for hardwood, representatives of each monophyletic softwood GUX group were denoted as GUX1 and GUX2 to indicate their clade membership.



	Welwitchia mirabilis	Gnetum montanum	Taxus baccata	Larix speciosa	Abies lasiocarpa	Juniperus scopulorum	Picea englemannii	Pinus ponderosa	Picea glauca	Picea abies	Pinus taeda	Picea sitchensis	Arabidopsis thaliana						Populus trichocarpa			
1. WmGUX1 (WM2013077)	1																					
2. GmGUX1 (GM2014663)	87.9	2																				
3. TbGUX1 (TB2012177)	80.9	81.5	3																			
4. LsGUX1 (LS2055446)	83.5	81.8	85.4	4																		
5. AlGUX1 (AL2058742)	84.7	83	85.5	94	5																	
6. JsGUX1 (JS2003994)	82.8	82.5	90.2	85.8	86.7	6																
7. PeGUX1 (PE2005588)	84.5	82.7	86.6	95.4	95.6	87	7															
8. PpGUX1 (PP2013068)	84.1	83.6	86.4	94	95.9	87.5	96.1	8														
9. PgGUX1 (GQ03239_L13)	84.4	82.3	86.4	95.2	95.5	86.9	99.7	96.1	9													
10. PaGUX1 (MA_84103g0010)	84.8	82.7	86.6	95.4	96.1	87.3	99.7	96.4	99.5	10												
11. PtGUX1 (PTA00041003)	84.8	84.4	86.3	94.3	96.1	87.3	96.3	98.9	96.3	96.6	11											
12. PtGUX2 (PTA00017485)	67	67.5	68.2	67.5	69.1	66.7	68.3	67.4	68	68.3	68	12										
13. PsGUX2 (PSI00008771)	67.6	67.4	68.2	66.9	67.5	66.4	67.3	66.9	67.5	67.3	67.2	96.6	13									
14. AtGUX1 (AT3G18660.2)	73.6	70.6	72.7	72.4	73.1	72.2	72.8	73.4	72.7	72.8	73.4	63.3	62.7	14								
15. AtGUX2 (AT4G33330.2)	60.9	60.7	60.4	59.5	60.9	61.2	59.5	60.5	59.5	60.2	59.1	65.3	65.3	59.3	15							
16. AtGUX3 (AT1G77130.1)	72.8	73.6	74.2	72.3	72.7	73.3	73.4	73.5	73.1	73.6	73.4	65.4	66.3	73	60.5	16						
17. AtGUX4 (AT1G54940.1)	56.3	57.2	57.5	54.9	56.3	57.3	56.3	56.3	56.4	56.7	56.3	60.7	58.8	54.3	59.6	57.8	17					
18. AtGUX5 (AT1G08990.1)	57	58.1	55.4	54.6	56.3	57.4	55.8	54.8	55.9	56.1	55.6	61	59.4	55.4	59.1	57.3	81.8	18				
19. PtrGUX1.1 (Potri.005G061600)	75.4	74.5	75	74	76.4	75.8	75.8	75.8	75.9	75.6	75.6	64	64.3	80.6	61.3	77.4	56.8	55.5	19			
20. PtrGUX3 (Potri.005G187900)	74.8	73.3	74.2	76	75.1	73.9	75.1	75.1	75.1	75	75.3	65	63.9	75.9	59.8	81.1	54.8	57.3	81.4	20		
21. PtrGUX1.2 (Potri.007G107200)	65.9	67	67.2	64.5	66.4	67	66.3	66.6	66.3	66.3	66.9	59.9	60.4	70.6	57.5	71.2	59.4	58.3	79.4	69.9	21	
22. PtrGUX2 (Potri.014G029900)	57.4	57.3	57.2	55.5	56.7	57.6	56.1	56.2	56.1	56.3	56.3	62.4	62.8	55.2	69.5	56.8	64.5	62	56.3	53.9	66.9	22
23. PGSIP8 (AT4G16600.1)	32.1	31.8	31.8	29.8	31.3	34.2	30.8	32.6	30.3	30.5	31.3	32.4	31.8	30.3	32.6	32.8	34.1	35.7	31.6	32.1	36.7	35

Figure 4.1 Phylogenetic analysis of angiosperm and gymnosperm GUX enzymes. Maximum likelihood phylogeny of the GT sequences of gymnosperm (blue) and angiosperm (green) GUX (A). Low bootstrap score (52) for the branch separating the PtGUX2 clade away from the angiosperm sequences indicates that this region of the phylogenetic tree might not be accurately represented. PgGUX1 and PtGUX2 studied as a part of this thesis are marked with a red arrow. Similarity matrix analysis for the gymnosperm and angiosperm GUX enzymes (B). Similarity score is indicated for each pair. Red colour denotes highest similarity. Pairs with lowest similarity are marked in green. PGSIP8, an Arabidopsis GT8 not in the GUX clade, was used as an outgroup.

4.3 Representatives of both conifer GUX clades are likely to be involved in wood biosynthesis.

To determine the likely biological role of the newly identified putative conifer GUX enzymes, expression profiling was performed using data available through the CONGENIE and Norwood services (Jokipii-Lukkari et al., 2017). The ExHeatmap tool, from the CONGENIE server, was used to profile the expression of PaGUX1 (MA_84103g0010, Figure 4.2) and PaGUX2 (MA_35705g0010, Figure 4.3). Reads encoding both PaGUX1 and PaGUX2 enzymes are enriched in a wood sample made from phloem, cambium and xylem of earlywood and latewood. High expression levels were also detectable in developing shoots. The exHeatmap tool provides only an overview of gene expression, a more specific analysis can be performed using the Norwood service. This tool allows for mapping of conifer gene expression across the stem, from cambium to mature xylem vessels. Importantly, the Norwood data is reported for three individuals and expression profiles for genes of interest can be placed in a biological context by comparing them with the expression of other genes likely to be involved in known biosynthetic pathways.

To further evaluate the biological role of putative conifer GUX enzymes, expression of *PaGUX1* and *PaGUX2*, together with the expression of other putative cell wall biosynthesis genes, was studied using the Norwood tool (Figure 4.4). Genes encoding proteins likely to be involved in biosynthesis of other cell wall components in Spruce were identified using CONGENIE BLASTP tool by looking for highest amino acid sequence similarity to Arabidopsis enzymes. Results of this analysis indicate a high degree of similarity in the expression profiles of genes involved in the same biosynthetic pathways. For example, genes MA_10429177g0010, MA_140410g0010 and MA_183130g0010, likely to encode three subunits of the secondary cell wall cellulose synthase, are co-expressed across the three biological replicates. More broadly, most of the studied cell wall biosynthesis genes share a similar expression profile, with high expression predominantly in the developing xylem vessels. Importantly, this xylem specific expression profile, which can be associated with gene's function in secondary cell wall biosynthesis, is also observed for both *PaGUX1* and *PaGUX2*. Therefore, the Norwood analysis supports the hypothesis that both putative conifer GUX enzymes are involved in softwood xylan biosynthesis.

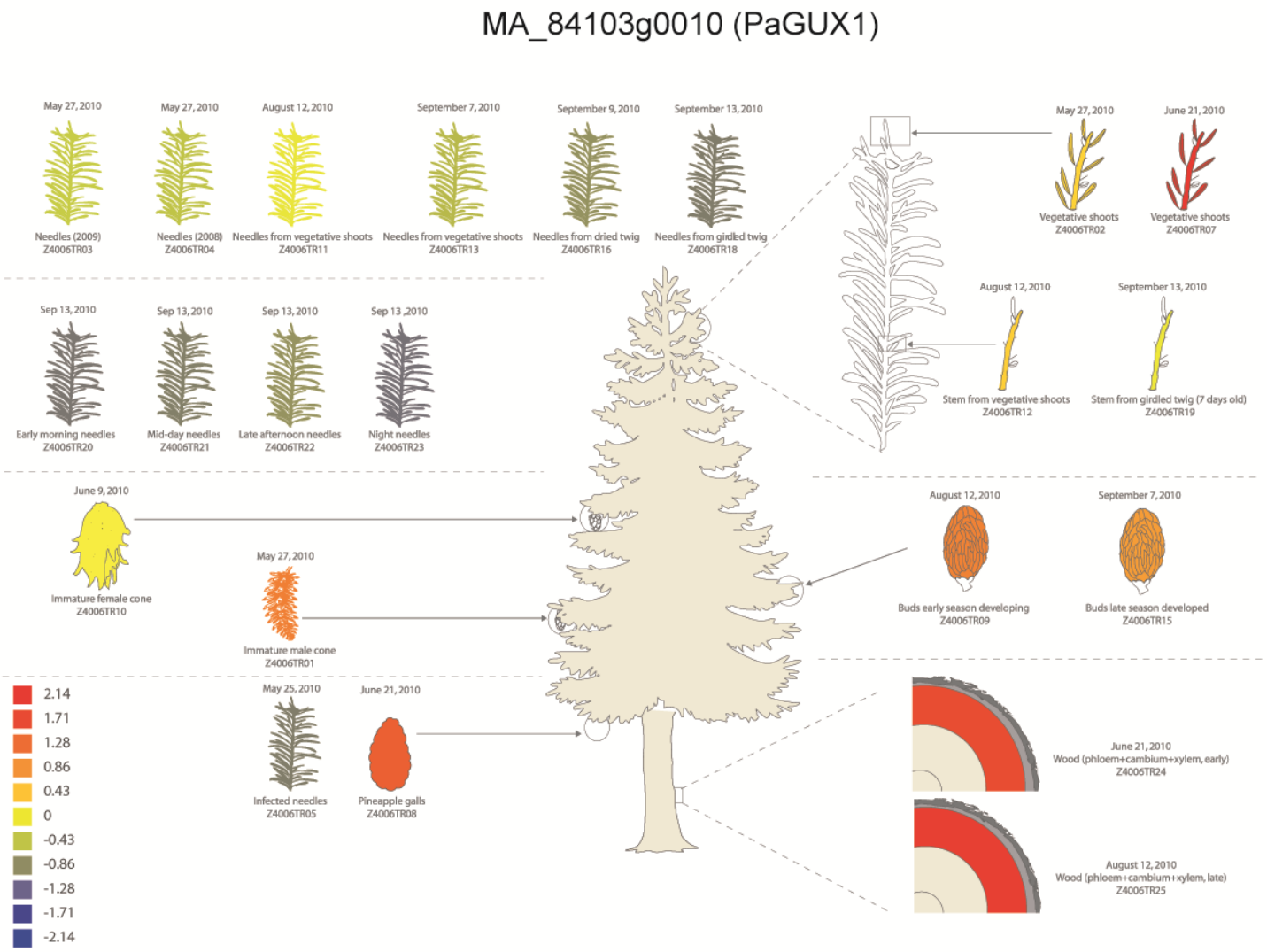


Figure 4.2 Expression profile of putative conifer GUX1. Expression profiling of PaGUX1 using the exHeatmap tool. Read enrichment is marked with red whereas read depletion is marked with blue.

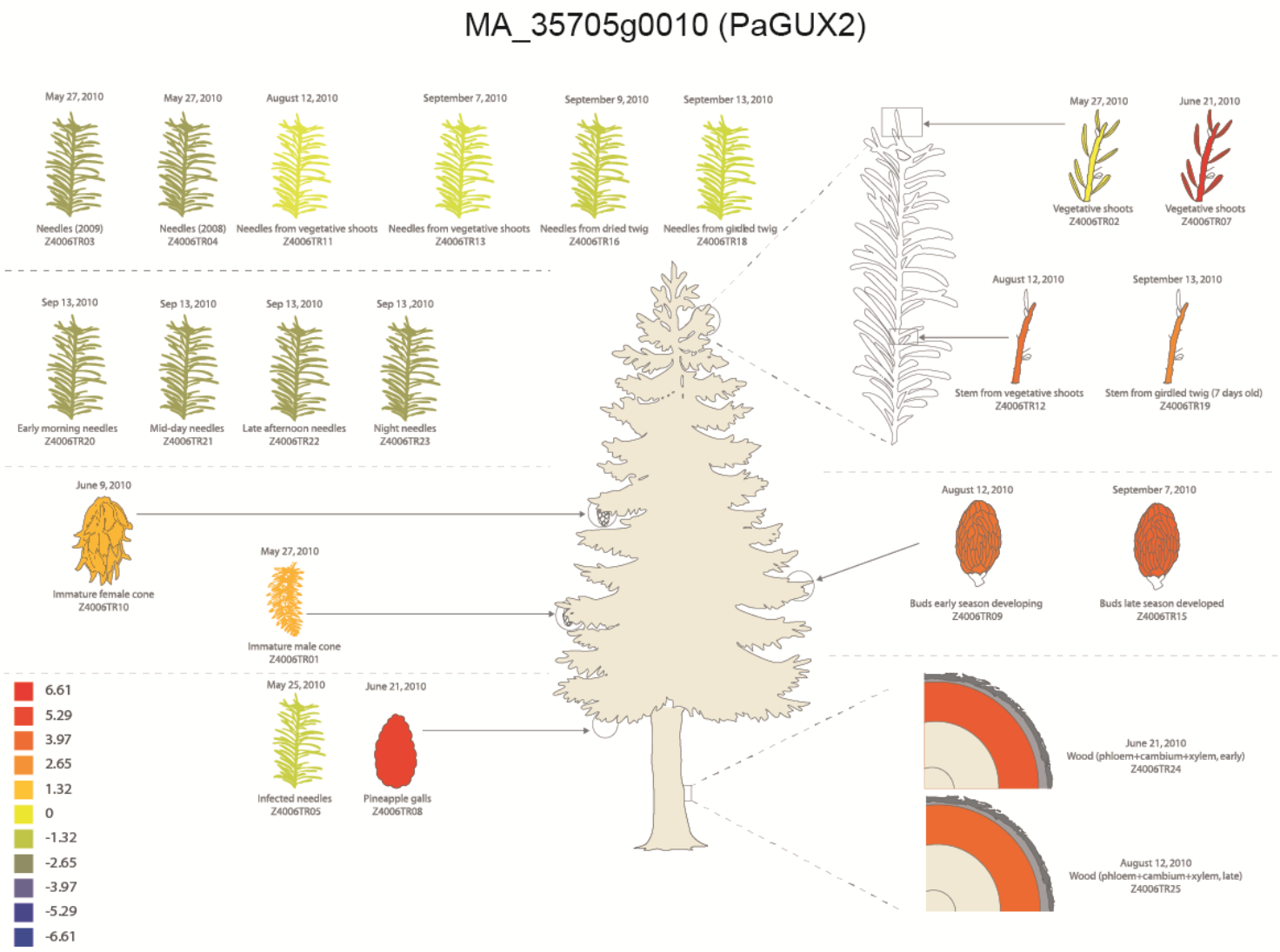
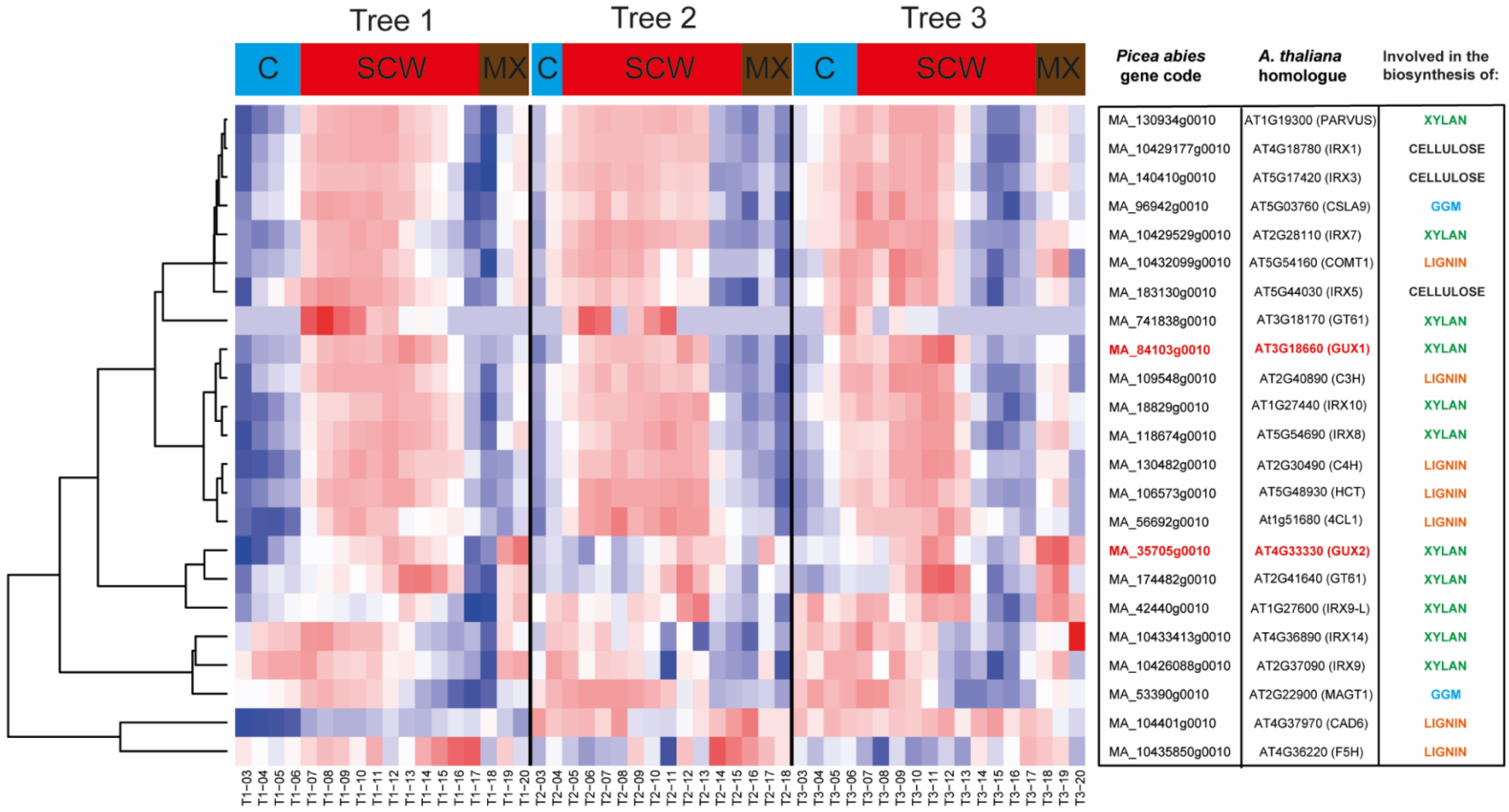


Figure 4.3 Expression profiles of putative conifer GUX2. Expression profiling of PaGUX2 using the exHeatmap tool. Read enrichment is marked with red whereas read depletion is marked with blue.



Picea abies gene code	*A. thaliana* homologue	Involved in the biosynthesis of:
MA_130934g0010	AT1G19300 (PARVUS)	XYLAN
MA_10429177g0010	AT4G18780 (IRX1)	CELLULOSE
MA_140410g0010	AT5G17420 (IRX3)	CELLULOSE
MA_96942g0010	AT5G03760 (CSLA9)	GGM
MA_10429529g0010	AT2G28110 (IRX7)	XYLAN
MA_10432099g0010	AT5G54160 (COMT1)	LIGNIN
MA_183130g0010	AT5G44030 (IRX5)	CELLULOSE
MA_741838g0010	AT3G18170 (GT61)	XYLAN
MA_84103g0010	AT3G18660 (GUX1)	XYLAN
MA_109548g0010	AT2G40890 (C3H)	LIGNIN
MA_18829g0010	AT1G27440 (IRX10)	XYLAN
MA_118674g0010	AT5G54690 (IRX8)	XYLAN
MA_130482g0010	AT2G30490 (C4H)	LIGNIN
MA_106573g0010	AT5G48930 (HCT)	LIGNIN
MA_56692g0010	At1g51680 (4CL1)	LIGNIN
MA_35705g0010	AT4G33330 (GUX2)	XYLAN
MA_174482g0010	AT2G41640 (GT61)	XYLAN
MA_42440g0010	AT1G27600 (IRX9-L)	XYLAN
MA_10433413g0010	AT4G36890 (IRX14)	XYLAN
MA_10426088g0010	AT2G37090 (IRX9)	XYLAN
MA_53390g0010	AT2G22900 (MAGT1)	GGM
MA_104401g0010	AT4G37970 (CAD6)	LIGNIN
MA_10435850g0010	AT4G36220 (F5H)	LIGNIN

Figure 4.4. Expression profiles of conifer genes likely to be involved in softwood biosynthesis. Expression profiling of *Picea abies* genes likely to be involved in the biosynthesis of cellulose, galactoglucomannan (GGM), xylan and lignin across cambium (C), secondary cell wall formation zone (SCW) and mature xylem/latewood (MX). Read enrichment is marked with red and read depletion is marked with blue. Most *P. abies* candidates (column 1 of the table) were identified based on having the highest homology score to *A. thaliana* enzymes (column 2). Due to their potential role in xylan branching two representatives were selected for putative xylan glucuronosyltransferases (GUX, marked in red) and putative arabinosyltransferases (GT61). Analysis reports data for three trees.

As the Norwood tool provides data only on the cambium and xylem specific expression for the genes of interest, further analysis is required to evaluate the possible role of conifer GUX1 and GUX2 in phloem or epidermal cell wall synthesis. To achieve this level of information, *in situ* hybridisation experiments were performed on young stems of *P. glauca.* Probes, corresponding to 1000 bp of the 3' end of *PgGUX1* and *PaGUX2* CDS, were ordered as synthetic DNA. *In vitro* transcription to generate the DIG labelled probes and *in situ* hybridisation was performed by Dr. Weibing Yang (Sainbury Lab, University of Cambridge). Following Anti-DIG probing with an AP conjugated antibody, 8 μm thick sections were visualised using a light microscope (Figure 4.5). The read-out in this *in situ* hybridisation is based on AP catalysed reduction of NBT. Therefore, the positive signal, indicating probe binding to mRNA, is associated with the development of a purple colour in the cells.

For both *Pg*GUX1 and *Pa*GUX2 high levels of signal were observed mostly from vascular cambium cell layer (Meents et al., 2018). Interestingly, unlike the Norwood results, the *in situ* data does not indicate strong expression of *PgGUX1* and *PaGUX2* genes in mature xylem tracheids. It is possible this may be resulting from the sensitivity of the *in situ* hybridisation experiment. *In situ* hybridisation requires a significant number of reads and probe penetration into the tissue thus its read-out may be limited by read abundance and accessibility (Yang et al., 2016). While in Norwood analysis only different layers of tracheids were evaluated, the *in situ* analysis looked globally at the entire stem. Given the limitations of the technique it is possible that the probe binding was limited to only certain parts of the cross-section. Alternatively, levels of *PgGUX1* and *PaGUX2* may be highest in vascular cambium which would not be surprising as this tissue is associated with polysaccharide biosynthesis (Meents et al., 2018). The in situ hybridisation experiments need to be repeated with the inclusion of controls including anti-sense probes and probes against genes expressed in other tissues.

In summary, considering limitations and advantages of all techniques used to probe expression of both conifer GUX enzymes, it is possible they are expressed predominantly in tissues associated with wood formation. This includes either vascular cambium or developing xylem vessels. Therefore, it is possible that *Pg*GUX1 and *Pa*GUX2 may be involved in the process of softwood xylan biosynthesis.

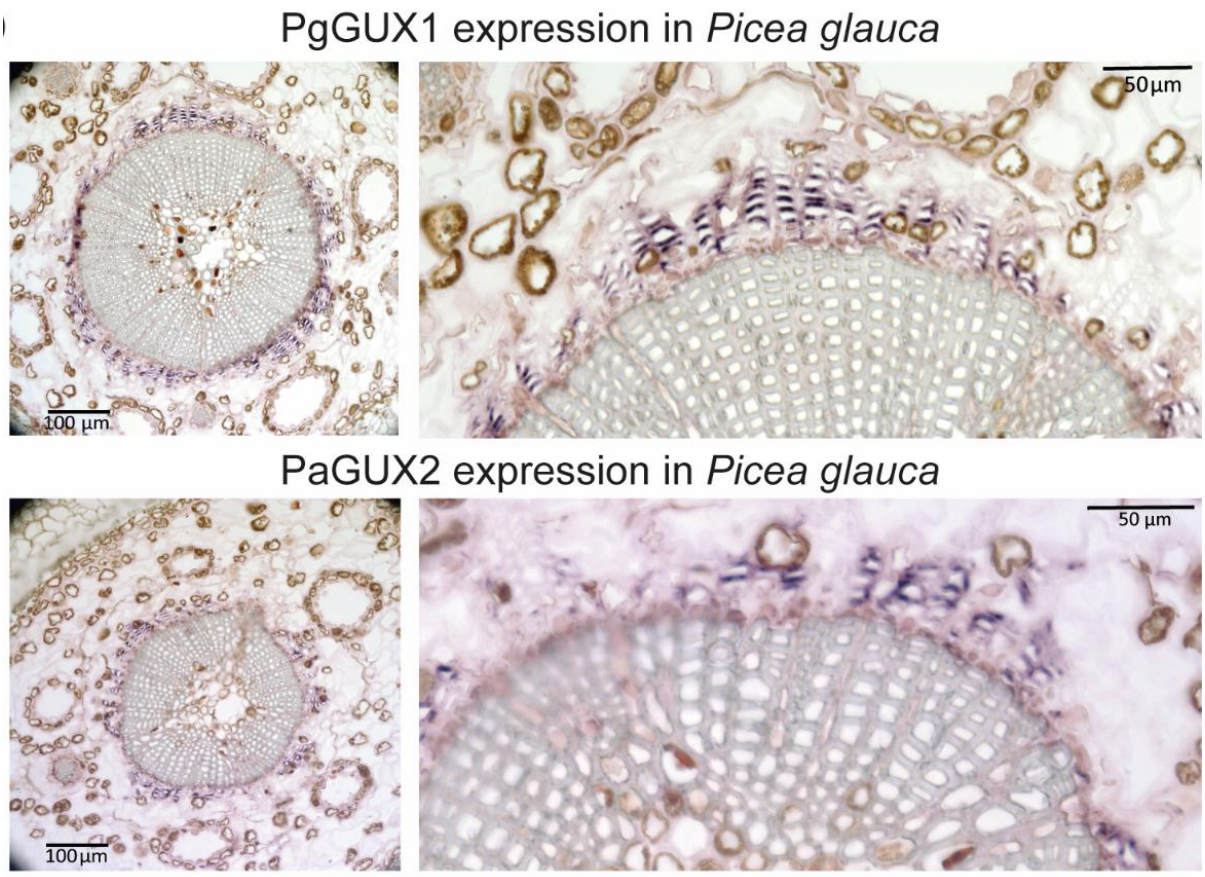


Figure 4.5 *In situ* hybridisation analysis of PgGUX1 and PaGUX2 expression in young stems of *Picea glauca*. Development of purple colouring is visible mostly in vascular cambium. Weaker signal, associated with pink staining, is observed from other cell types. Resin ducts contain strong brown pigmentation. Size bars are provided for reference. This experiment was performed only once.

4.4 Both PgGUX1 and PaGUX2 can be successfully produced using the tobacco pEAQ-HT expression system.

To evaluate the biochemical activity of conifer GUX enzymes, the CDS encoding PgGUX1 and PtGUX2, together with a C-terminal Myc tag sequence, were cloned into NruI digested pEAQ-HT vector. PgGUX1 and PtGUX2 are homologues of PaGUX1 and PaGUX2 from *Picea glauca* and *Pinus taeda* respectively. These genes were selected as, unlike for PaGUX enzymes, one could reconstruct a complete protein sequences from OneKP data for them. The pEAQ-HT vector family contains a 35S promoter together with 5' and 3' UTR viral enhancers and allows high levels of protein expression in the *Nicothiana benthamiana* leaf system (Sainsbury et al., 2009). Individual AGL-1 *Agrobacterium tumefaciens* strains transformed with pEAQ-

HT_PgGUX1-Myc, pEAQ-HT_PtGUX2-Myc and pEAQ-HT_AtMAGT1-Myc constructs were infiltrated into *N. benthamiana* (Figure 4.6A). AtMAGT1 encodes a galactoglucomannan galactosyltransferase (Yu et al., 2018) and was expressed to act as a negative control for any endogenous tobacco glucuronosyltransferase activity in the *in vitro* activity experiments.

Following 4 days of incubation, tobacco leaves were harvested and a membrane fraction was extracted to assay for the presence of the Golgi-localised glycosyltransferases. Membrane protein samples were separated on a SDS-PAGE gel and transferred onto a nitrocellulose membrane. An anti-Myc western blot was performed to evaluate the presence of PgGUX1-Myc, PtGUX2-Myc and AtMAGT1-Myc in the leaf protein extracts (Figure 4.6B). A band of a size corresponding to the Myc tagged proteins of interest was observed for all extracts. As previously reported a possible dimer band was observed for AtMAGT1 (Yu et al., 2018). No signal was detected for the sample from leaves infiltrated with the empty vector. These results indicate that the pEAQ-HT system, together with the membrane purification protocol can be used to obtain protein fractions containing the expressed conifer GUX enzymes and the AtMAGT1 control protein. These membrane fractions were used as crude enzyme preparations in the conifer GUX *in vitro* activity assays.

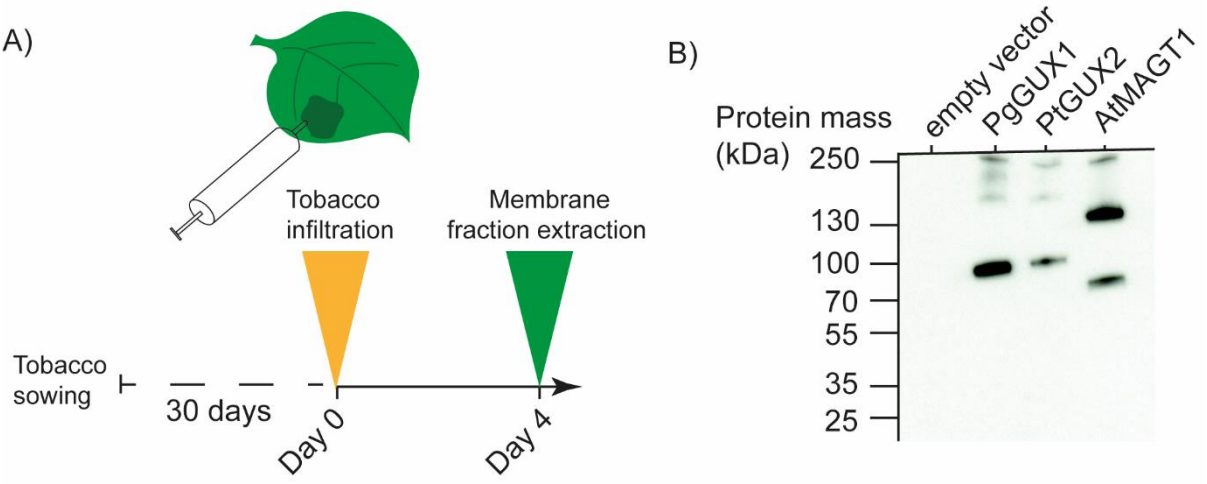


Figure 4.6 Transient tobacco expression using the pEAQ-HT system. Experimental scheme used in the transient expression experiments (A). Western Blot analysis of *N. benthamiana* membrane fraction extracted from leaves expressing PgGUX1, PtGUX2 and AtMAGT1 (B). Membranes extracted from leaves infiltrated with an empty vector were used as a control. Expected protein sizes: PgGUX1-3xMyc: 78.2 kDa; PtGUX2-3xMyc: 75.4 kDa; AtMAGT1-3xMyc: 59.1 kDa. Higher than expected protein mass may be due to glycosylation of GUX enzymes.

4.5 PgGUX1 is an active glucuronosyltransferase *in vitro*.

To evaluate if the PgGUX1 enzyme has any glucuronosyltransferase activity *in vitro,* the tobacco membrane fraction enriched for the protein was incubated with intact polymeric xylan without [Me]GlcA branches, extracted from *gux1/2* plants according to (Busse-Wicher et al., 2014), in the presence of UDP-GlcA. Since this xylan is insoluble without acetylation, the acetate branches were maintained on the acceptor throughout the *in vitro* reaction. Control reactions without UDP-GlcA and using AtMAGT1 were also performed. Following a 5h incubation period the reaction products were purified with methanol: chloroform and ethanol washes and de-acetylated with alkali treatment. This was followed by PACE analysis using xylanase GH11 digestion (Figure 4.7). Xylanase GH11 can digest unsubstituted xylan to xylobiose and xylose. Presence of [Me]GlcA branches inhibits GH11 digestion and gives rise to a [Me]GlcA-xylotetraose (UX_4) product. UX_4 can also be represented as XUXX to denote the position of the [Me]GlcA branch in the oligosaccharide. In this assay, UX_4 was released primarily from the reaction in which the PgGUX1 enzyme was incubated with the acetylated *gux1/2* xylan in the presence of UDP-GlcA. This result indicates that PgGUX1 is an active conifer glucuronosyltransferase *in vitro*. A minor amount of UX_4 was produced in the reaction using AtMAGT1 negative control indicating that a small proportion of activity detected in the PgGUX1 catalysed reaction may be endogenous to tobacco microsomes.

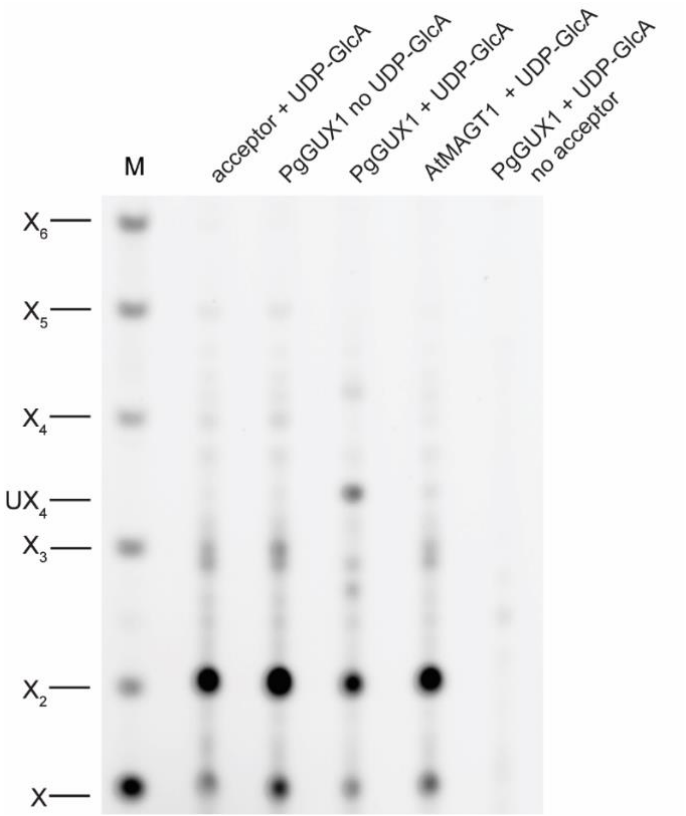


Figure 4.7 PgGUX1 has xylan glucuronosyltransferase activity *in vitro*. Products of the in vitro glucuronosylation were digested with xylanase GH11 and analysed by PACE. The enzyme generates xylose, xylobiose, plus the UX4 oligosaccharide if any GlcA is present on the xylan.

To evaluate the maximum amount of UDP-GlcA transfer on xylan by PgGUX1, the *in vitro* activity reaction was performed across a range of UDP-GlcA concentrations. Between 0.05 mM and 10 mM UDP-GlcA was used in the same assay conditions which utilise acetylated *gux1/2* xylan as an acceptor. Following incubation, the reaction products were extracted with ethanol and deacetylated with an alkali treatment. This was followed by a xylanase GH11 digestion and PACE analysis (Figure 4.8A). The amount of the XUXX oligosaccharide increased with the amount of UDP-GlcA used in the reaction. Percentage glucuronidation was quantified across the experiments by integrating the volume of peaks corresponding to UX_4, X2 and X structures. This was plotted against the concentration of UDP-GlcA used (Figure 4.8B). The reaction using 5 mM UDP-GlcA was repeated three times using one batch of microsomes. Therefore, the variation in the degree of glucuronidation obtained for this condition likely represents differences in experimental set-up across reaction replicates.

Once plotted, the data points were joined to generated a curve which enabled quantitation of the predicted maximal degree of glucuronation and the [UDP-GlcA]

required to achieve half the maximal amount transferred (Figure 4.8C). As in the assay conditions used the acceptor is unlikely to be glucuronidated further in any particular reaction these parameters are not K_M or V_{max} values reported for GUX enzymes previously (Rennie et al., 2012). However, the derived parameters can be used to compare the maximal amount of UDP-GlcA transferred by PgGUX1 with the other GUX enzymes studied in this project. Moreover, the data reported by Rennie et al., can be recalculated to obtain parameters derived for PgGUX1 in this study to evaluate the maximal degree of GlcA transferred by GUX enzymes in different reaction conditions. This calculation will be presented as a part of the discussion at the end of this chapter.

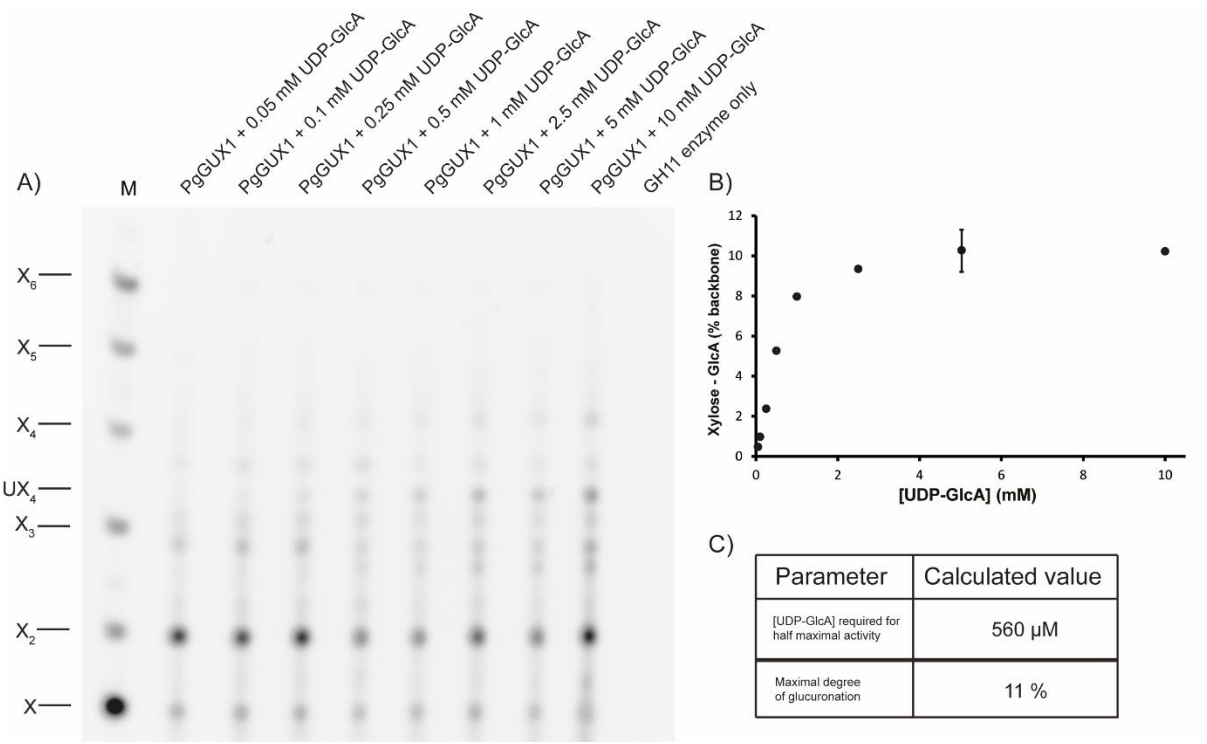


Figure 4.8 PgGUX1 activity across a range of UDP-GlcA concentrations. Products of *in vitro* glucuronosylation reaction with PgGUX1 and a range of [UDP-GlcA] were digested with xylanase GH11 and analysed by PACE (A). Degree of xylan glucuronidation for individual reactions (B). Reaction using 5 mM UDP-GlcA was repeated three times, error bars represent standard deviation of the glucuronidation degree obtained across the reactions. Activity parameters for PgGUX1 (C) quantified by non-linear regression analysis of the plotted data.

4.6 PgGUX1 can add biologically functional GlcA branches onto xylan *in planta.*

Having established that the PgGUX1 is an active glucuronosyltransferase *in vitro* the activity of the enzyme was tested using an *in vivo* assay. The coding sequence of PgGUX1 together with a C terminal Myc tag was placed under the control of the secondary cell wall cellulose synthase *IRX3* promoter and transformed into *gux1/2/3* mutant *A. thaliana* plants by Dr Marta Busse-Wicher. The investigation of the three homozygous lines (PgGUX1 lines) resulting from independent insertional events of the transgene into *A. thaliana* genome was carried out as a part of this PhD project. Chapter 2 describes the protocols used to identify homozygous lines in these and other transformation experiments performed as a part of this PhD project.

The use of *gux1/2/3 A. thaliana* plants as the over-expression background ensures lack of [Me]GlcA in both primary (*gux3*,(Mortimer et al., 2015)) and secondary (*gux1/2,* (Mortimer et al., 2010)) cell walls. Therefore, any [Me]GlcA branches on xylan detected in the transformed plants would result from PgGUX1 activity. To investigate the presence of [Me]GlcA branches on xylan, the Alcohol Insoluble Residues (AIR) made from basal stem sections of WT, *gux1/2/3* and three homozygous lines of PgGUX1 plants were digested with xylanase GH11 and analysed using PACE (Figure 4.9A). Basal stem sections from between 40 and 50 individual plants were used for each AIR preparation. Interestingly, the UX_4 band, indicative of xylan glucuronidation, was detected in WT plants and across all three PgGUX1 lines. The UX_4 oligosaccharide was not released from *gux1/2/3* AIR, confirming the lack of detectable glucuronidation in the over-expression background. The experiment was repeated for three independent matching biological replicates of AIR from WT, *gux1/2/3* and PgGUX1.1 to PgGUX1.3 plants and the average degree of xylan glucuronidation for each plant genotype studied was quantified by integrating the volume of peaks corresponding to UX_4, X2 and X structures (Figure 4.9B). Together, these data indicate that the cloned PgGUX1 enzyme is an active conifer glucuronosyltransferase *in vivo*. Interestingly, although all PgGUX1 transformants did have some xylan glucuronidation, its exact degree varied across the lines. This variation is likely a reflection of the level of transgene expression due to its genomic location.

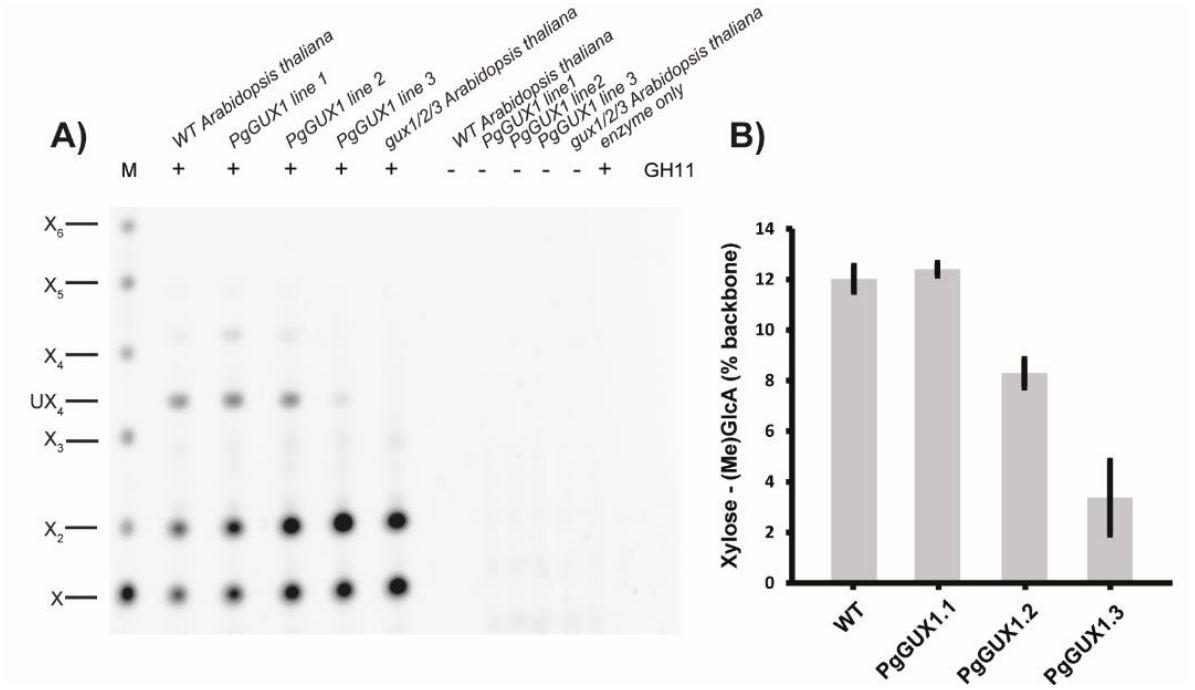


Figure 4.9 PgGUX1 is a functional xylan glucuronosyltransferase. PACE analysis of GH11 xylanase digests of WT, three independent transgenic lines of PgGUX1 in *gux1/2/3* and control *gux1/2/3* AIR from basal stems (A). Undigested AIR controls are marked with (−). The [Me]GlcA-xylotetraose band (UX4) was observed only in WT and PgGUX expressing lines. Quantitation of the degree of [Me]GlcA substitutions (B). Error bars represent standard deviation of three biological replicates of plant biomass.

Reduction of biomass recalcitrance is the only clear plant phenotype so far detected to be associated with the removal of xylan glucuronidation (chapter 3). To investigate if the [Me]GlcA added by the PgGUX1 enzyme is biologically functional, a saccharification assay was performed on WT, *gux1/2/3*, PgGUX1.1 and PgGUX1.2 AIR. Following incubation with Cellic® CTec2, the release of D-glucose (Figure 4.10A) and D-xylose (Figure 4.10B) was quantified. Similarly to analysis described in chapter 3 of this thesis, no additional chemical pre-treatment, other than AIR preparation, was used on the plant biomass prior to saccharification. Similarly to results obtained for *gux1/2* biomass (chapter 3) the *gux1/2/3 A. thaliana* AIR released significantly more D-glucose and D-xylose than the WT material. In both PgGUX1 lines studied, the monosaccharide release was reduced to WT levels, indicating that the reintroduced glucuronidation of the xylan backbone is able to restore recalcitrance of cell walls as effectively as native decorations. Thus, in this assay, branches added by the PgGUX1 enzyme are likely to play the same biological function as decorations introduced by the native Arabidopsis enzymes.

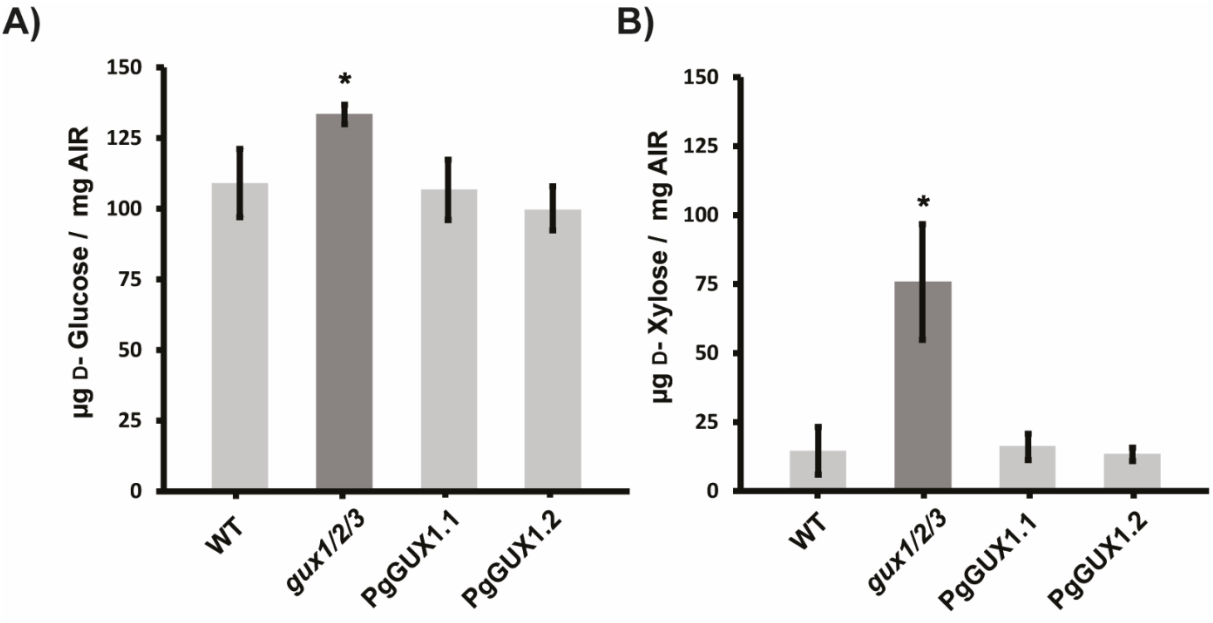


Figure 4.10 PgGUX1 saccharification analysis. D-glucose and D-xylose release following saccharification of WT, *gux1/2/3* and two lines of PgGUX1 AIR. Error bars represent standard deviation of saccharification assays of three biological replicates, *p value ≤0.05 in Student's t-test.

4.7 PtGUX2 is an active glucuronosyltransferase *in vitro* and can synthesise an unusual pattern of GlcA branches.

To investigate the activity of conifer enzymes clustering in the second clade distinct from the clade with PgGUX1 and highly similar enzymes, the *N. benthamiana* membranes enriched for the PtGUX2 enzyme were used in the *in vitro* GlcAT activity assay. Following a 5h incubation period, the xylan acceptor was purified with ethanol, de-acetylated with an alkali treatment and digested with xylanase GH11. Oligosaccharides released from reactions performed using PtGUX2 and the control protein AtMAGT1 were analysed with PACE (Figure 4.11A). The reaction in which the PtGUX2 enzyme was incubated with the acetylated *gux1/2* xylan acceptor in the presence of UDP-GlcA was the only one from which a significant amount of UX_4 oligosaccharide was released. Therefore, similarly to PgGUX1, the PtGUX2 enzyme is an active conifer glucuronosyltransferase *in vitro*.

Interestingly, an additional band (Band 1 on Figure 4.11A), not detected in PgGUX1 reactions, was present when the PtGUX2 enzyme was used in the *in vitro* assay. Since it migrates distinctly from UX_4, it is likely to have a different structure, such as additional GlcA and/or Xylosyl residues. To investigate the identity of Band 1 and to confirm

successful *in vitro* glucuronidation with PtGUX2 even further, α-glucuronidase GH115 treatment of reaction products containing UX4 and Band 1 was performed (Figure 4.11B). Prior to GH115 digestion xylanase GH11 was inactivated. Thus, any de-glucuronated oligosaccharide resulting from GH115 digestion will not be broken down to xylobiose and xylose, and should migrate exactly as markers corresponding to their backbone size. The GH115 digestion of PtGUX2 reaction products resulted in the release of X5 and X4. As X4 is a result of deglucuronidation of UX_4, Band 1 is also glucuronidated, and contains 5 xylose units in the backbone.

To further evaluate the identity of Band 1, the *in vitro* activity assay was repeated and the oligosaccharides released by xylanase GH11 from the PtGUX2 reaction were derivatised with 2-aminobenzoic acid and analysed using MALDI-ToF mass spectrometry (MS) (Figure 4.11C). The 866 *m/z* ion, corresponding to the mass of sodiated 2-AA labelled UX_4 structure was detected, confirming the successful GlcA transfer onto xylan by the PtGUX2 enzyme. A 744 m/z peak was also observed. This signal is consistent with the mass of the UX_3 structure. This was not expected, but may result from misdigestion by GH11 or xylanase GH10 contamination of the glycoside hydrolase preparation.

In addition to these expected MS signals, a pair of peaks of *m/z* equal to 1042 and 1174 was also detected. These correspond to mass of U_2X_4 and U_2X_5 oligosaccharides respectively. Similarly to the UX_4 and UX_3 structures, they differ by a single xylosyl unit. Thus, it is likely that the m/z 1042 peak is a result of GH11 miss-digestion or xylanase GH10 contamination. Canonical xylanase GH11 digestion requires a single unsubstituted xylose on position -1 and two unsubstituted xyloses on positions +1 and +2 from the branched monomer. Thus, in order to release a xylopentaose oligosaccharide with two GlcA substitutions (the U_2X_5, m/z 1174 structure) it is likely that consecutive xylosyl residues have the acidic substitution.

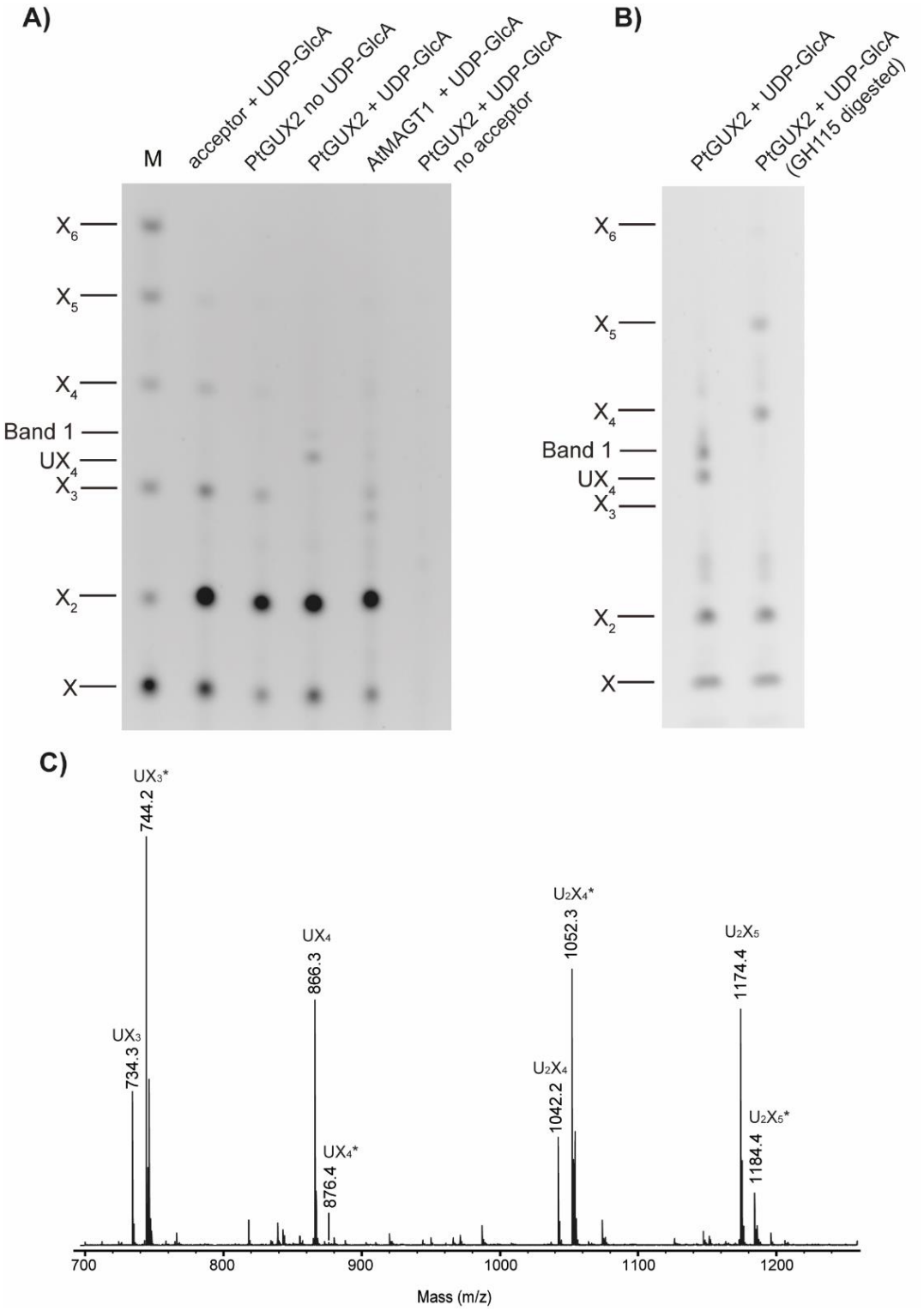


Figure 4.11 PtGUX2 has xylan glucuronosyltransferase activity *in* vitro. Products of the *in vitro* glucuronidation were digested with xylanase GH11 and analysed by PACE (A). In addition to UX_4 an unexpected band (Band 1) was observed in the PtGUX2 reaction with UDP-GlcA. B) GH115 treatment was performed on products of the reaction containing UX_4 and Band 1. C) MALDI-ToF MS spectrum of 2-AA labelled oligosaccharides released by xylanase GH11 digestion of PtGUX2 *in vitro* glucuronidation products. For each oligosaccharide signal a secondary +10 Da signal, marked with *, was also observed.

To investigate the exact molecular structure of the Band 1 oligosaccharide a MS-MS fragmentation analysis was performed. As the UX_3 structure was not observed with PACE, it is likely that the higher molecular mass products, corresponding to m/z 866 (UX_4) peak and the m/z 1174 (U_2X_5) peak within each pair, are the canonical xylanase GH11 digestion products detected on the gel. Moreover, X5 is released upon Band 1 deglucuronidation, further confirming its identity as U_2X_5. Thus, the m/z 1174 peak is highly likely to be the MS signal corresponding to the Band 1 PACE structure and therefore it was chosen for the MS-MS fragmentation analysis.

A fragmentation spectrum was obtained (Figure 4.12) and analysed for the presence of Y and B fragmentation ions to assign the order of branched and un-branched xylose monomers. The Y_α fragmentation ion, corresponding to the loss of GlcA was detected, further supporting successful glucuronidation by the PtGUX2 enzyme. Interestingly, the mass difference between both the Y_2 and Y_3 ions and between the Y_3 and Y_4 ions was equal to 308 Da, corresponding to a xylosyl residue carrying a GlcA branch. Together with the mass differences between the B_3 and B_2 ion also being equal to 308 Da, this data further supports the presence of GlcA branches on consecutive xyloses of the xylopentaose backbone. Therefore, the molecular structure of Band 1 is likely to be XUUXX, as presented in figure 4.12.

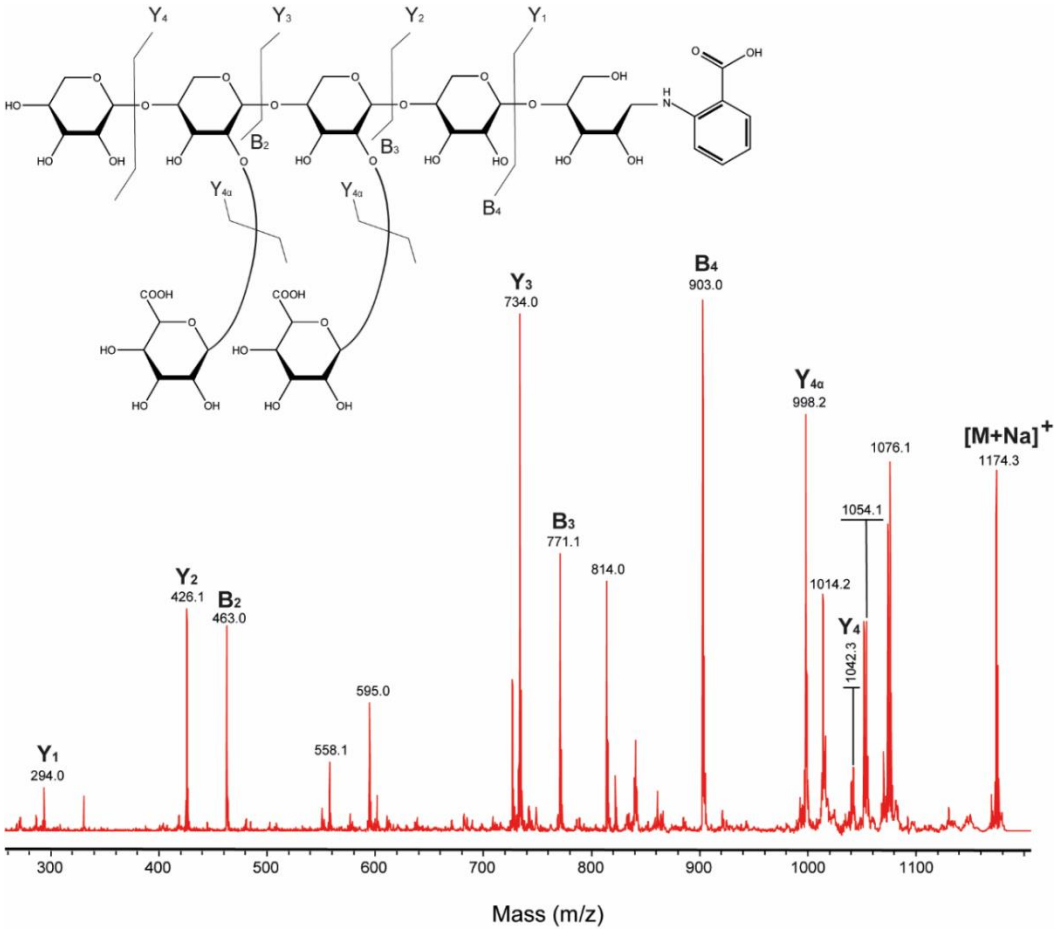
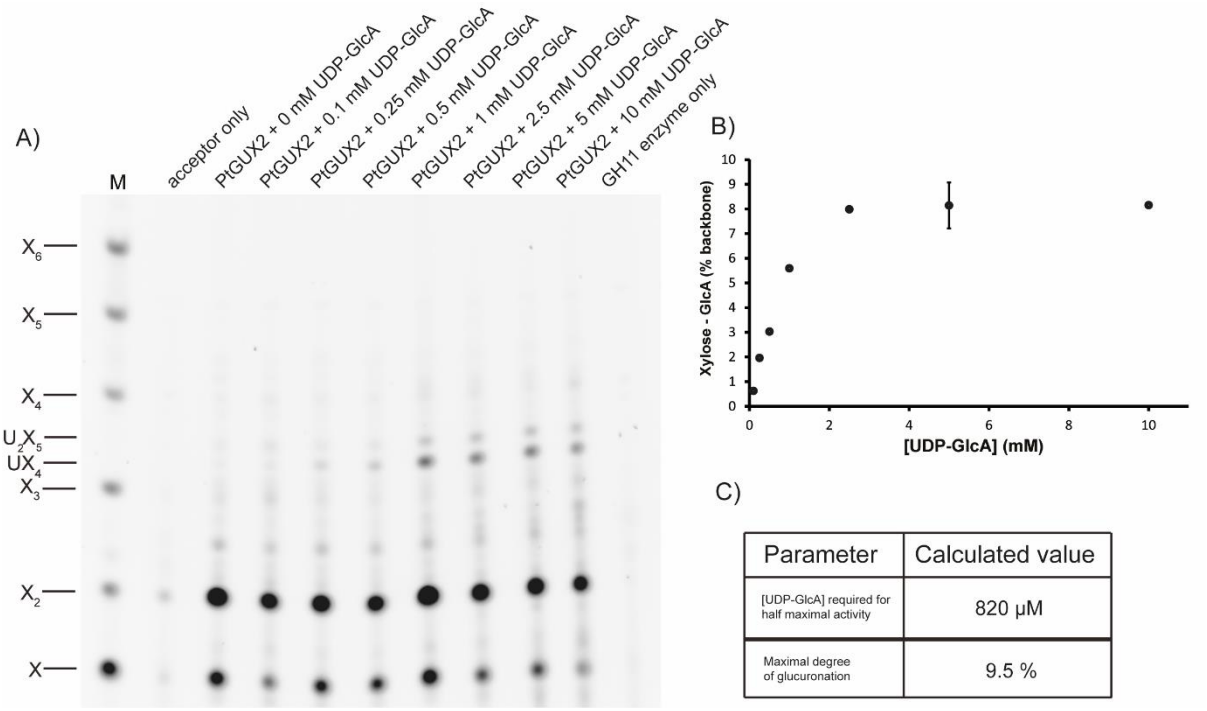


Figure 4.12 MALDI ToF MS-MS fragmentation analysis of the m/z 1174 structure. B and Y ions corresponding to the XUUXX structure are assigned in the spectrum. A 1174 parent ion is labelled as $[M+Na]^+$. A structure model with the fragmentation pattern marked is presented as an inset.

To compare maximum extent of xylan glucuronidation of the PtGUX2 enzyme with the one measured for the PgGUX1 in the 5h assay, a same set of *in vitro* activity assays was performed across a range of UDP-GlcA concentrations. Reaction products were extracted, deacetylated and digested with xylanase GH11. Resulting oligosaccharides were analysed with PACE (Figure 4.13A). Both UX4 and U2X5 structures were detected and their intensity increased with the amount of UDP-GlcA used in the assay. The total degree of glucuronidation was quantified by integrating the volume of bands corresponding to U2X5, UX4, X2 and X structures and plotted against the [UDP-GlcA] used (Figure 4.13B). The generated Michaelis-Menten like curve allowed for quantitation of the maximal possible degree of PtGUX2 dependant glucuronidation in the *in vitro* assay and the [UDP-GlcA] required to achieve half of it (Figure 4.13C).



Parameter	Calculated value
[UDP-GlcA] required for half maximal activity	820 µM
Maximal degree of glucuronation	9.5 %

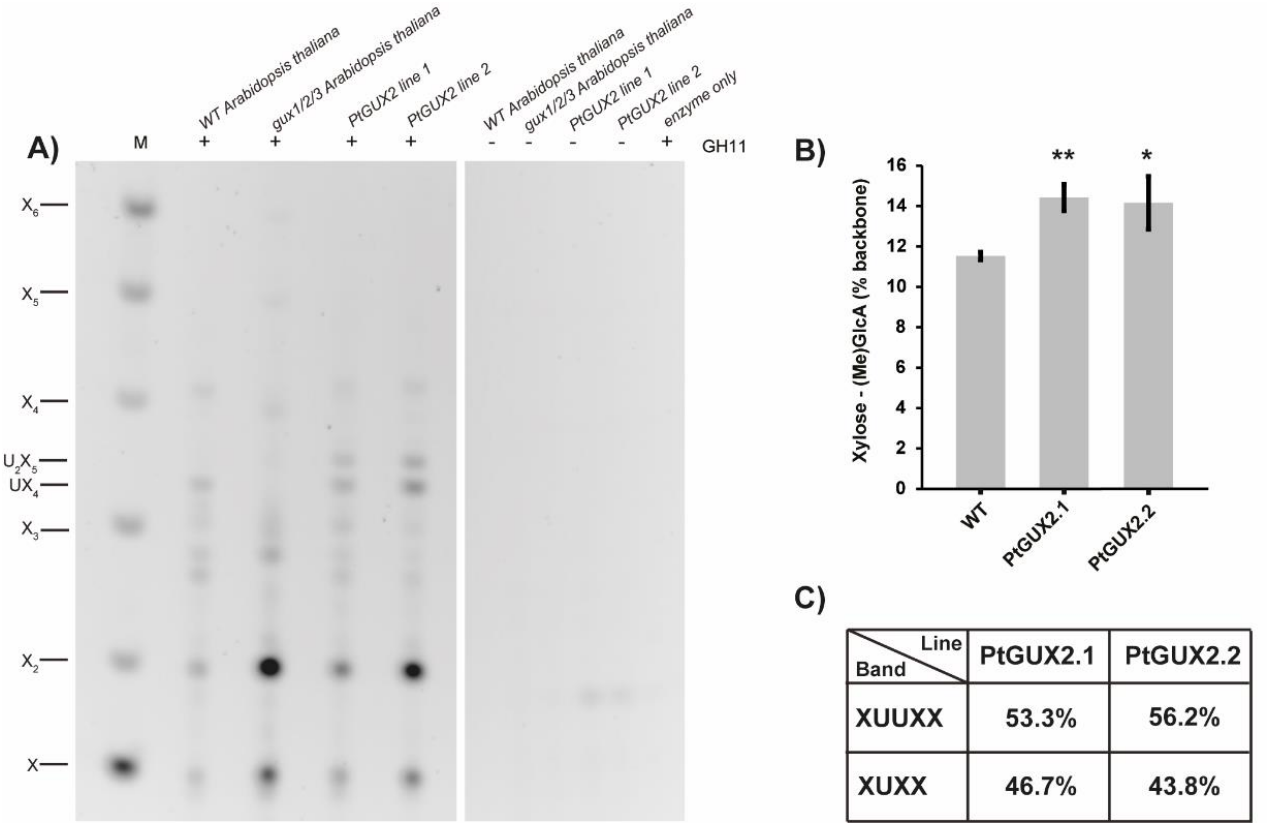
Figure 4.13 PtGUX2 activity across a range of UDP-GlcA concentrations. Products of *in vitro* glucuronosylation reaction with PtGUX2 and a range of [UDP-GlcA] were digested with xylanase GH11 and analysed by PACE (A). Degree of xylan glucuronidation for individual reactions, including both XUXX and XUUXX. (B) The amount of protein in the assay is unknown and the data was produced using a single batch of tobacco microsomes. Reaction using 5 mM UDP-GlcA was repeated three times, error bars represent standard deviation of the glucuronidation degree obtained. Activity parameters for PtGUX2 (C) quantified by non-linear regression analysis of the plotted data.

4.7 The GlcA branching pattern synthesised by the PtGUX2 enzyme *in vivo* is biologically functional.

After establishing that the PtGUX2 enzyme is an active glucuronosyltransferase *in vitro* a set of *in vivo* assays was performed to confirm the activity and in order to investigate if the enzyme is also capable of synthesising the consecutive pattern of GlcA branches *in planta.* To perform the *in vivo* assay, similarly to PgGUX1 experiments, *gux1/2/3 A. thaliana* plants were transformed with a construct in which *PtGUX2* expression is driven by the activity of the secondary cell wall cellulose synthase *IRX3* promoter. The *PtGUX2* CDS was followed by a 3' sequence encoding an in-frame 3xMyc tag. Two independent homozygous transgenic lines PtGUX2.1 and PtGUX2.2 were isolated.

The AIR obtained from stems of WT, *gux1/2/3,* PtGUX2.1 and PtGUX2.2 plants were treated with alkali and digested with xylanase GH11. The resulting oligosaccharides were analysed with PACE (Figure 4.14A). The biomass lacking both primary and secondary cell wall glucuronidation (*gux1/2/3* AIR) released only xylose and xylobiose upon xylanase GH11 treatment. The UX_4 oligosaccharide, indicative of xylan glucuronidation, was detected in digests performed using WT, PtGUX2.1 and PtGUX2.2 material. Interestingly, the PtGUX2 lines released an additional band with a migration like the U_2X_5 structure, synthesised in the *in vitro* assay. This annotation has to be confirmed by analysing products of *in vitro* reaction on the same PACE gel as the oligosaccharides released from the PtGUX2 lines. However, these results provide a strong indication that PtGUX2 is an active glucuronosyltransferase *in vivo* and that i*n planta*, similarly to the *in vitro* assay, it is likely to also synthesise the consecutive pattern of [Me]GlcA branches.

To investigate the efficiency of xylan glucuronation by the PtGUX2 enzyme *in vivo*, the PACE analysis was performed for three individual biological replicates of the WT, *gux1/2/3,* PtGUX2.1 and PtGUX2.2 biomass and the total degree of xylan glucuronidation was quantified by integrating the volume of X, X2, UX_4 and where applicable U_2X_5 structures (Figure 4.14B). The total degree of xylan glucuronidation obtained in both PtGUX2.1 and PtGUX2.2 lines was significantly increased over the WT levels and reached approx. 14%. To visualise the degree of consecutive GlcA branching detected in the PtGUX2 lines the fraction of [Me]GlcA contributed by the UX_4 and U_2X_5 structures was quantified (Figure 4.14C). This analysis indicated that the majority of [Me]GlcA is likely to be added in a consecutive manner by the PtGUX2 enzyme in both transgenic lines. Taken together, this data indicates that PtGUX2 is an active glucuronosyltransferase *in vivo* and that in this assay its likely main activity is to add GlcA to consecutive xylose units of xylan.



Band \ Line	PtGUX2.1	PtGUX2.2
XUUXX	53.3%	56.2%
XUXX	46.7%	43.8%

Figure 4.14 PtGUX2 is a functional xylan glucuronosyltransferase. PACE analysis of GH11 xylanase digests of WT, two independent transgenic lines of PtGUX2 in *gux1/2/3* and control *gux1/2/3* AIR (A). Undigested AIR controls are marked with (−). The [Me]GlcA-xylotetraose band (UX_4) was observed only in WT and PtGUX2 expressing lines. Doubly glucuronated U_2X_5 band was detected in both PtGUX2 lines. Quantitation of the degree of [Me]GlcA substitutions (B). Error bars represent standard deviation of three biological replicates of plant biomass. Contribution of both singly and doubly glucuronidated oligosaccharides to total xylan glucuronidation in both PtGUX2 lines is summarised in (C). *p value ≤0.05, **p value ≤0.01 in Student's t-test.

To investigate if the unusual pattern of GlcA branches synthesised by the PtGUX2 enzyme can confer biomass recalcitrance, saccharification assays were performed on three biological replicates of matching WT, *gux1/2/3*, PtGUX2.1 and PtGUX2.2 AIR. Both D-glucose (Figure 4.15A) and D-xylose (Figure 4.15B) release were measured after the treatment with the enzymatic saccharification cocktail. Similarly to PgGUX1 recalcitrance analysis, the *gux1/2/3* AIR released significantly more D-glucose and D-xylose than the WT material. For both PtGUX2 lines the amount of monosaccharides released was decreased and not significantly different from the sugar yield obtained from WT AIR. This indicates that the pattern of [Me]GlcA branches introduced by the PtGUX2 enzyme is capable of conferring biomass recalcitrance and thus it is likely to perform the same biological function as the WT xylan glucuronidation.

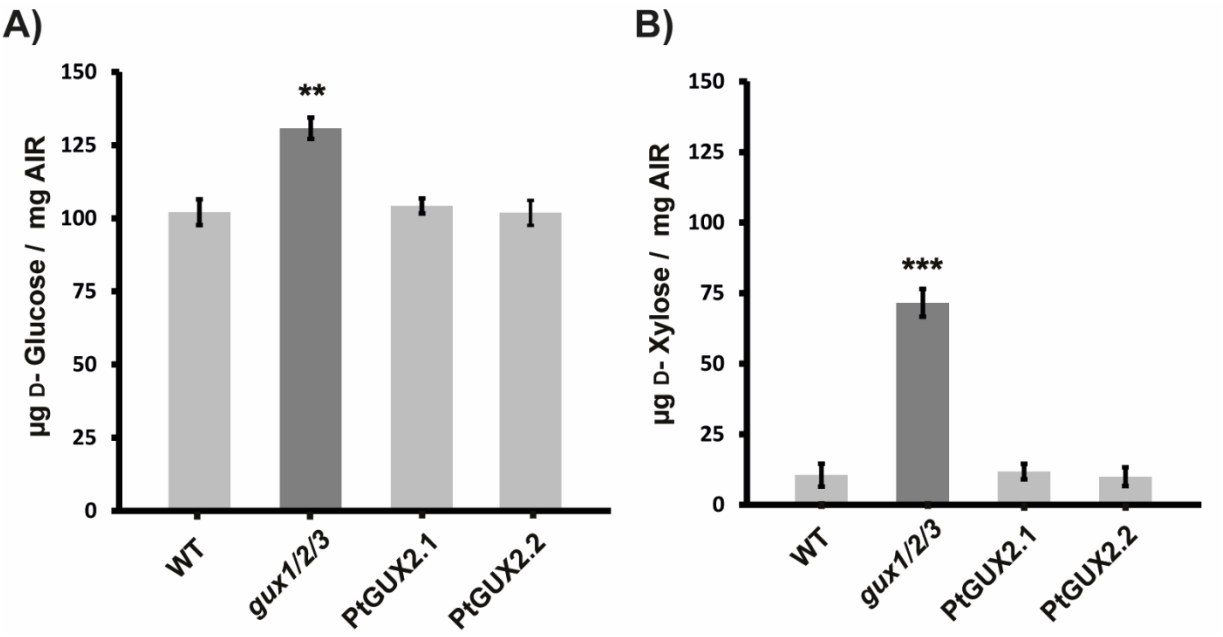


Figure 4.15 PtGUX2 saccharification analysis. D-glucose and D-xylose release from WT, *gux1/2/3* and two lines of PtGUX2 AIR. Error bars represent standard deviation of three biological replicates, **p value ≤0.01; ***p value ≤0.001 in Student's t-test.

4.8 Discussion

4.8.1 Representatives of both clades of conifer GUX enzymes are likely to contribute to softwood xylan biosynthesis

Globally, softwood is an important source of timber and biomass feedstock. Work described in chapter 3 of this thesis has demonstrated that in a model vascular plant *Arabidopsis thaliana,* removal of [Me]GlcA branches from xylan results in a significant reduction of biomass recalcitrance. To facilitate the application of this research into commercially relevant coniferous plants, this chapter has used both bioinformatics and biochemical analysis to identify and characterise conifer GUX enzymes likely to be responsible for glucuronidation of softwood xylan.

Use of data available via the CONGENIE service enabled identification of two distinct homologues of *A. thaliana* GUX enzymes encoded in the genome of *Picea abies.* Due to the large size and high complexity of conifer genomes the reference sequence available for *P. abies* is not complete (Nystedt et al., 2013). Indeed, sequence of the gene encoding PaGUX2 enzyme was incomplete and lacked the 5' part coding for the transmembrane region of the enzyme. Therefore, in order to include both GUX sequences from the model conifer species, *P. abies*, subsequent phylogenetic analysis was performed using the glycosyltransferase domain sequences of GUX enzymes. In addition to allowing for the inclusion of both *Picea abies* sequences, this approach excludes the highly variable N-terminal transmembrane region involved in Golgi targeting, and thus increases the confidence with which the phylogeny can be reconstructed.

After identification of the initial putative GUX sequences from *P. abies* and *P. glauca* the OneKP transcriptomic resource was applied to identify GUX sequences in a larger number of gymnosperm species. The OneKP database holds transcriptomic data for over 1000 plant species and has been successfully used to reconstruct evolution of hydroxyproline-rich glycoproteins (Johnson et al., 2017) or of the auxin response pathway (Mutte et al., 2018) across the tree of plant life, for example. The use of the OneKP database has enabled identification of GUX sequences in a range of coniferous and non-coniferous gymnosperm species. Phylogeny reconstructed using these sequences has suggested the presence of two distinct clades of conifer GUX enzymes. Interestingly, reads encoding the representatives of the second clade of

conifer GUX enzymes were only identified in a 3 out of 12 analysed gymnosperm transcriptomes. This might be caused by low expression levels of the clade two enzymes in tissues sampled for the transcriptomic analysis by OneKP. This lower number of clade two sequences is the likely cause for a lower, than for the clade one, bootstrap value when separating it from the angiosperm sequences. In order to resolve this issue it will be necessary to analyse a larger number of gymnosperm transcriptomes to identify additional members of both clades. Importantly, conifer GUX sequences share an exceptionally high degree of sequence similarity within each clade (Figure 4.1B). This is in line with previously published data indicating that the majority of gymnosperm evolution has happened prior to the lineage split with angiosperms and that after this event the pace of genome evolution has been slower in the gymnosperm lineage than in the angiosperm one (Pavy et al., 2012, Buschiazzo et al., 2012). Together with the reconstructed phylogeny the similarity analysis suggests the existence of two distinct clades of conifer GUX enzymes.

Similarly to Arabidopsis GUX1 and GUX2 enzymes, involved in secondary cell wall biosynthesis, the expression of both conifer *GUX* genes is enriched in stems. The expression profiling performed using the Norwood server has indicated that the conifer *GUX* co-expresses with a range of secondary cell wall biosynthesis enzymes, including putative secondary cell wall cellulose synthase components MA_10429177, MA_140410 and MA_183130. The co-expression with secondary cell wall cellulose synthesis was used to identify *A. thaliana GUX1* and *GUX2* in the first place (Mortimer et al., 2010). Moreover, the *in situ* hybridisation analysis indicated strong *PgGUX1* and *PaGUX2* mRNA signal in the vascular cambium. Taken together with the Norwood analysis, this is a strong indication that both PaGUX1 and PaGUX2 homologues are involved in glucuronidation of secondary cell wall xylan in softwood. It remains to be established if additional conifer GUX enzymes are involved in primary cell wall xylan glucuronidation, or if the lower *GUX* expression levels in the primary cell wall enriched tissues have led to lack of a strong *in situ* hybridisation signal.

4.8.2 *In vitro* and *in vivo* assays indicate that both conifer GUX homologues are active glucuronosyltransferases.

Analysis presented in this chapter used pEAQ-Hyper Trans (HT) *N. benthamiana* system to express glycosyl transferases from GT family 8 (GUX) and family 34 (AtMAGT1). The pEAQ-HT has been previously employed to express highly complex targets, including entire viral capsids (Thuenemann et al., 2013) or complete metabolic pathways for triterpene biosynthesis (Reed et al., 2017). Work described in this thesis extends the repertoire of proteins that can be expressed using the pEAQ-HT by developing methodology for production and harvesting of plant Golgi localised transmembrane glycosyltransferases. The expression of active protein obtained with the pEAQ-HT may enable structural studies or identification of residues involved in the specificity of glycosyltransferases. This will be particularly interesting for the PtGUX2 enzyme which is the only known GUX capable of catalysing glucuronidation of consecutive xylose monomers.

Work described in this chapter is the first to use polymeric acetylated *gux1/2* xylan as an acceptor for the GUX activity assay. These assay conditions allow high levels of *in vitro* xylan glucuronidation, exceeding 10% for PgGUX1 and 8% for PtGUX2. High efficiency of xylan glucuronation *in vitro* has enabled PACE and MS analysis, which facilitated characterisation of the consecutively glucuronidated xylan structure generated by PtGUX2. The efficiency of *in vitro* glucuronidation achieved in this assay exceeds previously reported results (Rennie et al., 2012). Emilie Rennie and colleagues used xylohexaose and radiolabelled UDP-GlcA to assay the activity of *A. thaliana* GUX1 - 5 enzymes. Increased incorporation of UDP-GlcA was detected for AtGUX1, AtGUX2 and AtGUX4 enzymes. No activity was detected for AtGUX3, which was later demonstrated to be an active glucuronosyltransferase *in vivo* (Mortimer et al., 2015). Low specific levels of GUX activity in the xylohexaose assay are a likely reason for this incorrect assignment. Specifically, AtGUX1 achieved only ~0.04% of total xylan glucuronidation, levels only 4 times higher than what was incorporated when control membranes, expressing PGSIP8, were used. Thus, the high levels of specific activity observed for PgGUX1 and PtGUX2 in the assay used performed as a part of this thesis, indicate that the use of polymeric xylan acceptor provides a superior environment to study the action of GUX enzymes *in vitro* when compared with the use of short oligosaccharides as described by Rennie et al. This efficient *in vitro* activity

assay will be used to study the spatial specificity of PgGUX1 and PtGUX2 in chapter 5 of this thesis.

High levels of xylan glucuronidation *in vitro* strongly indicate that PgGUX1 and PtGUX2 are active glucuronosyltransferases. This conclusion is further supported by the *in vivo* activity assays in which *gux1/2/3 A. thaliana* background was transformed with constructs in which *PgGUX1* and *PtGUX2* were expressed under the control of a strong secondary cell wall specific *IRX3* promoter. For each of the constructs a minimum of two individual homozygous lines were isolated, both showing restoration of [Me]GlcA branching. Saccharification assays performed on WT, *gux1/2/3* and transgenic feedstocks indicate that the [Me]GlcA introduced by the conifer enzymes is capable of restoring full levels of biomass recalcitrance. Interestingly, in the PgGUX1.2 line, in which the degree of xylan glucuronidation was more than 30% lower than in the WT plants, the recalcitrance was WT-like. This suggests that lower amount of [Me]GlcA branching is still capable of maintaining WT-like biomass recalcitrance. Moreover, the increase in [Me]GlcA levels observed in PtGUX2 plants did not result in a further increase in biomass recalcitrance. These observations are consistent with the proposed role of GlcA in the formation of ester linkages with lignin. In this model a small number of linkages might maintain full xylan-lignin cross-linking (Giummarella and Lawoko, 2016). Therefore, no further recalcitrance can be gained by introducing more [Me]GlcA and also the loss of a certain proportion of [Me]GlcA branches does not result in its reduction.

Importantly, both PgGUX1 and PtGUX2 dependant glucuronidations are capable of restoring the resistance of *A. thaliana* biomass to enzymatic depolymerisation. This observation may be important for conifer *GUX* mutagenesis experiments aiming at generation of softwood material without xylan glucuronidation and possibly with reduced recalcitrance. Results presented in this chapter indicate that even lower than WT levels of xylan glucuronidation are enough to provide WT-like biomass recalcitrance. Therefore, homologues of both conifer *GUX* may need to be targeted for mutagenesis to achieve full reduction in softwood recalcitrance.

4.8.3 Possible biological role of two individual GUX enzymes in softwood biosynthesis.

Some reports indicated that softwood has a simple xylan (Busse-Wicher et al., 2016b) structure. Consistent spacing of six xylose units was the only pattern of MeGlcA branches detected by Busse-Wicher et al. in pine and Douglas fir. This is in contrast with a more complex pattern of the acidic branches observed on hardwood xylan and synthesised by two distinct GUX enzymes in the *A. thaliana* model plant (Bromley et al., 2013). Interestingly, other reports indicate the presence of more complex xylan patterns in softwood (Shimizu et al., 1978, Yamasaki et al., 2011, Martinez-Abad et al., 2017). Martinez-Abad et al., have confirmed that the main spacing between MeGlcA branches is equal to six xylose units in spruce wood. However, similarly to Shimizu et al., and Yamasaki et al., the work has also reported the presence of MeGlcA branches on consecutive xylose monomers. In comparison to alkali extraction, also used by Busse-Wicher et al, these structures were greatly enriched in xylan fractions extracted with subcritical water. This may indicate that the two patterns of MeGlcA branches may be present on different xylan molecules.

Discoveries presented in this chapter indicate a possible link between the presence of two distinct clades of GUX enzymes in conifers and the two patterns of MeGlcA branching observed in softwood. It is possible that similarly to GUX1 and GUX2 enzymes in *A. thaliana* the conifer GUX1 and GUX2 enzymes might be responsible for the biosynthesis of individual xylan structures. One hypothesis is that the PgGUX1 clade of enzymes would be responsible for the synthesis of the main xylan pattern in softwood, with a spacing of six xylose units between the acidic branches. In this scenario, the PtGUX2 clade of enzymes would be responsible for addition of GlcA branches to consecutive xylose residues, forming the structure described by Martinez-Abad et al., among other papers, in spruce wood. To evaluate this hypothesis chapter 5 of this thesis will use a mixture of *in vitro* and *in vivo* assays to evaluate the activity of PgGUX1 and PtGUX2 on a more softwood-like xylan acceptor. In addition to that, it might be necessary to evaluate the specificity of further members of each clade to confirm functional assignments.

The biological role of the two types of softwood xylan patterning remains unknown. In *A. thaliana* the GUX1 enzyme is responsible for the biosynthesis of xylan molecules with an even pattern of GlcA branches. The GUX2 enzyme catalyses the addition of GlcA with a more dense, largely uneven, pattern. Thus, the hardwood xylan structure glucuronidated by GUX1 is predominantly compatible with interaction with the cellulose microfibril and the GUX2 glucuronidated structure is largely considered to be incompatible with interaction with the hydrophilic faces cellulose microfibril as a two-fold screw (see Busse –Wicher et al., 2016 and chapter 1 and Figure 1.13 for detailed review of this model). It is possible that the conifer GUX1 and GUX2 have a similar function as the angiosperm enzymes. If PtGUX2 clade is indeed responsible only for the synthesis of consecutively glucuronidated xylan the current xylan-cellulose interaction model indicates that the PtGUX2 dependant xylan structure would be unable to interact with the hydrophilic face of the cellulose fibril (Busse-Wicher et al., 2016a, Grantham et al., 2017). However, molecular dynamics simulations performed by Martinez-Abad et al., indicate that the consecutive pattern of xylan branching may also allow interaction of xylan with certain parts of the hydrophilic face of the cellulose microfibril. The impact this possible binding may have on coating of different faces of the cellulose fibril is discussed further in chapter 5. Interestingly, unpublished solid state NMR experiments on softwood indicate that both two-fold, cellulose bound, and three-fold, soluble, xylan structures are present in softwood (Terrett et al., in preparation). Thus, future NMR experiments using softwood from single *GUX* mutant conifer plants will be required to evaluate the exact role of each enzyme in biosynthesis of xylan that interacts with cellulose a two-fold and three-fold screw conformation.

Chapter 5: Determinants of gymnosperm and angiosperm GUX activity

5.1 Introduction

The specific pattern of GlcA and acetyl branches on xylan is important for the establishment of the xylan-cellulose interaction (Grantham et al., 2017). In the *A. thaliana* model, formation of this xylano-cellulose complex is facilitated by the presence of xylan regions with even spacing of substitutions over the majority of the polymer (Chapter 1 section 1.10.2 and Busse-Wicher et al., 2016a). In addition to this even spacing some xylan regions have incompatible, oddly spaced decorations. It is known that in Arabidopsis the even pattern of [Me]GlcA is synthesised by AtGUX1 and the incompatible pattern by AtGUX2 (Bromley et al., 2013). However, the reason for this difference in GUX activity and the specific regions of the xylan polymer which are decorated with the different branching patterns remain unknown.

Work described in Chapter 4 of this thesis has indicated that there are two distinct clades of GUX enzymes encoded in conifer genomes. The expression patterns of the genes encoding these glycosyltransferases support the hypothesis that they are both involved in wood formation. Similarly to hardwood, the softwood xylan has regions, or distinct molecules, with different MeGlcA patterns. The majority of conifer xylan has MeGlcA decorations in a compatible pattern but a small proportion has MeGlcA on consecutive xylosyl residues (Section 1.7.3, Chapter 1). It is unknown how these different conifer patterns of MeGlcA decorations are synthesised and how many GUX enzymes are involved in this process. Results presented in Chapter 4 indicate that conifer GUX enzymes have distinct activities *in vitro*, with conifer GUX clade 2 being capable of adding GlcA onto consecutive xylosyl units.

Importantly, unlike in hardwoods, in addition to [Me]GlcA, softwood xylan has Ara decorations rather than acetyl branches. In Chapter 4 of this thesis the *in vitro* and *in vivo* activity assays for conifer GUX enzymes PgGUX1 and PtGUX2 were performed on fully acetylated xylan molecules. Thus, the specific activity of the enzymes might have been altered in these conditions. It is known that changes in acetylation pattern alter the activity of hardwood GUX1 (Grantham et al., 2017) and that additional branches may sterically impede GUXs. Therefore, in order to fully understand the

synthesis of conifer xylan, it will be important to perform further characterisation of these enzymes on different xylan acceptors.

By evaluating any specificity of the conifer GUX enzymes towards generating compatible or incompatible xylan molecules one might identify a possible basis for control of the xylan-cellulose interaction in softwood. This may be relevant for the maintenance of softwood properties which has significant industrial importance.

Taking all this into consideration, this chapter will aim to investigate the specific activity of conifer GUX enzymes. In particular, work presented in this chapter will attempt to:

- Evaluate the impact of xylan acetyl branches on the activity of conifer GUX enzymes;

- Analyse the pattern of GlcA branches added by the conifer GUX enzymes when acting on fully acetylated and partially deacetylated xylan acceptors;

- Compare the impact of xylan acetylation on softwood and hardwood (eudicot) GUX activities.

5.2 The pattern of [Me]GlcA in PgGUX1 lines is different to the one observed in softwood and in WT *A. thaliana*

In Chapter 4, it was shown that PgGUX1 transfers GlcA to xylan in Arabidopsis plants with a maximal frequency of 12%, but whether this is in a regular pattern or randomly arranged is unknown. The enzyme might produce xylan with GlcA spaced strictly every 6 xylosyl residues, as in *Picea spp.*, or it might be influenced by eudicot xylan synthesis machinery or other modifying enzymes (e.g. ESK1) to produce alternative patterns. In addition to [Me]GlcA branches, Gnetophyte xylan carries acetyl decorations (Busse-Wicher et al., 2016b). This is in contrast with conifer xylan, which is not acetylated and has Ara branches. Therefore, the Gnetophyte xylan structure is more similar to that of hardwoods than of softwoods (Busse-Wicher et al., 2016b). On the other hand, the similarity between the sequence of Gnetophyte GUX1 and PgGUX1 exceeds 80% (see figure 4.1 in Chapter 4). Therefore, the biosynthetic environment of the Gnetum gymnosperm GUX1 enzyme in *G. montanum* may be similar to PgGUX1 expressed in *A. thaliana*.

To compare the pattern of [Me]GlcA branches in softwoods and in *A. thaliana* lines expressing the PgGUX1 enzyme, AIR was isolated from WT *A. thaliana*, two independent PgGUX1-expressing *gux1/2/3 A. thaliana* lines, *Picea glauca* de-barked

branches and *Gnetum montanum* stems. AIR isolated from the samples was de-acetylated and digested with a GH30 glucuronoxylanase. Xylanase GH30 requires [Me]GlcA branches to catalyse the hydrolysis of the xylan backbone. The enzyme specifically targets the β-1,4 linkage between the +1 and +2 xylose towards the reducing end from the [Me]GlcA carrying residue (Chapter 1, section 1.9.4). Thus, the DP of oligosaccharides released by GH30 is equal to the distance between the glucuronidated residues and the digestion can be used to analyse the pattern of [Me]GlcA decorations on xylan (Bromley et al., 2013).

Isolated oligosaccharides were derivatised with ANTS and analysed on PACE (Figure 5.1). Similarly to what was previously observed by Busse-Wicher et al., 2016 the *P. glauca* AIR released predominately UX_6 and UAX_6 structures and a small proportion of other oligosaccharides, including some UX_7. Also consistent with published data (Bromley et al., 2013), WT *A. thaliana* AIR produced a range of oligosaccharides with UX_8 dominating and the glucuronidated oligosaccharides with an even backbone DP more abundant than those with an odd backbone DP. As previously reported (Busse-Wicher et al., 2016b) the Gnetophyte material produced predominantly UX_6 and there was little preference for even DP in the other released oligosaccharides. Interestingly, AIR from both lines of PgGUX1 plants released a different oligosaccharide pattern than the WT *A. thaliana* material. UX_6 was the most frequent oligosaccharide in the PgGUX1 digests and the preference for an even DP of longer oligosaccharides was largely lost. The GH30 oligosaccharide profile of PgGUX1 biomass showed some similarities to that from *G. montanum.* Specifically, in both digests UX_6 was the most frequent oligosaccharide and the intensity of the UX_7 oligosaccharide was lower than that of UX_8. In addition to that, the longer DP oligosaccharides (>UX_8) largely lacked preference for the even length of the backbone.

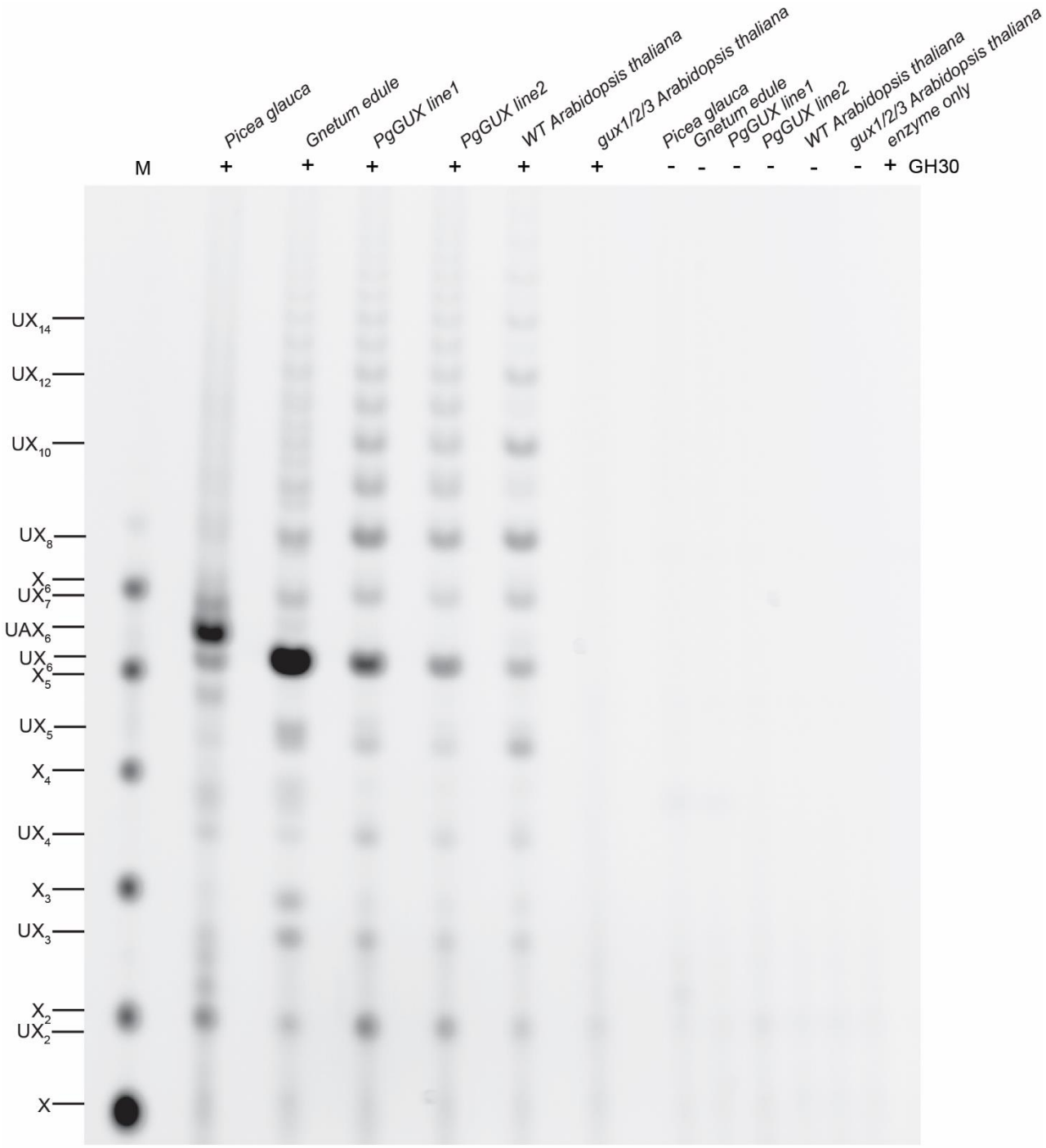


Figure 5.1 [Me]GlcA distribution on xylan in PgGUX lines, WT and *gux1/2/3* Arabidopsis, Spruce (*Picea glauca*) and a gnetophyte (*Gnetum montanum*). AIR of all plants was hydrolyzed with glucuronoxylanase GH30 and analysed by PACE. Undigested material was used as a control.

The PACE results indicate that the pattern of [Me]GlcA branches differs between WT and PgGUX1 *gux1/2/3 A. thaliana*, indicating that the PgGUX1 enzyme specificity is different to that of AtGUX1 and AtGUX2. Also, the pattern of [Me]GlcA decorations is different in PgGUX1 lines than it is in *Picea glauca.* This suggests that either PgGUX1 is not the glucuronosyltransferase involved in synthesis of the digested softwood xylan or that the PgGUX1 expression alone is not able to generate a conifer-like pattern of [Me]GlcA branches. If the second interpretation is correct, this result may suggest that factors other than the protein sequence (Rennie et al., 2012) influence GUX enzyme specificity. One hypothesis is that the backbone acceptor structure can affect the [Me]GlcA patterning, such as the presence of acetate or absence of Ara branches in Arabidopsis. This is supported by the similarity between the patterning in PgGUX1 lines and in *G. montantum.* In both cases, highly similar gymnosperm GUX enzymes act on acetylated xylan acceptors. In addition, the presence of acetate branches was already demonstrated to influence hardwood GUX activity (Grantham et al., 2017). Alternatively, specific interactions with other GUX enzymes or the xylan biosynthesis complex may be required to set the precise [Me]GlcA patterning.

5.3 The *in vitro* activity of conifer GUX enzymes is stimulated when the xylan acceptor is partially deacetylated

The results of the GH30 analysis presented in section 5.2 of this chapter indicate that factors other than the protein sequence or structure may contribute to the GUX enzyme specificity. The comparison between the [Me]GlcA pattern in PgGUX1 lines and in *G. montanum* xylan indicated that the acetylation of the xylan backbone may influence the activity of gymnosperm GUX enzymes. To evaluate the influence of xylan acceptor acetylation on the maximal UDP-GlcA transfer efficiency of PgGUX1 and PtGUX2 enzymes the *in vitro* GlcA transfer (GlcAT) was performed using xylan acceptors with reduced acetylation. Unfortunately, deacetylated xylan without [Me]GlcA substitutions is insoluble, and would therefore not be a suitable acceptor. Thus, the GlcAT onto the *gux1/2* acetylated xylan acceptor was performed in the presence or absence of acetylxylan esterase. The esterase CE4 removes single 2 or 3 linked acetyl groups from xylose monomers which do not carry GlcA decorations (Taylor et al., 2006). In order to maintain xylan solubility the GlcAT assay was performed for 3h prior to the addition of the esterase enzyme. To allow for further GlcA

transfer onto xylan by the GUX enzymes the concentration of the UDP-GlcA was increased to 10mM upon addition of the acetylxylan esterase enzyme.

Following over-night incubation of *gux1/2* xylan with UDP-GlcA and PgGUX1 or PtGUX2 in the presence or absence of acetylxylan esterase the reaction products were extracted, fully de-acetylated with an alkali treatment, digested with xylanase GH11 and the released oligosaccharides were analysed by PACE (Figure 5.2A). All reactions catalysed by both PgGUX1 and PtGUX2 released XUXX in addition to xylose and xylobiose. This confirms successful GlcA transfer in all of the conditions. In addition to XUXX the PgGUX1 catalysed reaction released a novel band (Band 1) migrating between xylopentaose and xylotetraose standards. Upon re-analysis low intensity Band 1 was observed to be also detectable in the GH11 digestion products of PgGUX1 *in vitro* GlcAT reactions performed in Chapter 4. Interestingly, the intensity of this band increased with higher UDP-GlcA concentration and longer reaction time and was further increased in the reaction that included the esterase enzyme. For the PtGUX2 catalysed glucuronidation both XUXX and XUUXX products, in addition to xylose and xylobiose, were present in the reaction lacking acetylxylan esterase. In the reaction which included the esterase enzyme, the XUUXX became the dominant glucuronidated product. Moreover, the relative amount of xylobiose decreased in the product profile of GlcAT reactions performed with the esterase. This indicates increased xylan glucuronidation levels.

To investigate the identity of the novel band generated in the PgGUX1 reactions the oligosaccharides released by the xylanase GH11 from the enzymatic reaction in the presence of acetylxylan esterase were digested with alpha-glucuronidase GH115. The enzyme is able to remove α-1-2 linked GlcA from xylose monomers present at any position within the xylooligosaccharide (Tenkanen and Siika-aho, 2000). Following GH115 deglucuronidation the oligosaccharides were analysed by PACE (Figure 5.2B). In addition to xylose and xylobiose, two new oligosaccharides xylotetraose and xylohexaose were detected following GH115 treatment. As the xylotetraose is likely to have originated from XUXX deglucuronidation, the novel oligosaccharide might have six xyloses in the backbone.

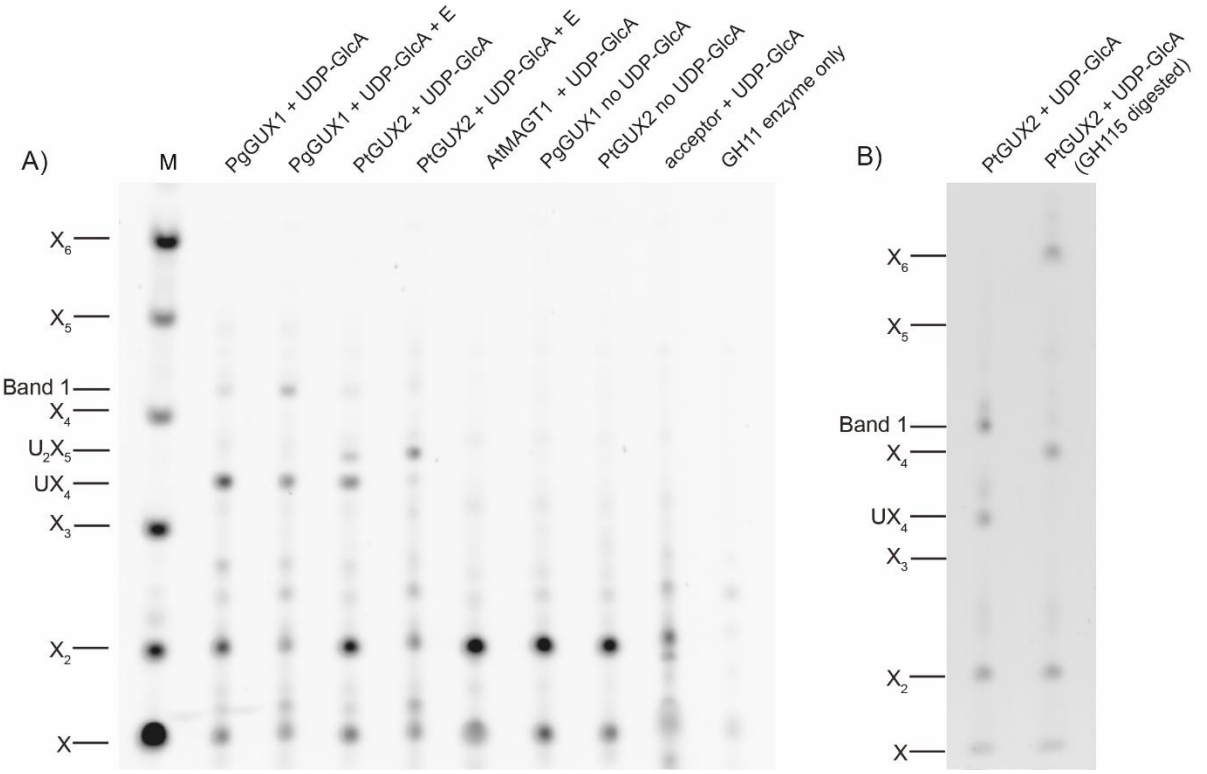


Figure 5.2 *In vitro* GlcA transfer by PgGUX1 and PtGUX2 onto acetylated and partially deacetylated xylan. A) PACE analysis of xylanase GH11 digestion of *in vitro* glucuronidation products catalysed by the two conifer enzymes. The reaction was carried out in the presence or absence of xylan acetyl esterase (E). Reactions with AtMAGT1 GT34, no UDP-GlcA and no xylan acceptor were analysed as controls. B) PACE of GH115 glucuronidase digested oligosaccharides released by GH11 from xylan glucuronidated by PgGUX1 in the presence of the esterase.

In order to investigate the molecular structure of the band 1 oligosaccharide the PgGUX1 dependant *in vitro* glucuronidation in the presence of acetylxylan esterase was repeated. Reaction products were fully deacetylated and digested with xylanase GH11. Released oligosaccharides were derivatised with 2-AA and analysed with MALDI-ToF MS (Figure 5.3A). A signal at 866 m/z corresponding to a sodiated 2-AA labelled UX_4 (XUXX) oligosaccharide was detected, further confirming successful glucuronidation in these reaction conditions. A peak at 745 corresponding to sodiated UX_4 without label suggests the 2AA derivatisation was incomplete. In addition to this canonical GH11 digestion product a weaker signal at m/z 734 was also detected. This mass is consistent with a GH11 mis-digestion product: UX_3. In addition to these expected signals, a new set of peaks at m/z equal to 1174, 1185 and 1306 was detected. As the novel oligosaccharide product (Band 1 on Figure 5.2A) of PgGUX1 *in vitro* glucuronidation is likely to have 6 xylose units in the backbone the m/z of these

signals was matched to X6 backbone with a varying number of GlcA branches. The 1306 m/z signal is consistent with the mass of sodiated 2-AA labelled xylohexaose with two GlcA branches (U_2X_6) and the 1174 corresponds to a structure of $U_2 X_5$, which is likely to be a matching 2AA labelled GH11 mis-digestion product. For both 2-AA labelled UX_4 and U_2X_6, matching sodiated peaks without the 2-AA derivative were also detected at m/z 745 and 1185 respectively. These results indicate that the Band 1 structure detected on PACE is highly likely to have a m/z equal to 1306 following 2-AA labelling.

In order to investigate the structure of the novel oligosaccharide an MS-MS fragmentation analysis of the m/z 1306 structure was performed. It is important to remember that a difference of 132 Da indicates the loss of an unsubstituted xylosyl residue, whereas a mass difference of 132 + 176 = 308 Da indicates the loss of a glucuronidated xylosyl residue. A clear fragmentation pattern was recorded (Figure 5.3B) and a set of Y and B fragmentation ions was matched to the spectrum. Interestingly, the mass difference between both Y_2 and Y_3 and between Y_4 and Y_5 fragmentation ions was consistent with loss of a xylosyl monomer carrying a GlcA branch. Further analysis confirmed that, all Y and B ions detected on the spectrum were consistent with a structure of xylohexaose with GlcA branches on the 3rd and 5th xylosyl unit from the reducing end of the oligosaccharide (Figure 5.3B top). This structure has GlcA branches on two xylose units spaced only by one unsubstituted residue. Hence, due to steric hindrance, GH11 cannot digest it to smaller oligosaccharides.

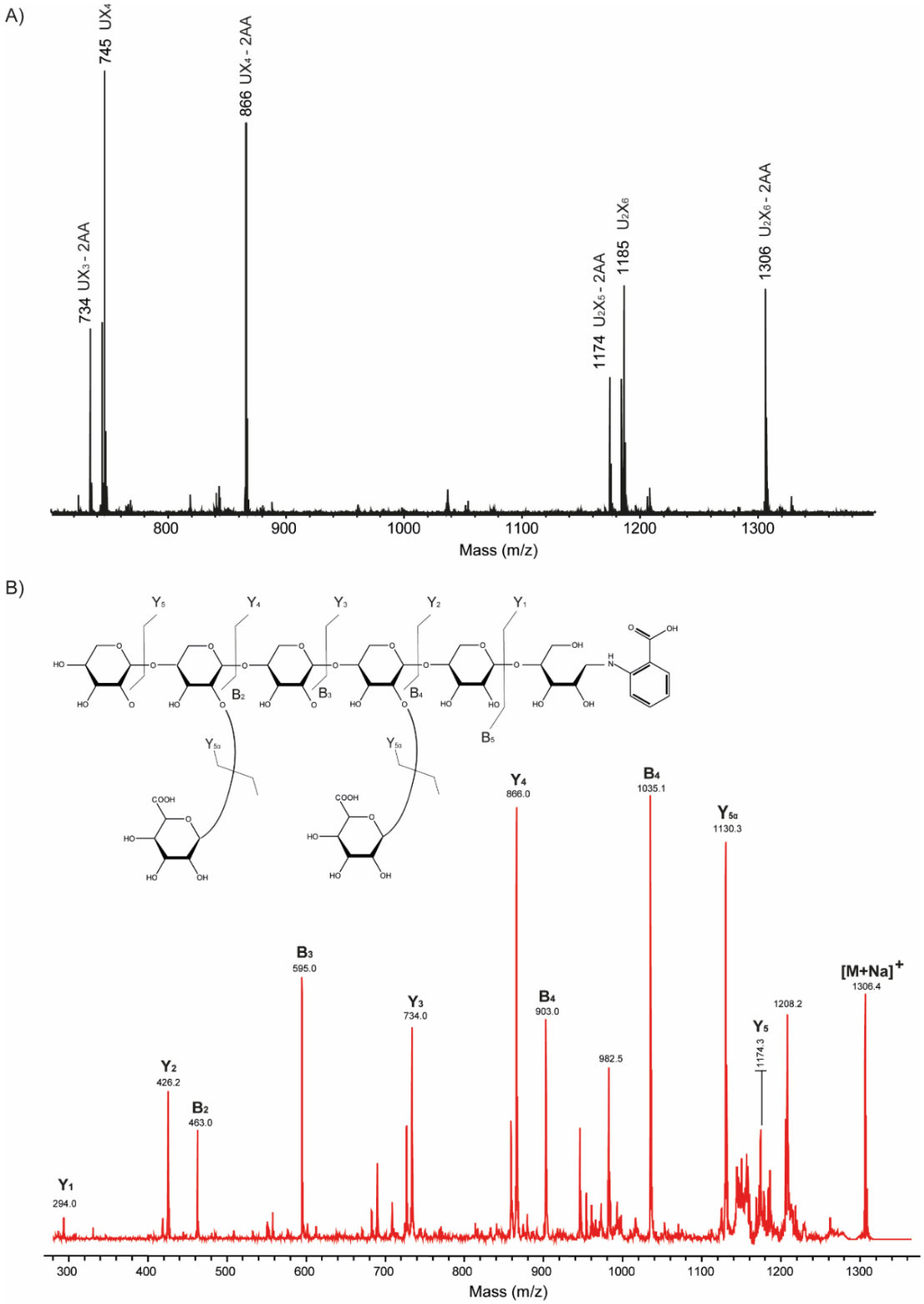
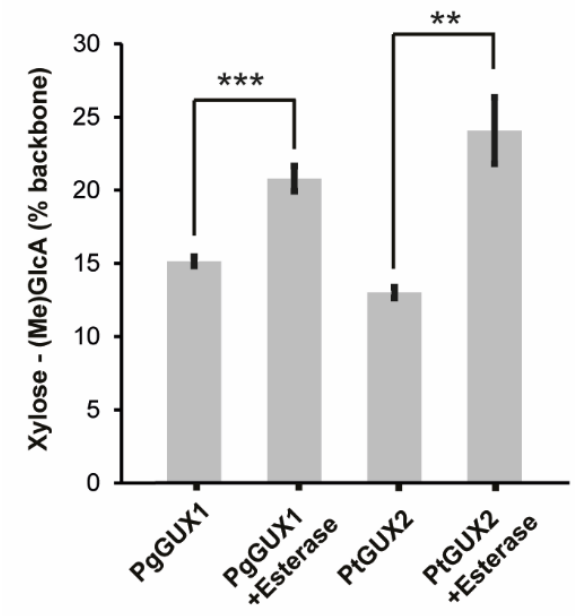


Figure 5.3 MS analysis of PgGUX1 *in vitro* glucuronidation products. A) Mass spectrum of 2-AA labelled oligosaccharides released by GH11 from PgGUX1 catalysed glucuronidation performed in the presence of the acetyl xylan esterase. B) Fragmentation spectrum of the m/z 1306 peak, marked as $[M+Na]^+$ and deduced oligosaccharide structure. All B and Y fragmentation ions are marked on the structure and the spectrum.

Having established that the molecular structure of the novel band generated by the PgGUX1 enzyme is XUXUXX, the degree of glucuronidation could be quantified across PgGUX1 and PtGUX2 *in vitro* glucuronidation reactions in the presence and the absence of acetylxylan esterase. To achieve this, the *in vitro* activity assays were repeated three times using one batch of tobacco microsomal proteins expressing the enzymes and the generated glucuronoxylan was digested with xylanase GH11. Released oligosaccharides were analysed on PACE. The intensity of the XUXUXX, XUXX, XX and X bands was quantified for PgGUX1 and for PtGUX2 the XUUXX, XUXX, XX and X bands were quantifed. The resulting data allowed quantitation of the average maximal degree of xylan glucuronidation obtained in the overnight PgGUX1 and PtGUX2 reactions in the presence and absence of the acetylxylan esterase (Figure 5.4A). As the assay used is likely to be an end-point measure of the glucuronidation process these values should represent the maximal degree of glucuronidation in the reaction conditions used.

For both enzymes the degree of xylan glucuronidation was increased when the reaction was performed together with the xylan acetylesterase enzyme. For PgGUX1 the activity increased by 25% and for PtGUX2 it has nearly doubled. In order to evaluate the contribution of the doubly glucuronidated oiligosaccharides to the total degree of glucuronidation across the reactions, the contribution of XUXUXX in each PgGUX1 reaction and the contribution of XUUXX in each PtGUX2 reaction was expressed as a percentage of the total GlcA content (Figure 5.4B). The GlcA originating from the doubly glucuronidated oligosaccharides amounted to half of the total GlcA in PgGUX1 reactions in the presence of acetylxylan esterase. For the PtGUX2 reaction in the presence of esterase, over 95% of GlcA was contributed by the doubly glucuronidated XUUXX structure. The stimulation of XUXUXX formation by PgGUX1 and XUUXX formation by PtGUX2 in the presence of the acetylxylan esterase enzyme is likely to result from the removal of acetylation, which due to its predominant positioning on alternating xylosyl residues in *gux1/2* xylan (Busse-Wicher et al., 2014), may result in inhibition of the glucuronidation process required for formation of these oligosaccharides.

A)



B)

Reaction	% GlcA contributed by the doubly glucuronated oligosaccharide
PgGUX1 + UDP-GlcA	16 (± 2) %
PgGUX1 + UDP-GlcA + esterase	52 (± 3) %
PtGUX2 + UDP-GlcA	38 (± 4) %
PtGUX2 + UDP-GlcA + esterase	95 (± 2) %

Figure 5.4. Degree of xylan glucuronidation by PgGUX1 and PtGUX2 *in vitro*. A) Total degree of glucuronidation by PgGUX1 and PtGUX2 in the presence or absence of acetyl esterase enzyme in the *in vitro* reaction. B) Contribution of the doubly glucuronidated GH11 oligosaccharide products (XUXUXX for PgGUX1 and XUUXX for PtGUX2) to the total degree of glucuronidation in the different reaction conditions. Error bars represent standard deviation of three replicates of the *in vitro* reaction, **p value ≤0.01; ***p value ≤0.001 in Student's t-test.

The results of the GH11 analysis suggest that the removal of acetyl branches from the xylan acceptor stimulates the activity of both conifer GUXs *in vitro*. A significant proportion of the GlcA added by both enzymes onto partially de-acetylated xylan is densely packed, with either one or no unbranched xyloses between the branched monomers. This suggests that the PgGUX1 and PtGUX2 activity is not only stimulated but also the pattern of GlcA branches added by the enzymes *in vitro* may change upon partial removal of acetyl branches from the acceptor. As this change results from the enzymatic removal of acetyl branches during the glucuronidation process, it is possible that on the native *gux1/2* acceptor the O-2 linked acetate branches in proximity of the newly glucuronidated xylosyl monomer may prevent addition of further decorations to neighbouring or closely positioned backbone units.

5.4 The pattern of GlcA branches generated by PgGUX1 and PtGUX2 *in vitro* is different when the xylan acceptor is partially deacetylated.

To evaluate the pattern of GlcA branches generated by PgGUX1 and PtGUX2 *in vitro* the xylan products of reactions in the presence or absence of the xylan acetylesterase were deacetylated and digested with xylanase GH30. Released oligosaccharides were derivatised with the ANTS fluorophore and analysed by PACE (Figure 5.5). In addition to the xylo-oligosaccharide standard, the WT *A. thaliana* AIR was digested with GH30 to release known glucuronidated xylooligosaccharides.

The pattern of GlcA branches generated by PgGUX1 in the absence of xylan acetylesterase was largely similar to the pattern detected for PgGUX1 transgenic *A. thaliana* lines (Figure 5.1). The dominant spacing of GlcA was equal to 6 xylose units but longer DP oligosaccharides were also released, with a slight preference for even spacing not observed in PgGUX1 transgenic *A. thaliana* lines. The pattern changed substantially upon the addition of xylan acetylesterase to the reaction. Partial deacetylation of the acceptor resulted in PgGUX1 generating a denser GlcA patterning *in vitro* with spacings of 6, 4 and 2 preferred.

GH30 digestion of the glucuronoxylan generated by PtGUX2 in the absence of the xylan acetylesterase released a mixture of oligosaccharides. Interestingly, each known singly glucuronidated oligosaccharide, co-migrating with the standard released by GH30 from WT Arabidopsis AIR, had a second band migrating close to it. The GH11 digestion of the products of the same reaction resulted in production of singly and doubly glucuronidated oligosaccharides. Therefore, presumably, the GH30 oligosaccharides migrating close to known digestion products may be carrying GlcA on consecutive xyloses, i.e. the second and third xylose from the reducing end. This structure is unlikely to be digested further due to the steric hindrance of GH30. Addition of the xylan acetylesterase to PtGUX2 catalysed glucuronidation has led to a change in the resulting GH30 pattern. In these conditions shorter oligosaccharides were released and almost all of them migrated like putative doubly glucuronidated oligosaccharides. This alteration in the ratio of doubly and singly glucuronidated oligosaccharides, is in line with the GH11 digestion results which indicated that the dominant glucuronidated oligosaccharide generated by PtGUX2 when the xylan acceptor is partially deacetylated is XUUXX (Figure 5.4).

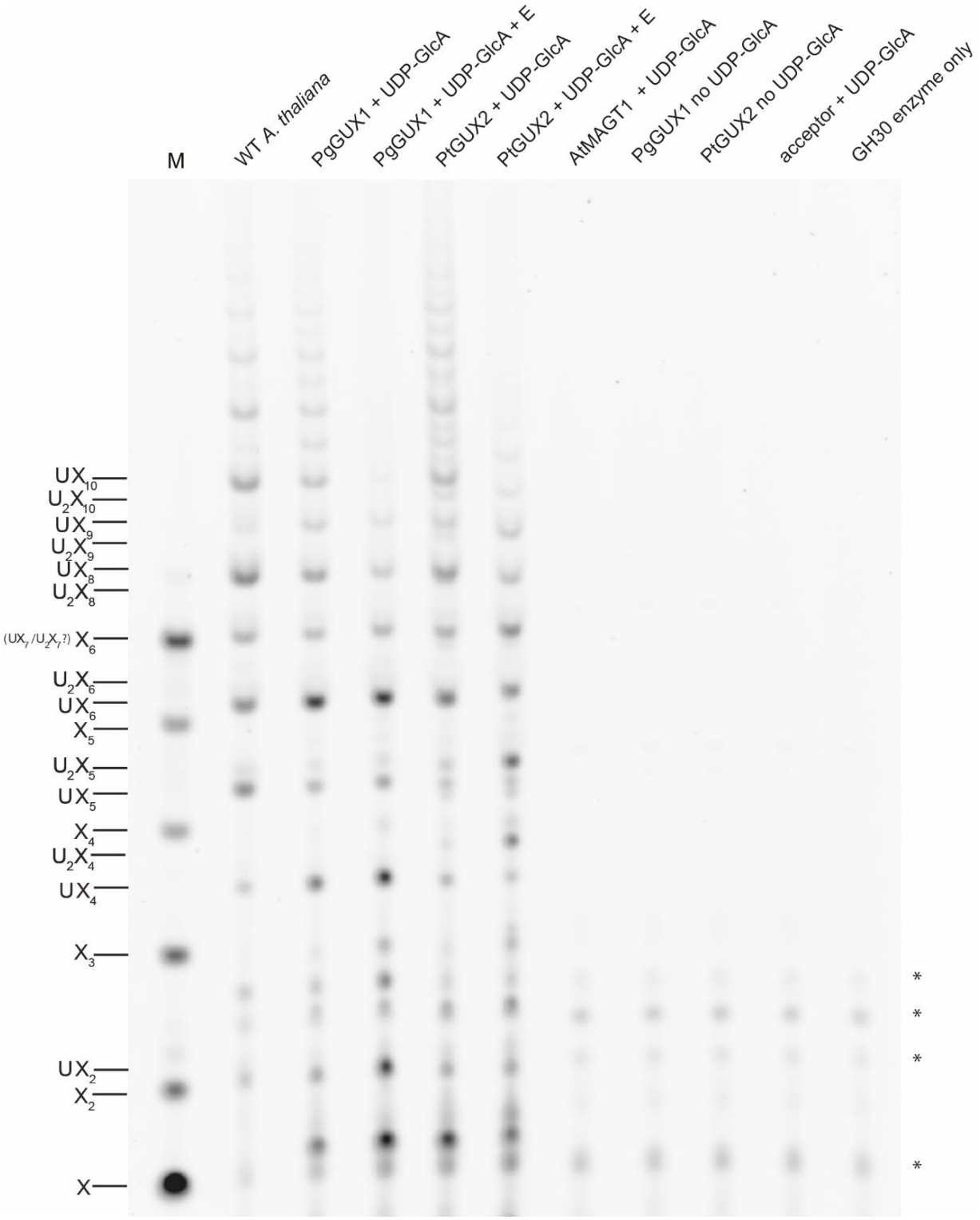


Figure 5.5 GlcA pattern generated *in vitro* by PgGUX1 and PtGUX2 in the presence or absence of xylan acetylesterase. Xylan product was hydrolyzed with glucuronoxylanase GH30 and analysed by PACE. Xylan resulting from reactions performed in the presence of glucomannan galactosyltransferase AtMAGT1 or in the absence of UDP-GlcA was digested as a control. UX_7 and U_2X_7 oligosaccharides may be co-migrating with X6 and were marked with a question mark. Background bands prevent part of the annotation and were marked with *.

Together with the GH11 data the results of the GH30 analysis provide further evidence that the pattern of GlcA branches added by both PgGUX1 and PtGUX2 *in vitro* may be sensitive to the presence of acetyl branches on the xylan acceptor. Interestingly, for both enzymes, partial removal of acetyl branches resulted in generation of more densely glucuronidated xylan molecules. This is indicated by both a significant increase in the degree of glucuronidation in the GH11 analysis, and the release of shorter oligosaccharides in the GH30 digestion. In addition to that, the addition of the acetylesterase enzyme enabled PtGUX2 to add GlcA to neighbouring xylose units with very high efficiency, suggesting this might be a preferred mode of glucuronidation for this enzyme. To further evaluate the determinants of PgGUX1 and PtGUX2 activity, *in vivo* experiments may be required to support the proposed relationship between xylan acetylation and the activity of conifer GUX enzymes.

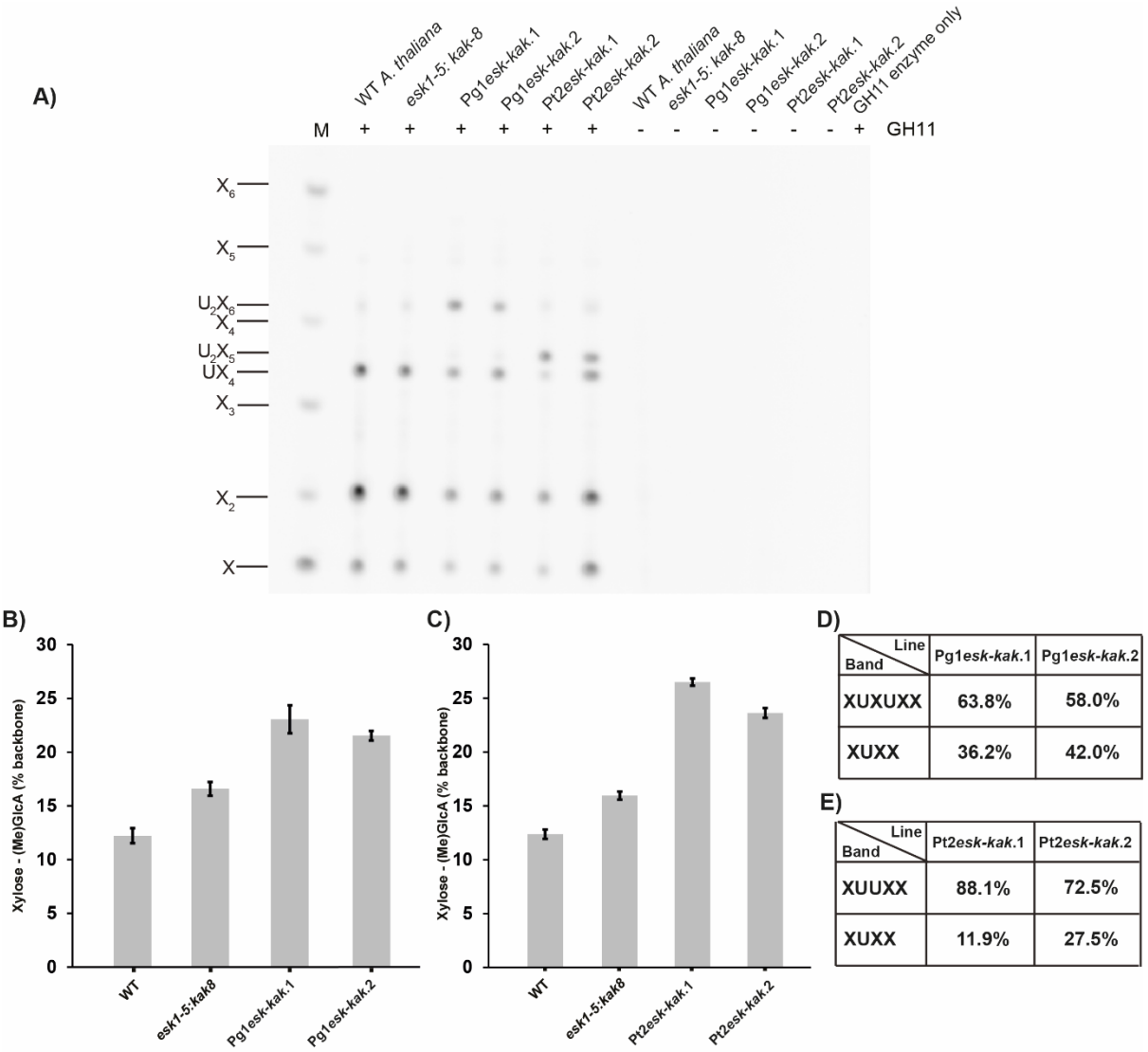
5.5 The activity of conifer GUX enzymes is changed when they act on partially un-acetylated xylan *in vivo.*

The analysis of GUX activity using the Arabidopsis model can provide additional information not observable in *in vitro* assays. In particular, in addition to indicating how changes in the acceptor structure influence GUX activity the *in planta* analysis may indicate the impact of other xylan biosynthesis enzymes on the process of glucuronidation. Therefore, to test the importance of xylan acetylation for PgGUX1 and PtGUX2 activity *in planta*, the constructs in which genes encoding both enzymes are expressed under the control of a strong secondary cell wall *IRX3* promoter were transformed into *esk1-5:kak-8 A. thaliana* plants. The *esk1-5* mutant background has a ~60% reduction in xylan acetylation (Xiong et al., 2013) which leads to the loss of xylan compatible patterning (Grantham et al., 2017) and dwarfing. The *kak8* suppressor mutation partially restores plant growth, but has no effect on the level of xylan acetylation or xylan patterning (Bensussan et al., 2015, Grantham et al., 2017). Therefore, the *esk1-5:kak8* plant is an easily, thanks to improved growth, transformable *A. thaliana* with reduced xylan acetylation. Two independent homozygous transgenic *esk1-5:kak-8* lines were isolated for both PgGUX1 and PtGUX2 transformation events by screening 10 independent hemizygous insertional lines for the degree of xylan glucuronidation. The two lines with the most GlcA were used in the analysis.

For the **PgGUX1** expression the two independent p*IRX3*::*PgGUX1 esk1-5:kak-8* lines were named **Pg1*esk-kak*.1** and **Pg1*esk-kak*.2**. For the **PtGUX2** expression the two independent p*IRX3*::*PtGUX2 esk1-5:kak-8* lines were named **Pt2*esk-kak*.1** and **Pt2*esk-kak*.2.**

To evaluate the impact of reducing the acetylation of xylan on the *in vivo* activity of PgGUX1 and PtGUX2, AIR isolated from matching WT, *esk1-5:kak-8* and two independent transgenic *esk1-5:kak-8* lines overexpressing either PgGUX1 or PtGUX2 was deacetylated and digested with xylanase GH11. The released oligosaccharides were analysed with PACE (Figure 5.6A). For all the digestion reactions xylose, xylobiose and UX_4 were released. In addition to UX_4 the U_2X_6 doubly glucuronidated oligosaccharide was abundant in the products released from AIR isolated from both Pg1*esk-kak*.1 and Pg1*esk-kak.*2 stems. In the case of PtGUX2 over-expression, both Pt2*esk-kak*.1 and Pt2*esk-kak*.2 release U_2X_5, with line one digestion releasing nearly only the doubly glucuronidated product.

To compare the amount of [Me]GlcA present on xylan of WT, *esk1-5:kak-8* and conifer GUX expressing *esk1-5:kak-8* plants the intensity of peaks corresponding to the glucuronidated and unbranched oligosaccharides was quantified (Figure 5.6B and C) for three biological replicates of basal stem AIR. As previously reported (Grantham et al., 2017), the degree of xylan glucuronidation was observed to be increased in *esk1-5:kak-8* plants. For both lines of the Pg1*esk-kak* and Pt2*esk-kak* plants the degree of [Me]GlcA branching was elevated even further, with the *Pg1esk*-kak.1 line exceeding 23% xylan glucuronidation and the Pt2*esk-kak*.1 line reaching more than 25% of xylan [Me]GlcA branching. Interestingly, for both PgGUX1 and PtGUX2 expressing *esk1-5:kak-8,* the majority of [Me]GlcA was contributed by the doubly glucuronidated GH11 digestion product (Figure 5.6D and E). This suggests possible changes in the [Me]GlcA patterning for both independent lines for both Pg1*esk-kak* and Pt2*esk-kak* genotypes. It is important to consider that in both Pg1*esk-kak* and Pt2*esk-kak* genotypes the conifer GUX enzymes do not act in isolation. Unlike in the case of PgGUX1 and PtGUX2 genotypes, in the Pg1*esk-kak* and Pt2*esk-kak* plants some xylan glucuronidation may be contributed by the Arabidopsis GUX enzymes. This may be somewhat limited by the use of an IRX3 promoter, which is stronger than the AtGUX ones (Schmid et al., 2005), to drive the expression of the conifer enzymes.



Band \ Line	Pg1esk-kak.1	Pg1esk-kak.2
XUXUXX	63.8%	58.0%
XUXX	36.2%	42.0%

Band \ Line	Pt2esk-kak.1	Pt2esk-kak.2
XUUXX	88.1%	72.5%
XUXX	11.9%	27.5%

Figure 5.6 Activity of PgGUX1 and PtGUX2 in the *esk1-5:kak-8 A. thaliana*. A) PACE analysis of GH11 xylanase digests of WT, two independent transgenic lines of *Pg1esk-kak* and *Pt2esk-kak* genotypes and control *esk1-5:kak-8* AIR. Undigested AIR controls are marked with (−). Quantitation of the degree of [Me]GlcA substitutions across WT, *esk1-5:kak-8* and Pg1*esk-kak* (B) and Pt2*esk-kak* (C) genotypes. Error bars represent standard deviation of three biological replicates of plant biomass. D) Degree of [Me]GlcA contributed by XUXUXX and XUXX in the two lines of Pg1*esk-kak* genotype. E) Degree of [Me]GlcA contributed by XUUXX and XUXX in the two lines of Pt2*esk-kak* genotype.

5.6 The pattern of [Me]GlcA branches on the xylan of the *esk1-5:kak-8 A. thaliana* is altered by PgGUX1 and PtGUX2 expression.

In order to investigate the pattern of [Me]GlcA branches in WT, *esk1-5:kak-8* and conifer GUX expressing acetylation mutants, AIR isolated from their biomass was digested with GH30. Released oligosaccharides were analysed on PACE (Figure 5.7). For WT and *esk1-5:kak-8* results matching those previously published were obtained, with WT predominantly releasing oligosaccharides of even DP (Bromley et al., 2013) and *esk1-5:kak-8* plants lacking this preference (Grantham et al., 2017).

For AIR isolated from the two lines of the Pg1*esk-kak* genotype the DP of released oligosaccharides was markedly decreased, suggesting more dense patterning of [Me]GlcA decorations. For both lines the dominant oligosaccharides had DP equal to 6, 4 and 2. The decrease in the abundance of longer oligosaccharides, compared to both WT and the *esk1-5:kak-8* transformation background, was proportional to the total degree of glucuronidation obtained in the transgenic lines. The Pg1*esk-kak*.1 line, with a higher total degree of glucuronidation (Figure 5.6) had a lower proportion of long DP oligosaccharide products than the Pg1*esk-kak*.2 line. The pattern of GlcA branches observed in both lines of Pg1*esk-kak* genotype is similar to the results obtained during PgGUX1 mediated *in vitro* xylan glucuronidation in the presence of xylan acetylesterase (see Figure 5.5 lane 4). In the case of AIR from Pt2*esk-kak* genotypes the [Me]GlcA pattern obtained is also similar to the one observed when the enzyme was used *in vitro* in the presence of the acetylesterase enzyme (Figure 5.5 lane 6). In Pt2*esk-kak*.1, which has a higher total degree of xylan glucuronidation, only doubly glucuronidated oligosaccharides were detected. In Pt2*esk-kak*.2, which has a lower degree of xylan glucuronidation, a mixture of doubly and singly glucuronated oligosaccharides was detected. This is consistent with the GH11 digestion results, where line one released predominantly doubly glucuronated oligosaccharides and line two released both XUXX and XUUXX.

Taken together with the *in vitro* activity results, this *in vivo* analysis, based on the expression of PgGUX1 and PtGUX2 in an *A. thaliana* mutant with decreased xylan acetylation, suggests that both the total amount of glucuronidation and the GlcA pattern generated by conifer GUX enzymes may be affected by the presence of acetyl branches on the xylan acceptor.

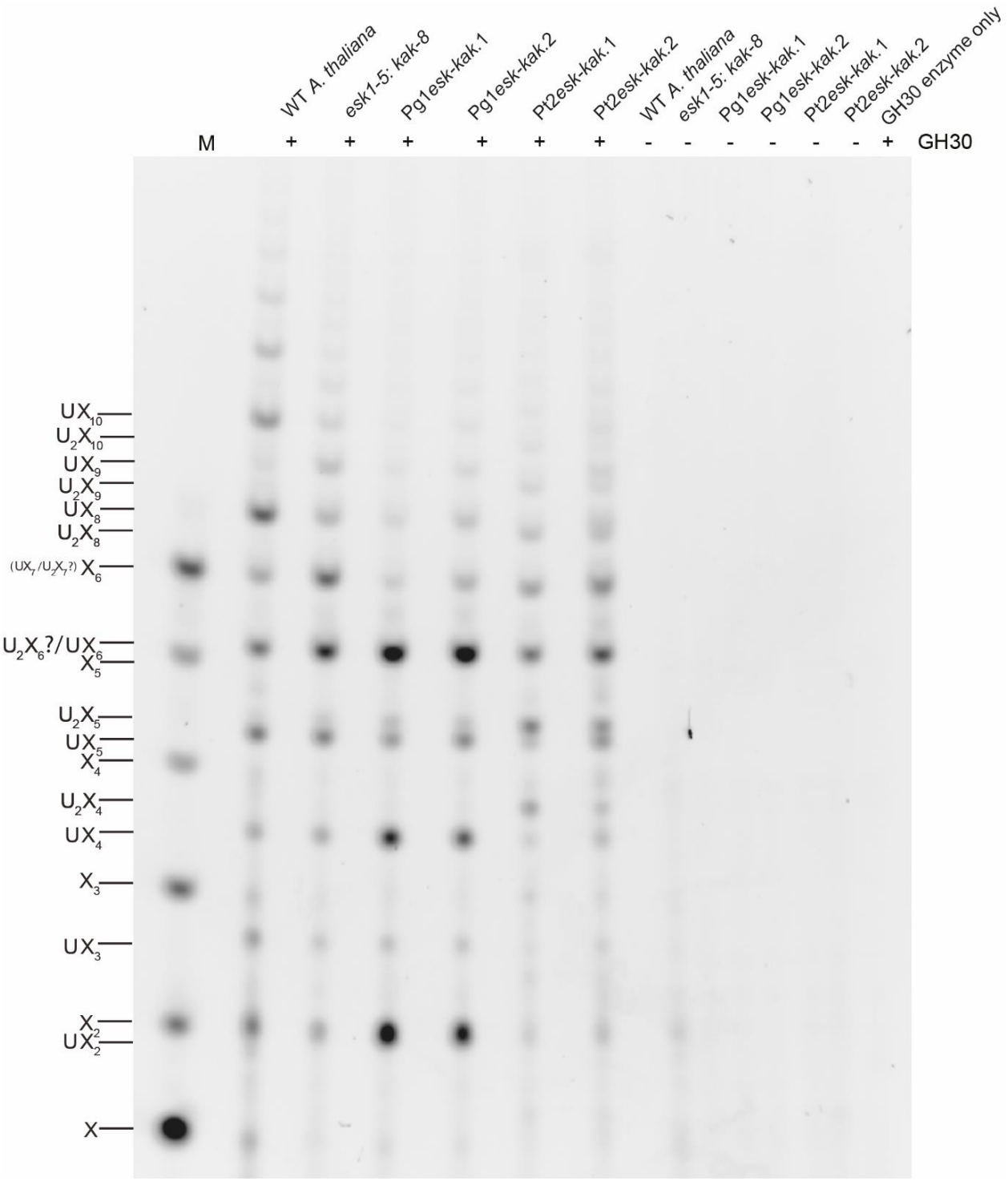


Figure 5.7 [Me]GlcA distribution on WT, *esk1-5:kak-8* and *esk1-5:kak-8 A. thaliana* plants expressing PgGUX1 and PtGUX2. AIR of all plants was hydrolyzed with glucuronoxylanase GH30 and analysed by PACE. Undigested material was used as a control. UX_3 and UX_2 bands was assigned by inference. Assignment of some doubly glucuronidated oligosaccharides was not possible due to co-migration with other structures and they were marked with a question mark.

5.7 Determinants of hardwood GUX activity

In order to compare the activity of the two conifer GUX enzymes to hardwood GUXs, *in vitro* activity assays were performed using Arabidopsis (At) GUX1 and GUX2 enzymes. The Arabidopsis proteins were expressed in the *N. benthamiana* system alongside the AtMAGT1 control protein. Successful expression of all three enzymes was confirmed using western blot (Figure 5.8A). Membranes enriched for both AtGUX1 or AtGUX2 were incubated with acetylated *gux1/2* xylan and a range of UDP-GlcA concentrations. Following 5h of incubation the polysaccharide was extracted from the reaction, de-acetylated and digested with xylanase GH11. Products of AtGUX1 (Figure 5.8C) and AtGUX2 (Figure 5.8E) dependent glucuronidation were analysed on PACE alongside products of a control reaction performed with AtMAGT1 enriched membranes (Figure 5.8E). For both AtGUX1 and AtGUX2 the glucuronidated XUXX product was observed in addition to xylobiose and xylose. No XUXX was detected in the MAGT1 catalysed reaction confirming that both AtGUX1 and AtGUX2 have specific activity in these reaction conditions.

The degree of xylan glucuronidation for each of the AtGUX1 and AtGUX2 reaction conditions was quantified and plotted against the concentration of UDP-GlcA used (Figure 5.8D and F). Obtained data points were evaluated with regression analysis to calculate the maximal theoretically possible degree of xylan glucuronidation by the two enzymes and half the concentration of UDP-GlcA required to achieve it (Figure 5.8B). The *in vitro* activity results indicate that in the assay conditions used AtGUX1 is less active than AtGUX2, but that they have similar requirements for the concentration of UDP-GlcA.

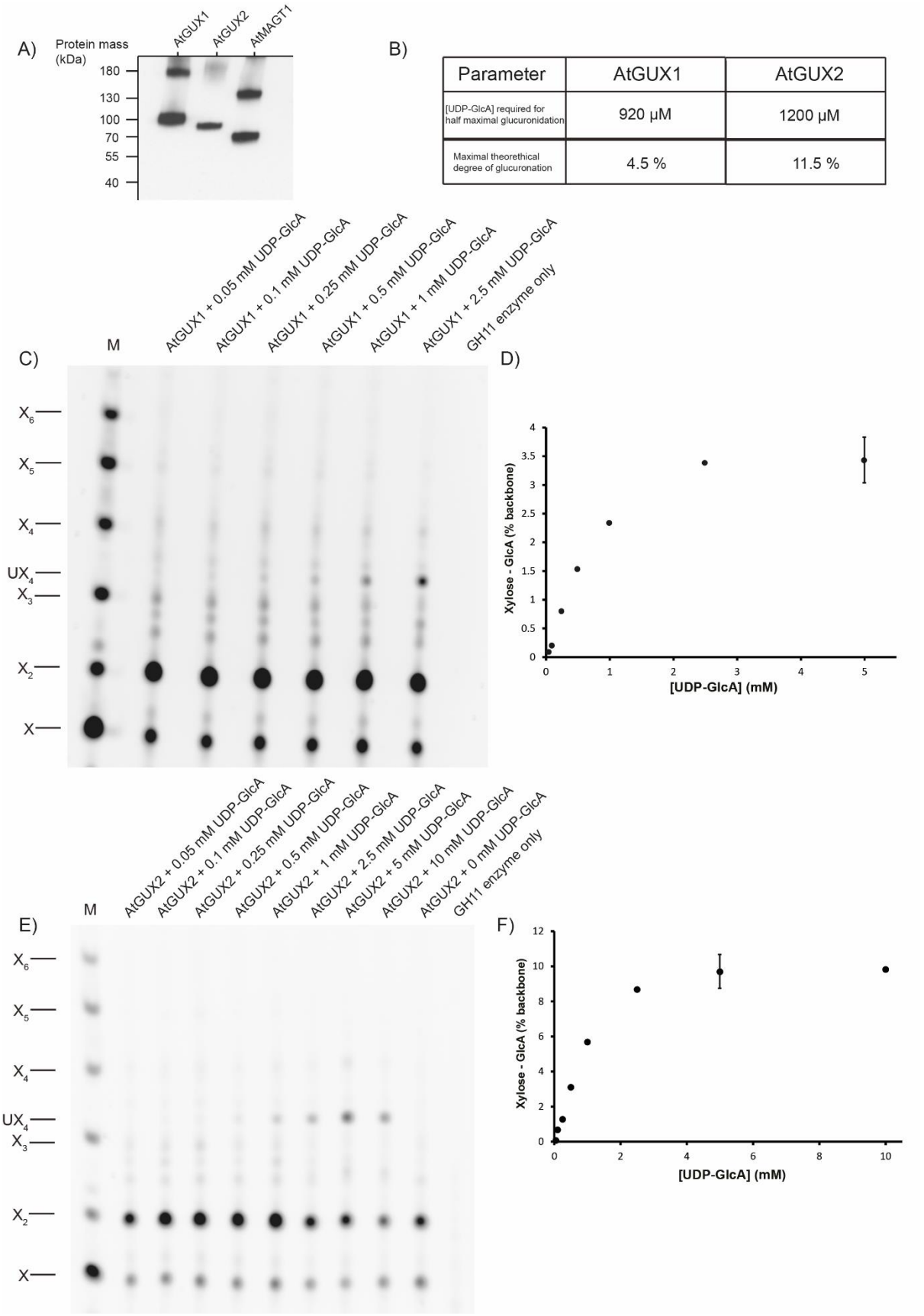


Figure 5.8 AtGUX1 and AtGUX2 activity *in vitro*. See next page for figure legend.

Figure 5.8 AtGUX1 and AtGUX2 activity *in vitro*. A) western blot analysis of membranes extracted from *N. benthamiana* infiltrated with the AtGUX1, AtGUX2 and AtMAGT1 expression constructs. Some dimers are evident. B) Table summarising theoretical *in vitro* activity parameters for AtGUX1 and AtGUX2 calculated using regression analysis C) Products of *in vitro* glucuronidation reaction with AtGUX1 and a range of [UDP-GlcA] digested with xylanase GH11 and analysed on PACE D) Dependence on [UDP-GlcA] of the degree of xylan glucuronidation for individual AtGUX1 reactions quantified using XUXX, XX and X band intensities. Data for 5 mM [UDP-GlcA] was obtained on a separate gel E) Products of *in vitro* glucuronidation reaction with AtGUX2 and a range of [UDP-GlcA] digested with xylanase GH11 and analysed on PACE F) Degree of xylan glucuronidation for individual AtGUX2 reactions quantified using XUXX, XX and X band intensities. For D) and F) the reaction using 5 mM UDP-GlcA was repeated three times using one batch of microsomal proteins, error bars represent standard deviation.

The lower activity of AtGUX1 is a surprising result as *in vivo* it is responsible for about 70% of total secondary cell wall xylan glucuronidation (Bromley et al., 2013). Hardwood xylan is also acetylated and in *esk1* plants, where the degree of Arabidopsis xylan acetylation is reduced, the AtGUX1 activity is changed (Grantham et al., 2017). Therefore, xylan acetylation may also be relevant for AtGUX1 and AtGUX2 dependant glucuronidation *in vitro*. In order to investigate this influence and see if it is one of the reasons for the low activity of AtGUX1 in this *in vitro* assay the reaction for both enzymes was repeated in the absence or presence of acetylxylan esterase. Inclusion of this enzyme in the assay had a significant impact on conifer GUX activity. The same reaction conditions as for the softwood GUX esterase supplementation experiments were used in the AtGUX1 and AtGUX2 experiments. Following completion of the reaction the polysaccharide product was extracted with ethanol, fully de-acetylated with an alkali treatment and digested with xylanase GH11. Released oligosaccharides were analysed on PACE (Figure 5.9A). For both AtGUX1 and AtGUX2 xylose, xylobiose, xylotetraose with one GlcA branch and the previously characterised (Figures 5.2 and 5.3) xylohexaose with two GlcA branches were detected. Little difference in the overall band intensity was observed between the AtGUX1 reactions carried out in the presence and absence of acetyl xylan esterase. The U_2X_6 band had stronger intensity in the oligosaccharides released from the xylan product of the AtGUX2 dependant glucuronidation performed in the presence of acetyl esterase, when compared to the product generated without it.

In order to quantify any differences in the total degree of glucuronidation obtained in the reactions the intensity of the bands corresponding to all the oligosaccharide products released was measured and the total average degree of glucuronidation was quantified for three replicates of the experiment performed with one batch of microsomal proteins (Figure 5.9B). For both AtGUX1 and AtGUX2 the degree of xylan glucuronidation obtained in the over-night reaction was increased compared to the 5 h long reaction used to generate data presented on Figure 5.8. This suggests that extended reaction time and/or addition of UDP-GlcA during the glucuronidation process stimulates AtGUX activity. No significant difference in the degree of glucuronidation was measured for AtGUX1 reactions performed in the presence or absence of acetylxylan esterase. On the other hand, the inclusion of the esterase enzyme had a significant, positive, effect on the activity of AtGUX2. Enzymatic removal of acetyl branches increased AtGUX2 activity by approximately 30%. To further investigate the reason for the difference between AtGUX1 and AtGUX2 activity it will be necessary to repeat the experiments performed in the presence and absence of the esterase enzyme. Ideally, subsequent analysis should include standardisation of the reaction outcome for the total amount of enzyme used.

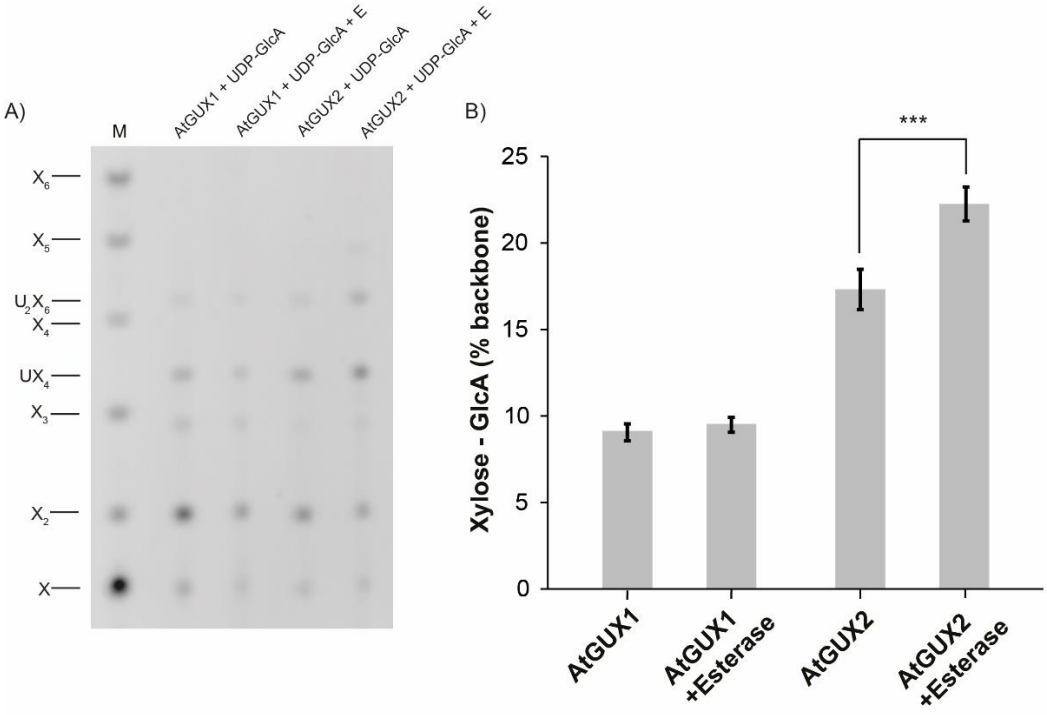


Figure 5.9 *In vitro* GlcA transfer by AtGUX1 and AtGUX2 onto acetylated and partially deacetylated xylan. A) PACE analysis of xylanase GH11 digestion of *in vitro* glucuronidation products catalysed by the two Arabidopsis enzymes. The reaction was carried out in the presence and absence of xylan acetyl esterase. B) Total degree of glucuronidation by AtGUX1 and AtGUX2 in the presence or absence of acetyl esterase enzyme in the *in vitro* reaction. Error bars represent standard deviation of three replicates of the reaction, *** is p value ≤0.001 in Student's t-test

In order to evaluate the pattern-specific activity of AtGUX1 and AtGUX2, xylanase GH30 digestion was performed on products of *in vitro* glucuronidation performed in the presence or absence of acetylxylan esterase. Released oligosaccharides were labelled with a fluorophore and analysed on PACE (Figure 5.10). In addition to a xylohexaose ladder, the released oligosaccharides were compared to previously annotated products of xylanase GH30 digestion of WT *A. thaliana* AIR.

Overall, the intensity of bands released from AtGUX1 dependant glucuronidation reactions was lower than the one of AtGUX2. This is in line with the lower overall degree of glucuronidation obtained in the AtGUX1 catalysed reactions (Figure 5.9). Digestion of the xylan produced by the AtGUX1 enzyme in the absence of the acetyl esterase resulted in the release of predominately UX_5, UX_6, UX_7 and UX_8 oligosaccharides. Some, very weak, bands corresponding to glucuronidated xylooligosaccharides of longer DP were also detected. A very similar oligosaccharide

profile was obtained when xylan products of AtGUX1 dependant glucuronidation performed in the presence of the esterase enzyme were digested with GH30. The somewhat weaker intensity of the longer DP oligosaccharides may have been due to the presence of esterase in the AtGUX1 reaction performed. This may be an artefact of the small amounts of oligosaccharides released and thus could reflect un-even loading across the wells. Therefore, to evaluate this possible reduction in the quantity of longer DP oligosaccharides further the digestion will need to be repeated.

For the AtGUX2 enzyme dependant glucuronidation, significant differences in the oligosaccharide profile released from the xylan product of the reaction performed in the absence or presence of acetylxylan esterase were observed. In the absence of the esterase enzyme AtGUX2 has catalysed the addition of the GlcA mainly onto every 4^{th}, 6^{th}, 8^{th} and 10^{th} xylosyl unit. Longer DP oligosaccharides were detected and they also showed a preference for an even DP. This suggests that in these reaction conditions AtGUX2 is capable of producing a xylan molecule with largely evenly spaced GlcA branches. Addition of the esterase enzyme resulted in a significant shift of the oligosaccharide profile. In these conditions AtGUX2 produced a more clustered pattern of GlcA branches, with the decorations present predominantly on every 2^{nd}, 3^{rd}, 4^{th} and 6^{th} xylosyl unit. This tight clustering of GlcA branches added by the AtGUX2 enzyme in the presence of acetylxylan esterase is in line with the increase in the total degree of glucuronidation when compared to the reaction lacking the esterase enzyme (Figure 5.9B). In addition to the described bands additional oligosaccharides were observed for both AtGUX2 reactions. These might be a result of GH30 side activity or possibly could correspond to doubly glucuronidated oligosaccharides observed previously for PtGUX2 reactions. The second hypothesis is unlikely as no evidence of XUUXX structure was seen in the GH11 digestion of AtGUX2 products (Figure 5.9). Considering these caveats the AtGUX2 reaction will need to be repeated. However, if confirmed, the results may indicate that the presence of acetyl branches on the xylan acceptor may have a different effect on the *in vitro* activity of AtGUX1 and AtGUX2. While AtGUX1 activity seems not to be altered by the partial removal of the acetyl esters the activity and specificity of AtGUX2 is changed by this procedure.

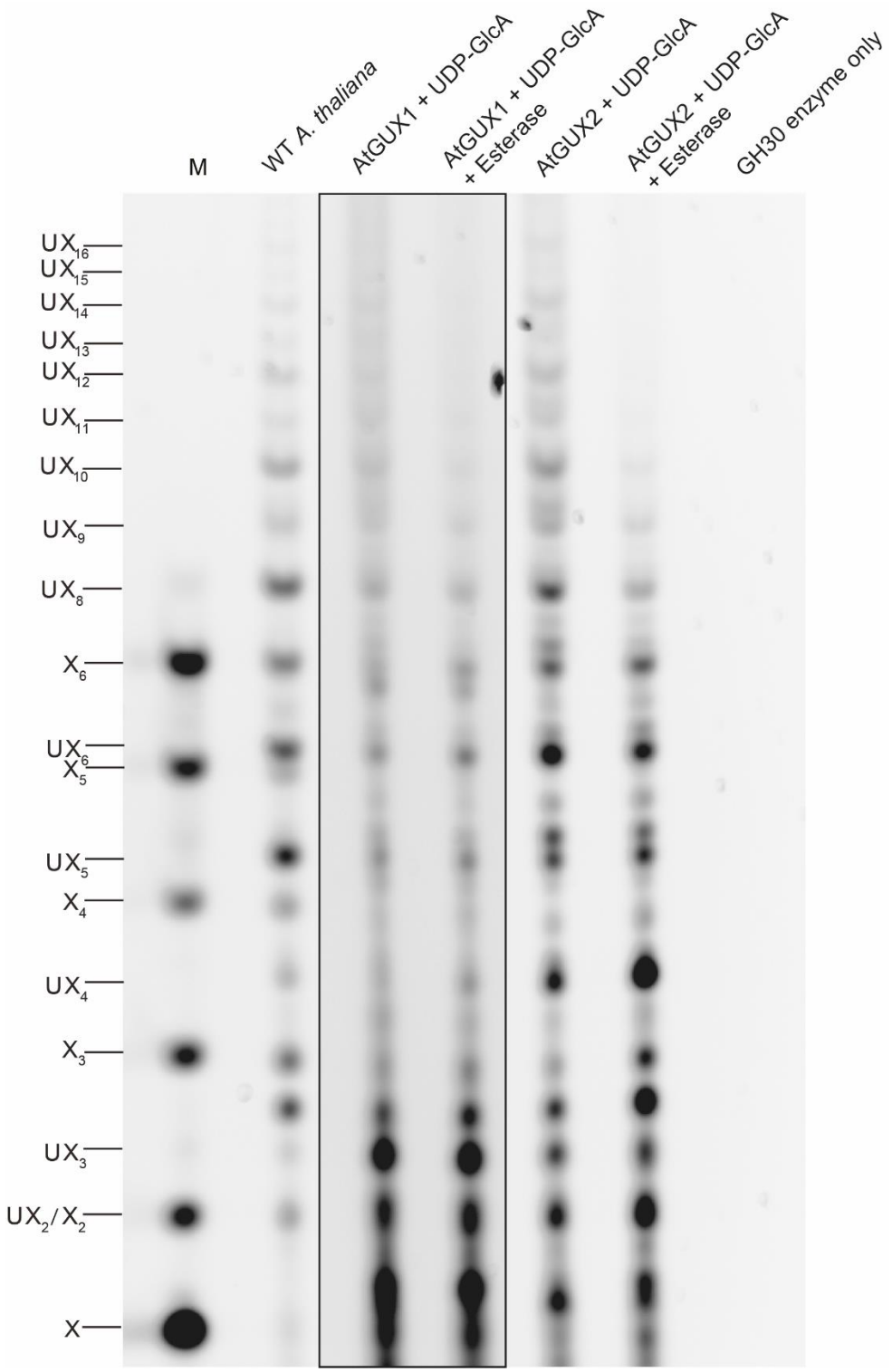


Figure 5.10 GlcA pattern generated *in vitro* by AtGUX1 and AtGUX2 in the presence or absence of xylan acetylesterase. Xylan product was deacetylated with NaOH and hydrolysed with glucuronoxylanase GH30 and analysed by PACE. Gel exposure of 5s was used for AtGUX1 reaction products (marked in black box) rather than 2s used for the remaining part of the gel. WT *A. thaliana* AIR was digested to act as a mobility standard for glucuronoxylooligosaccharides

5.8 Discussion

This chapter evaluated the specific activity of conifer GUX enzymes. Using a set of *in vitro* and *in vivo* activity assays it was demonstrated that the structure of the xylan acceptor influences GUX enzyme activity. Importantly, results discussed in this chapter indicate that both conifer GUX enzymes seem to have activities that may be related to the biological function of xylan regions that they glucuronidate. Additionally, they are distinct from the Arabidopsis AtGUX1 and AtGUX2 activities.

Some experimental limitations need to be overcome to provide more conclusive data on the activity of the GUX enzymes from two distinct clades. Specifically, it may be necessary to evaluate further members of each conifer GUX clade prior to the assignment of clade-specific activities. In Arabidopsis AtGUX1 and AtGUX3 belong to the same clade (Figure 4.1 Chapter 4). However, the enzymes, though they have broad similarities, show distinct activities and act in different tissues (Mortimer et al., 2015). As in each conifer transcriptome analysed only one protein product has been identified for each of the GUX clades a situation similar to the one in Arabidopsis is unlikely. However, to conclude with a greater degree of certainty, analysis of further gymnosperm transcriptomes and further members of each conifer GUX clade is needed.

In addition to analysis of further members of each conifer GUX clade it may be necessary to perform additional *in vitro* activity assays to understand the determinants of conifer GUX activity better. Assays performed as part of this chapter were not standardised for the amount of the protein used and therefore each interpretation assumes that the reaction has run to completion, irrespective of the amount of enzyme used. To overcome this issue, purified GUX enzymes with known concentrations could be used to perform assays in which kinetic parameters such as K_M and V_{max} can be measured. In addition to that, it would be interesting to perform assays using different xylan acceptors. In particular, it would be relevant to investigate the impact of xylan arabinosylation on conifer GUX activity. This may be performed using commercially available wheat arabinoxylan as an acceptor or by extracting spruce xylan and removing MeGlcA branches with GH115. The analysis of AtGUX1 and AtGUX2 enzymes would also benefit from further assays. In particular, by knowing the amount

of protein in the assay one could be more confident when concluding that in the conditions used AtGUX2 is more active than AtGUX1.

The results of the *in planta* analysis presented in this chapter also require further analysis. It is particularly important to consider that in both the Pt1*esk-kak* and Pt2*esk-kak* genotypes the Arabidopsis AtGUX1 and AtGUX2 enzymes may also contribute to the observed pattern of [Me]GlcA branches. Interestingly, the pattern of GlcA branches on xylan generated by PgGUX1 and PtGUX2 *in vitro,* when acting on a partially de-acetylated acceptor, is similar to the one observed in Pg1*esk-kak* and Pt2*esk-kak* xylan (Figures 5.5 and 5.7). Therefore, the contribution of AtGUX1 and AtGUX2 enzymes to the pattern may be minimal. In order to investigate this further it may be beneficial to transform *esk1-5:gux1:gux2* Arabidopsis mutant with the p*IRX3*::*PgGUX1* and p*IRX3*::*PtGUX2* constructs. The transformation of the triple mutant may be impossible as the plant is severely dwarfed, probably due to aggregation of unbranched xylan (Grantham et al., 2017). However, it may be possible to transform a plant heterozygous for one of the mutated Arabidopsis genes and select a transformed triple mutant within its progeny.

5.8.1 The structure of the xylan acceptor influences GUX activity

Experiments presented at the start of this chapter, together with previously published work (Busse-Wicher et al., 2016b) indicate that gymnosperm GUX1 enzymes seem to generate a different pattern of [Me]GlcA branches depending on the presence or absence of xylan acetylation and arabinosylation (Figure 5.1). Gnetophyte and conifer GUX1 enzymes have sequence similarity exceeding 80% (Figure 4.1 from chapter 4). Yet, when the conifer gymnosperm enzymes act in the native conifer background, in which in addition to being glucuronidated xylan is also arabinosylated, the pattern of GlcA branches is different compared to what is observed when a Gnetophyte gymnosperm GUXs, sharing a high amino acid sequence similarity to the conifer ones, catalyse the addition of GlcA branches onto acetylated, non-arabinosylated, gnetophyte xylan. Interestingly, the pattern of GlcA branches observed in the *A. thaliana gux1/2/3* plants expressing the PgGUX1 enzyme seems to closely resemble the one observed in gnetophyte plants (Figure 5.1) suggesting activity of gymnosperm enzymes may be influenced by presence of acetyl branches on xylan. Evidence exists that in *Arabidopsis* some xylan acetylation precedes glucuronidation (Grantham et al.,

2017). Thus, in the case of both PgGUX1 lines and gnetophyte plants it is possible that the gymnosperm enzyme may be adding GlcA branches onto an acetylated xylan molecule. To investigate this further, it may be necessary to evaluate the pattern of [Me]GlcA branches generated by Gnetophyte GUX1 enzyme when expressed in *gux1/2/3 A. thaliana.* It may also be interesting to study it using *in vitro* activity assays performed in the presence and absence of the acetylxylan esterase enzyme. Ara branches may also be important for the regulation of conifer GUX patterning. It is possible that, similarly to the role of acetylation for hardwood GUX, Ara substitutions may guide the positioning of conifer GlcA branches. This can be evaluated using *in vitro* activity assays using arabinoxylan as an acceptor as discussed in the first part of this discussion. Results presented in the first section of this chapter do not evaluate the importance of Ara branches but as discussed they do indicate that xylan acetylation may have an effect on the pattern of GlcA decorations added by conifer GUX1 in PgGUX1 lines.

To evaluate this hypothesis experiments in which xylan acetylation is decreased prior or during conifer GUX dependant glucuronidation were performed. Results indicate that *in vitro* and *in vivo* the maximal amount of GlcA transferred onto xylan by both PgGUX1 and PtGUX2 is increased when xylan is partially deacetylated. The pattern of GlcA branches added by the two enzymes has also changed upon xylan deacetylation. Interestingly, the PgGUX1 enzyme, when acting on a partially deacetylated xylan acceptor, generated largely compatible xylan molecules. In the same conditions, PtGUX2 catalysed the addition of GlcA branches almost solely onto consecutive xylosyl units, spaced 5, 6 and 7 monomers apart. The importance of this discovery is evaluated in parts 5.7.2 of this discussion section.

The work described in this chapter is not the first one to observe the effect of polysaccharide acceptor structure on the activity of branching enzymes. One striking example of such specificity was recently described in galactoglucomannan biosynthesis (Yu et al., 2018). In this work a range of glucomannan oligosaccharides was used to assay the activity of the AtMAGT1 galactosyltransferase. Interestingly, the enzyme was capable of adding galactose branches solely onto acceptors with a minimum of two consecutive glucose-mannose disaccharides in the structure. No activity was detected when oligosaccharides lacking this structure were used as acceptors. In xylan biosynthesis, the activity of xylan acetyltransferases, TBL

enzymes, may also require specific polysaccharide structures (Zhong et al., 2017). In particular, the TBL32 and TBL33 enzymes may be only capable of adding acetyl branches to carbon 3 of xylose monomers which carry GlcA decorations. Specificity of branching enzymes was also observed in the biosynthesis of other hemicelluloses. For example, MUR3 is responsible for addition of galactose specifically onto the third Xylose of the repeating XXXG structure of xyloglucan (Pauly and Keegstra, 2016).

5.8.2 Conifer GUX enzymes may be responsible for the synthesis of different xylan structures *in planta*.

Patterning of GlcA branches is a significant determinant of compatibility of xylan molecules for their interaction with the hydrophilic face cellulose fibril. In the *A. thaliana* model plant, the AtGUX1 enzyme is responsible for the addition of evenly spaced GlcA branches whereas the AtGUX2 enzyme adds GlcA branches without this preference (Bromley et al., 2013). Therefore, in a two-fold screw conformation, AtGUX1 patterned compatible xylan regions have GlcA only on one face of the polysaccharide and the other surface is free for hydrogen bonding with the cellulose fibril (see Busse-Wicher et al., 2016a and section 1.10.2 of Chapter 1). The AtGUX2-dependant pattern of GlcA branches is known as incompatible as it places the decorations on both faces of the two-fold screw and thus is likely to inhibit the interaction between xylan and the hydrophilic surface of the cellulose microfibril.

Similarly to the *A. thaliana* model, both compatible and incompatible GlcA patterns have been detected in softwood. The majority of the softwood xylan is compatible and has the MeGlcA branch on every 6^{th} monomer (Busse-Wicher et al., 2016b). However, a small proportion of the molecule has an incompatible branching pattern with GlcA on consecutive monomers (Martinez-Abad et al., 2017). Interestingly, the proportion of xylan with consecutive MeGlcA branches differs between the alkali and hot water extracted xylans, with the latter having a greater amount of the possibly incompatible structure. This may indicate that, unlike in hardwood, the compatible and the incompatible xylan patterns may exist on different polymers in softwood or that the proportion of these structures varies across the xylan molecules.

Results presented in this chapter indicate that the two conifer GUX enzymes may be responsible for the synthesis of different patterns of GlcA branches. Specifically, clade one enzymes (such as PgGUX1) might synthesise compatible xylan whereas clade

two enzymes (such as PtGUX2) may be involved in the biosynthesis of xylan molecules with incompatible patterning. This conclusion is supported by the results of *in vitro* and *in vivo* GUX activity experiments presented in this chapter. In particular, by decreasing the acetylation of the xylan acceptor it was possible to shift conifer GUX activities towards generating specific types of patterning. As the conifer xylan is not acetylated (Scheller and Ulvskov, 2010) it is possible that these low acetylation conditions are more alike to the ones in which conifer GUX enzymes exert their activity *in planta.* In order to be confident that these activity differences are specific to different clades of conifer GUX enzymes it is necessary to evaluate further members of PgGUX1 and PtGUX2 clades. Conclusive evidence regarding the specificity of conifer GUX enzymes may be obtained using conifer plants mutated selectively in either GUX1 or GUX2. This approach is discussed further in the final chapter of this thesis.

When acting in the low acetylation *esk1-5:kak-8* background, or upon enzymatic removal of acetyl branches from the xylan acceptor *in vitro*, the PgGUX1 enzyme catalysed addition of GlcA branches mainly onto every 6^{th}, 4^{th} and 2^{nd} xylose monomer (Figures 5.5 and 5.7). These are the dominant distances between Ara and GlcA branches observed over the majority of mature softwood xylan (Busse-Wicher et al., 2016 and Figure 5.11). In this model Ara branches are present either 2 or 4 units away from MeGlcA and the acidic branches are spaced by 6 monomers. Therefore, the result of *in vivo* and *in vitro* activity assays indicate that clade one conifer GUX enzymes may be responsible for the addition of GlcA branches with a compatible pattern to the majority of conifer xylan and that their activity may be influenced by the presence of Ara decorations. Of course, this hypothesis assumes that in the assay conditions used the spacing defined *in planta* by Ara branches would be provided by other MeGlcA decorations, possibly due to interactions with binding sites in PgGUX1. Therefore, the hypothesis has to be tested further by analysing PgGUX1 activity on arabinoxylan. For PtGUX2, the removal of xylan acetylation enabled this enzyme to place GlcA predominately on consecutive xylose monomers. This may indicate that clade two of conifer GUX enzymes may be responsible for generating these incompatible xylan structures *in planta*.

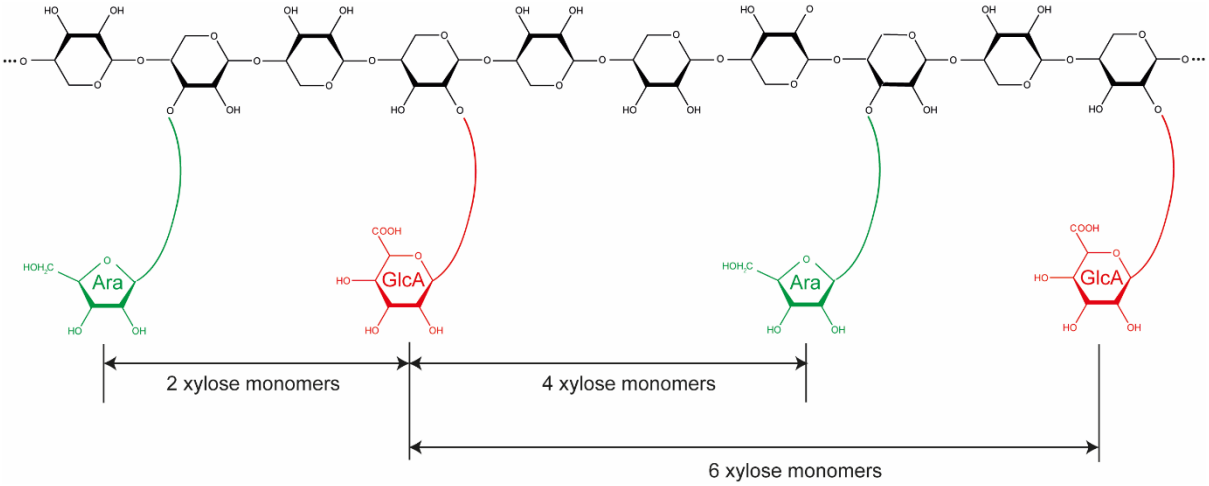


Figure 5.11 Dominant xylan structure in softwood (modified after Busse-Wicher et al., 2016). Distances of 2, 4 and 6 xylose residues can be detected between GlcA (red) and arabinose (Ara, green) xylan branches.

Results presented in this chapter indicate that the two conifer GUX enzymes may be responsible for the synthesis of different xylan structures *in planta.* Two independent groups have used molecular dynamics simulations to evaluate the interaction between the compatible regions of conifer xylan and cellulose (Busse-Wicher et al., 2016b, Martinez-Abad et al., 2017). Both reports indicate that the compatible structure, which based on the results presented in this chapter is likely to be synthesised by the PgGUX1 enzyme, can form a stable interaction, as a two-fold screw, with the hydrophilic face of the cellulose microfibril. This may suggest that PgGUX1-dependent glucuronidation can be involved in the synthesis of conifer xylan regions which interact with the cellulose fibril. Interestingly, the work described by (Martinez-Abad et al., 2017) also evaluated the interaction between cellulose and a xylooligosaccharide with GlcA branches on consecutive monomers. The group was able to detect stable interaction of this structure with both hydrophobic and hydrophilic surfaces of the cellulose fibril. However, unlike in the case of the evenly spaced xylo-oligosaccharide the presence of GlcA on consecutive monomers restricts binding of further xylan molecules to the cellulose (Figure 5.12A). Moreover, the group has only used a hexagonal shaped model of the cellulose crystal to evaluate formation of the xylano-cellulose complex. Previous work has indicated that the conifer cellulose may form a different crystalline structure with a square cross-section (Figure 5.12B) (Fernandes et al., 2011). Therefore, further work is required to evaluate the interaction between different conifer xylan structures and the square cellulose crystal. However, it is unlikely that the GlcA branches present on consecutive xylose monomers could be

accommodated in grooves of the hydrophilic surface of the square cellulose crystal. Taken together, this analysis and the data presented in this chapter indicate that the PtGUX2 enzyme is likely to be responsible for the synthesis of polysaccharide structures which limit the number of xylan chains that can bind the cellulose fibril, interacts only at hydrophobic surfaces, or cannot interact with the cellulose microfibril at all.

Figure 5.12. Impact of consecutive glucuronidation on xylan-cellulose interaction. A) Adapted after (Martinez-Abad et al., 2017), when bound to hydrophilic face of the cellulose fibril xylooligosaccharide with consecutive glucuronidations renders an additional binding surface (red arrow) inaccessible. This does not happen for compatible conifer xylan regions (green arrow). B) Presented after (Busse-Wicher et al., 2016b), binding of acetylxylan to the hydrophilic surface of the square cellulose fibril. Unlike patterned acetylxylan, it is unlikely that xylan with consecutive glucuronidations could fit into the binding groove on the hydrophilic surface of this cellulose crystal.

5.8.3 Hardwood GUX enzymes have a different activity *in vitro* and *in vivo*

In addition to investigating the specificity of conifer GUX enzymes, work discussed in this chapter has also analysed the activity of *A. thaliana* GUX1 and GUX2 enzymes using the *in vitro* assay in which the GlcA is added onto a polymeric, acetylated xylan acceptor. The *in vitro* activity of AtGUX1 and AtGUX2 was previously reported. Rennie et al., 2012 used a similar tobacco expression system to produce all five putative *A. thaliana* GUX enzymes and studied their activity using xylohexaose as an acceptor and ^{14}C UDP-GlcA as a substrate. Measurements of UDP-GlcA incorporation have indicated that, when incubated with AtGUX1 and AtGUX2, more signal was transferred

onto the acceptor when compared with the control reactions. Although significant, the difference between control and GUX mediated glucuronidation efficiency was small. AtGUX1 transferred four times more GlcA than the control and AtGUX2 three times more. In the work described in this chapter no measurable UDP-GlcA transfer onto xylan is detected in the control reactions, possibly because the sensitivity of the detection method is low. When the total degree of glucuronidation obtained *in vitro* is measured, the assay used in the work described in this chapter achieves more total transfer efficiency than what was obtained by Rennie et al., 2012 (Table 5.1). This discrepancy may be partially associated with protein amounts as Rennie et al., 2012 used a different vector for the *N. benthamiana* expression than pEAQ-HT used in this work. Alternatively, this data may indicate that the use of polymeric acetylated xylan as an acceptor allows more efficient measurement of GUX mediated GlcA transfer *in vitro*. This may suggest that, rather than glucuronidating short oligosaccharides, hardwood GUX enzymes act on polymeric xylan *in planta.* To evaluate this hypothesis further, *in vitro* activity assays using AtGUX expressed with pEAQ-HT could be performed using xylohexaose as an acceptor.

Table 5.1 AtGUX1 and AtGUX2 activity in this study and reported by (Rennie et al., 2012)

Parameter	AtGUX1	AtGUX2
Maximal theoretical degree of glucuronidation on acetylated polymeric xylan (This study)	4.5%	11.5%
Maximal degree of glucuronidation on xylohexaose (X6) (Rennie et al., 2012)	~0.004%	~0.003%

The results presented in this chapter indicate unexpectedly that AtGUX2 has higher *in vitro* activity than AtGUX1. While the maximal calculated degree of glucuronidation obtained with AtGUX1 is equal to 4.5%, the AtGUX2 enzyme can add, in the 5h

reaction, GlcA branches to as much as 11.5% of the backbone. These results differ from the observations of Rennie et al., 2012 where AtGUX1 was more active than the AtGUX2. What is more, *in vivo* the AtGUX1 enzyme is responsible for the majority of secondary cell wall xylan glucuronidation (Bromley et al., 2013). In addition to these differences in the total degree of xylan glucuronidation obtained with the two Arabidopsis enzymes studied, their specific activity also differed from what is observed *in vivo*. In particular, in the *in vitro* assay performed without the esterase both AtGUX1 and AtGUX2 were able to synthesise xylan molecules with a largely compatible pattern of GlcA branches. This is different to *in vivo* conditions where AtGUX1 produces compatible GlcA patterning while AtGUX2 adds the acidic decorations in an incompatible manner.

Interestingly, partial removal of acetyl branches *in vitro* led to an increase in AtGUX2 activity and enabled the enzyme to generate a pattern largely similar to the one it synthesises *in vivo*. The removal of acetyl branches had little impact on the activity of AtGUX1. This observation is different to the reports of *in vivo* analysis where AtGUX1 activity was detected to be altered in *esk1-5* which has reduced levels of xylan acetylation, whereas AtGUX2 was relatively unaffected (Grantham et al., 2017). Together, these observations indicate that in hardwood factors such as the timing of synthesis, presence of other xylan biosynthesis enzymes or localisation of GUX proteins to different Golgi stacks may contribute to the process of xylan pattern formation more than the GUX protein sequence itself. Importantly, *esk1* mediated xylan acetylation is required for the generation of a complete hardwood xylan [Me]GlcA patterning. It is unclear what the acetylation state is of *gux1/2* acceptor used in the presented *in vitro* analysis. It is possible that the *gux1/2* acceptor is over-acetylated due to absence of GUX activity during its synthesis. To evaluate if these factors might have influenced AtGUX1 and AtGUX2 activity it might be beneficial to use WT xylan as an acceptor following removal of [Me]GlcA branches using GH115, possibly in conjunction with a regiospecific acetylxylan esterase.

5.8.4 Synthesis of patterned xylan molecules in softwood and hardwood *in vivo*.

Taken together, the results presented in this chapter indicate that a range of factors are likely to define the patterning of xylan molecules in hardwood and softwood. The work has indicated that some specific differences in the synthesis of the xylan molecule exist between the two groups. In particular, it appears that the softwood xylan patterning may be set by the activity of GUX enzymes themselves, with a possible contribution from Ara decorations guiding conifer GUX1 activity. In hardwood xylan synthesis the structure of the final polysaccharide product might be influenced by other phenomena to a greater extent than in conifers, with acetylation having a particularly significant influence on GlcA patterning.

For both hardwood and softwood xylan it is unknown how the proportion of compatible and incompatible patterns is maintained. The total degree of glucuronidation obtained by both conifer GUX enzymes studied seems to be similar *in vitro* (Figure 5.4). However, in the final polysaccharide product in the conifer, the abundance of the structures synthesised by the two enzymes differs. While the compatible GlcA patterning, likely to be synthesised by PgGUX1, covers the majority of the xylan molecules the PtGUX2 dependant consecutive glucuronidation can only be detected at low levels. This indicates that specific processes govern the softwood xylan glucuronidation process *in vivo*. In particular, there appears to be a need for tight control of the extent to which the two GUX enzymes exert their catalytic activity. This may be achieved by localisation of the enzymes to different stacks of the Golgi Apparatus or by having specific backbone biosynthesis processes, dedicated to the manufacturing of PgGUX1 and PtGUX2 patterned xylans. Alternatively expression of the genes encoding specific enzymes may be controlled on a temporal basis.

For the hardwood GUX, the activity of AtGUX1 and AtGUX2 differed significantly from each other in the *in vitro* assay. In particular, unlike *in planta,* the AtGUX2 had much higher activity than AtGUX1 and it was able to synthesise a compatible xylan pattern *in vitro*. This indicates that in the *in vivo* conditions it is likely that AtGUX1 is the first to add GlcA branches onto xylan, in this way not an entire xylan molecule is glucuronidated by the more active AtGUX2 enzyme. This may be achieved by localisation of the different hardwood GUX enzymes to different stacks of the Golgi Apparatus. Moreover, results presented in this chapter indicate that it is likely that the

AtGUX2 acts on partially de-acetylated xylan. This partial de-acetylation may be achieved following TBL mediated acetyl group addition by GDSL family of esterases which were reported to influence xylan structure in rice (Zhang et al., 2017).

Overall, the work presented in this chapter indicates strongly that hardwood and softwood xylan is likely to have compatible and incompatible xylan regions synthesised by distinct GUX enzymes. This in turn raises important questions about the biological role for these xylan structures. Are wood properties altered by the proportion of compatible and incompatible xylan regions? Can conifer plants, in response to external stimuli, alter the structure of xylan molecules that are being synthesised? In order to evaluate some of these questions the next chapter will use the *A. thaliana* model to investigate possible biological roles of xylan-cellulose interaction.

Chapter 6: The Biological function of xylan- cellulose interaction

6.1 Introduction

The results presented in the previous chapter indicate that conifer xylan molecules have structures which may both facilitate and inhibit its interaction with the hydrophilic surface of the cellulose microfibril, according to the model presented in the introduction. The exact biological role of the arising xylan–cellulose complex, also known as the xylanocellulose fibril, remains unknown, but the structures were proposed to be important for the maintenance of cell wall network integrity (Simmons et al., 2016). To investigate the role of this xylan-cellulose interaction it may therefore be beneficial to use a plant model in which these polysaccharides do not form stable complexes. Recent reports indicate that in *A. thaliana esk1* mutant, the xylan-cellulose interaction is likely to be lost completely or may occur in a different manner than in WT plants (Grantham et al., 2017). The two-fold screw xylan, indicative of bound xylan, was absent in these plants, according to solid state NMR studies. *A. thaliana ESK1* is a xylan specific O-acetyltransferase responsible for the majority of xylan acetylation. Plants with *ESK1* mutation show an overall ~60% reduction in acetate branching of xylan (Xiong et al., 2013). This is associated with changes in the activity of GUX enzymes. Specifically, in *A. thaliana esk1* plants, GUX1 may not be able to generate compatible [Me]GlcA patterning (Grantham et al., 2017). Thus, the pattern of GlcA branches in xylan of *esk1* plants is largely incompatible, which was proposed to be the reason for loss of the normal xylan-cellulose interaction in this mutant (Grantham et al., 2017). In addition to that, *esk1* plants have collapsed vessels and are dwarfed (Lefebvre et al., 2011, Xiong et al., 2013, Yuan et al., 2013), supporting the view that xylan, and perhaps its normal interaction with cellulose, is important for cell wall strength.

Three different approaches have been developed to rescue the growth phenotype of *esk1* plants. Firstly, a secondary mutation in *KAKTUS* (*KAK)* gene was observed to allow *esk1kak* plants to grow significantly better than the *esk1* alone (Bensussan et al., 2015). This rescue is believed to be achieved mainly by the formation of a larger number of xylem vessels allowing for water transport to be more dispersed. This may prevent vessel collapse. Importantly, the xylan of *esk1kak* plants has the same, incompatible, structure as the *esk1* hemicellulose and it does not interact with cellulose (Grantham et al., 2017). In addition to the *kak* mutation, mutation in a promoter region

leading to the reduction in *MORE AXILLARY BRANCHES 4* (*MAX4*) expression was recently reported to improve the growth of *esk1:max4* plants compared to *esk1* single mutant (Ramirez et al., 2018). In addition to these two approaches, over-expression of AtGUX1 was also observed to rescue *esk1* growth (Xiong et al., 2015). It was proposed this may result from improved solubility of *esk1* xylan due to the presence of additional branches upon AtGUX1 over-expression. AtGUX2 was not reported as a suppressor of the *esk1* growth in that work. Interestingly, work described in chapter 5 of this thesis indicates that overexpression of conifer GUXs in *esk1-5:kak-8* plants leads to a change in [Me]GlcA patterning, with PgGUX1 generating largely compatible and PtGUX2 creating more incompatible xylan molecules. Therefore, it is possible that a change in [Me]GlcA patterning, leading to changes in association with cellulose, may be involved in AtGUX1-mediated rescue of the *esk1* growth phenotype.

Solid state nuclear magnetic resonance (ssNMR) is one of the techniques used in this chapter in order to evaluate this hypothesis by probing the molecular architecture of plant cell walls. In the ssNMR experiments a solid sample is spun at a high frequency (>12000 Hz in the presented experiments) at an angle (known as the magic angle) with respect to the magnetic field. This process greatly increases resolution of signal obtained from solid materials (Andrew, 2010). The analysis presented in this chapter uses stable ^{13}C-labelled plant material generated by growth of Arabidopsis plants in a $^{13}CO_2$ enriched atmosphere as described in the Methods chapter. Early ssNMR experiments on wood structure could not take advantage of complete ^{13}C labelling and relied on natural abundance of the ^{13}C isotope to generate signal. Due to low abundance of ^{13}C in plant biomass (<1%), the only possibly type of experiments involved analysis of 1D spectra of plant material which enabled identification of two distinct cellulose hardwood environments (Larsson et al., 1997) and annotation of some signals likely to be associated with hemicelluloses in softwood pulp (Hult et al., 2000). The use of 2D solid state NMR to analyse plant cell wall material was pioneered by the lab of Prof. Mei Hong at the Massachusetts Institute of Technology. Their work analysed mainly cell walls of plant seedlings grown on solid media containing ^{13}C-glucose. This allowed for efficient labelling of plant material and for example enabled experiments which provided indications that both pectins and xyloglucan may be interacting with the cellulose microfibril in the primary cell walls of Arabidopsis (Dick-Perez et al., 2011).

The ssNMR experiments looking at the secondary cell wall material, which are performed as part of this PhD project, were modified from existing methods for application to biomass by Prof. Ray Dupree and the lab of Prof. Paul Dupree (Dupree et al., 2015). In this research two main types of experiments are used to probe the different conformations of polysaccharides within the cell wall and evaluate their spatial proximity. Changes in the bond conformation of polysaccharides is associated with the alteration of chemical shifts attributed to their monosaccharide carbons in spectra obtained from cross-polarisation (CP)- refocussed INADEQUATE ssNMR experiments where magnetisation is transferred between covalently linked atoms. For xylan, the shift of carbon 4 is particularly affected by the change between the three-fold and two-fold conformation (Dupree et al., 2015). Specifically, the *in muro* three-fold (3f) xylan has a C4 signal at 77.4 ppm while the C4 of two-fold (2f) xylan has a peak at 82.2 ppm (Simmons et al., 2016). The CP-PDSD experiment in which magnetisation is transferred through space is used to probe the proximity of polysaccharides and the results of this analysis indicated that in WT Arabidopsis stems the 2f xylan is likely to be in direct contact with the cellulose microfibril (Simmons et al., 2016).

In order to investigate the possible role of the xylan-cellulose interaction, the work described in this chapter employs ssNMR and other techniques such as plant phenotyping and electron microscopy to:

- Investigate the impact of over-expressing various GUX enzymes on the growth of *esk1* plants;

- Evaluate the effect of GUX over-expression on the structure of *esk1* xylan molecules and their interaction with the cellulose microfibril;

- Explore the nanoscale cell wall structures which may be the location for the formation of xylanocellulose complexes;

6.2. Over-expression of some GUX enzymes rescues *esk1-5* growth phenotype

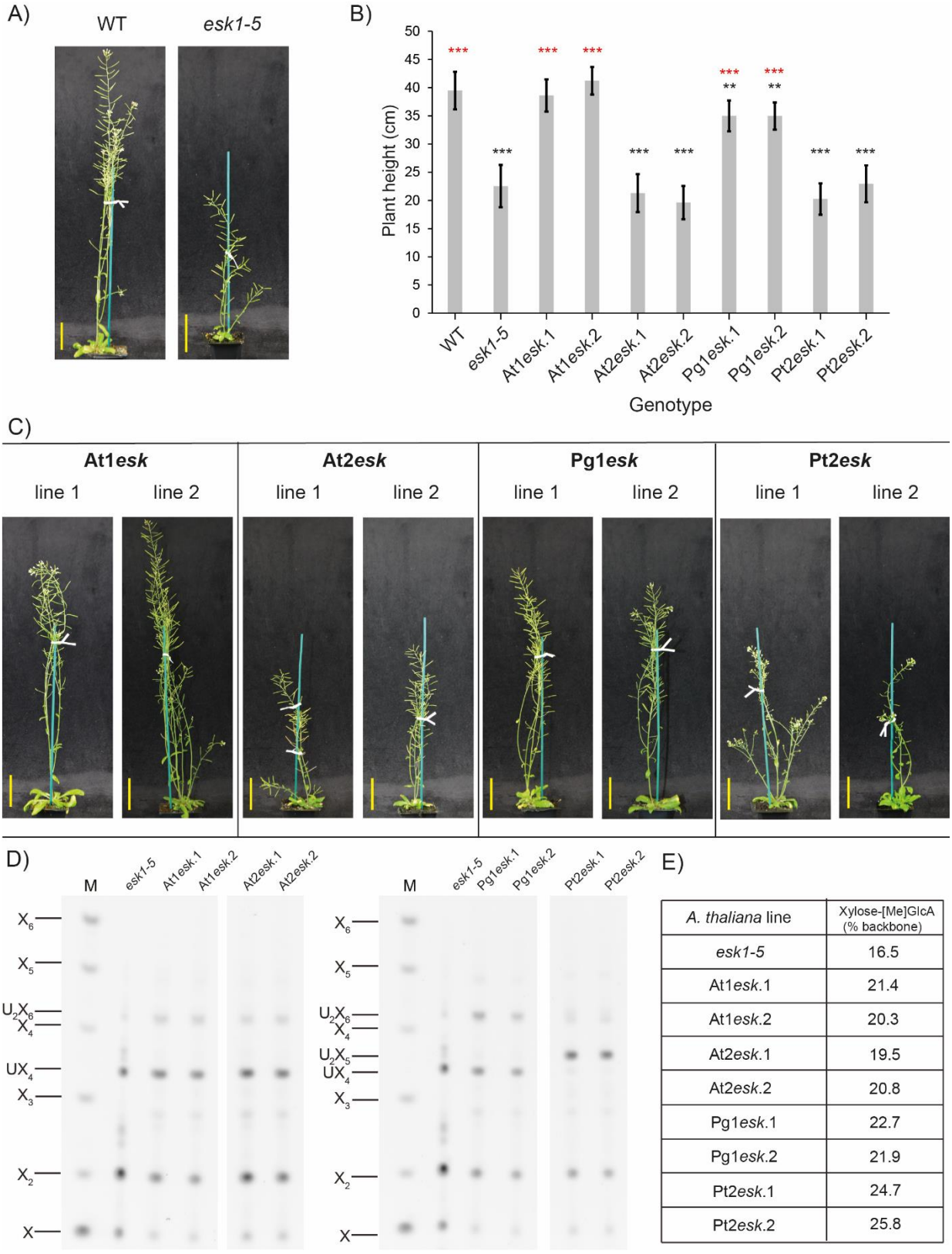
Expression of AtGUX1 enzyme in *esk1* mutant background has a positive effect on plant growth and restores it to be WT-like (Xiong et al., 2015). Interestingly, results of experiments presented in chapter 5 of this thesis suggest that the expression of conifer GUX enzymes changes the structure of *esk1-5:kak-8* xylan, with the plants expressing PgGUX1 generating largely compatible xylan molecules. However, as the *esk1-5:kak-8* plants used in this experiment show little to no dwarfing phenotype (Bensussan et al., 2015) it remains unknown if these changes in xylan structure have any positive effect on plant growth. In order to investigate this possible suppression of the growth phenotype, *esk1-5* dwarfed plants were transformed with constructs in which over-expression of AtGUX1, AtGUX2, PgGUX1 and PtGUX2 is driven by a strong secondary cell wall promoter: p*IRX3*. For each of the constructs two independent homozygous lines were isolated by screening the xylan of ten independent hemizygous lines for the highest degree of glucuronidation. Two independent hemizygous plants with the most [Me]GlcA on xylan were selected from each transformation event and used to generate homozygous lines with code names detailed in Table 6.1. The growth of the resulting homozygous plants was compared to that of WT and *esk1-5 A. thaliana.*

Table 6.1 Transgenic lines resulting from the transformation of *esk1-5* plants. p*IRX3* denotes a promoter region of the *IRX3* gene. Nos terminator was marked as t*NOS*

***GUX* gene expressed**	**Genetic construct used**	**Resulting transgenic lines**
At*GUX1*	p*IRX3*::AtGUX1-Myc::t*NOS*	**At1*esk*.1** and **At1*esk*.2**
At*GUX2*	p*IRX3*::AtGUX2-Myc::t*NOS*	**At2*esk*.1** and **At2*esk*.2**
Pg*GUX1*	p*IRX3*::PgGUX1-Myc::t*NOS*	**Pg1*esk*.1** and **Pg1*esk*.2**
Pt*GUX2*	p*IRX3*::PtGUX2-Myc::t*NOS*	**Pt2*esk*.1** and **Pt2*esk*.2**

In accordance with previously published work, the mutation of *ESK1* was observed to have a negative effect on plant growth (Grantham et al., 2017, Lefebvre et al., 2011, Xiong et al., 2013, Xiong et al., 2015). Stems of 7 week old *esk1-5* plants reached ~50% of the WT height (Figure 6.1A and B). Interestingly, transformation of *esk1-5* plants with constructs encoding different GUX enzymes had varying impact on the height of mutant plant stems (Figure 6.1B and 6.1C). As already reported (Xiong et al., 2015), transformation with a construct allowing for expression of AtGUX1 enabled restoration of *esk1-5* stem growth to WT levels for both At1*esk*.1 and At1*esk*.2 lines. Similarly, transformation with a construct enabling expression of PgGUX1 also had a significant positive effect on *esk1-5* stem height for both Pg1*esk*.1 and Pg1*esk*.2 genotypes. However, unlike in the case of AtGUX1, the *esk1-5* plants transformed with a construct enabling expression of PgGUX1 developed stems that were ~10% shorter than the ones of WT plants. Interestingly, transformation with genetic constructs enabling expression of both AtGUX2 and PtGUX2 had no significant influence on the height of *esk1-5* stems.

In order to investigate if the expressed GUX enzymes were active in the *esk1-5* background, AIR isolated from *esk1* plants hemizygous for the GUX over-expression constructs was de-acetylated with an alkali treatment and digested with xylanase GH11. The released oligosaccharides were analysed by PACE (Figure 6.1D). All analysed samples released U_2X_6, UX_4, xylobiose and xylose. In addition to that, AIR from the stems of Pt2*esk*.1 and Pt2*esk*.2 plants released the U_2X_5 oligosaccharide. Quantitation of the intensity of bands corresponding to glucuronidated oligosaccharides, xylobiose and xylose indicated that in the case of all transgenic lines the degree of xylan glucuronidation was increased when compared to the *esk1-5* background (Figure 6.1E). Therefore, all the GUX enzymes were active in all the lines studied. Interestingly, since only some of these lines showed improved growth, these experiments indicate that an increase in the degree of xylan glucuronidation alone is insufficient to rescue *esk1-5* growth.



A. thaliana line	Xylose-[Me]GlcA (% backbone)
esk1-5	16.5
At1*esk*.1	21.4
At1*esk*.2	20.3
At2*esk*.1	19.5
At2*esk*.2	20.8
Pg1*esk*.1	22.7
Pg1*esk*.2	21.9
Pt2*esk*.1	24.7
Pt2*esk*.2	25.8

Figure 6.1 Impact of GUX over-expression on *esk1-5* growth and xylan structure. See next page for full figure legend.

Figure 6.1 Impact of GUX over-expression on *esk1-5* growth and xylan structure. **A)** Representative images of WT and *esk1-5 A. thaliana* plants. **B)** Average height of WT, *esk1* and GUX overexpressing *esk1-5* plants. N = 60 plants per genotype grown once. Red asterisks indicate significant height difference when compared with *esk1-5.* Black asterisks indicate significant height difference when compared with WT. **p value ≤0.01; ***p value ≤0.001 Tukey test after ANOVA **C)** Representative images of GUX over-expressing *esk1-5* plants. **D)** PACE analysis of oligosaccharides released by GH11 digestion of *esk1-5* and GUX over-expressing *esk1-5* AIR **E)** Total quantified degree of glucuronidation of xylan in *esk1-5* and hemizygous GUX over-expressing *esk1-5* plants.

6.2 Over-expression of AtGUX1 and AtGUX2 has a different effect on the structure of *esk1-5:kak-8* xylan

To investigate the impact of GUX over-expression on the molecular architecture of *esk1-5* cell walls it was necessary to generate well growing plants over-expressing each GUX enzyme studied. This was required as the growth of biomass suitable for the solid state NMR analysis is challenging and dwarfed plants, like *esk1-5* over-expressing AtGUX2, do not perform well in the ^{13}C enrichment chamber (Grantham et al., 2017). In order to generate material suitable for this analysis, *esk1-5:kak-8* plants were transformed with the AtGUX1 and AtGUX2 over-expression constructs used in section 6.1. The use of *esk1-5:kak-8* background enables better plant growth while maintaining reduction in cell wall acetylation and associated disruption in xylan cellulose interaction (Grantham et al., 2017). For both transformation events two independent homozygous lines were isolated, following the same criteria as those applied in section 6.1 of this chapter, and studied. For the **AtGUX1** expression the two independent p*IRX3*::AtGUX1 *esk1-5:kak-8* lines were named **At1*esk-kak*.1** and **At1*esk-kak*.2**. For the **AtGUX2** expression the two independent p*IRX3*::AtGUX2 *esk1-5:kak-8* lines were named **At2*esk-kak*.1** and **At2*esk-kak*.2.**

To evaluate the impact of AtGUX1 and AtGUX2 over-expression on the structure of *esk1-5:kak-8* xylan, AIR isolated from WT, *esk1-5:kak-8* and homozygous over-expression lines was de-acetylated with an alkali treatment and digested with xylanase GH11. The released oligosaccharides were labelled with a fluorophore and analysed with PACE (Figure 6.2B). All analysed samples released glucuronidated xylotetraose and xylotriose oligosaccharides alongside unbranched xylose and xylobiose. In addition to that, AIR from *esk1-5:kak-8* mutant and the AtGUX overexpressing mutants released doubly glucuronidated xylohexaose. In order to investigate the degree of

glucuronidation in all the analysed plants the GH11 digestion experiment was repeated for three independent biological replicates of the plant material and the intensity of glucuronidated and non-glucuronidated bands was quantified. The results of this analysis (Figure 6.2B) indicated that expression of both AtGUX1 and AtGUX2 leads to a significant increase in the degree of *esk1-5:kak-8* xylan glucuronidation. Similarly to results presented in section 6.1, this indicates that both enzymes can catalyse further addition of GlcA onto xylan with a lower acetylation degree. Interestingly, similar experiments on fully acetylated xylan showed that only AtGUX1 over-expression could increase the degree of xylan glucuronidation (Bromley et al., 2013).

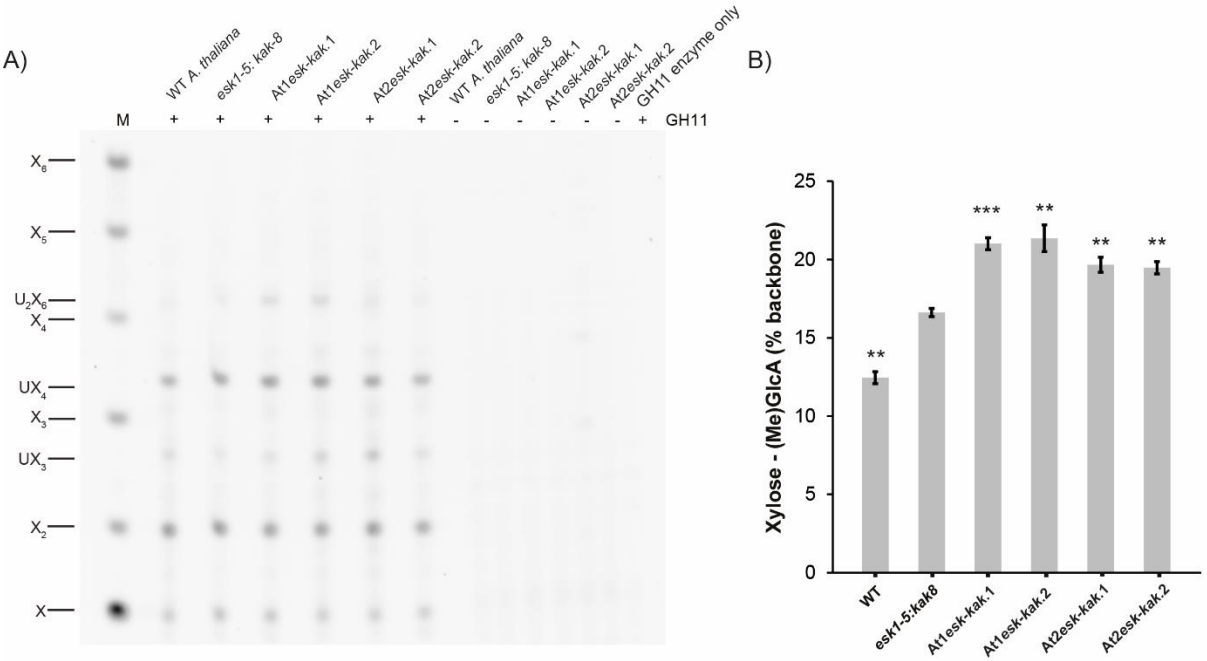


Figure 6.2 Impact of GUX over-expression on the degree of *esk1-5:kak-8* xylan glucuronidation. A) PACE analysis of oligosaccharides released by GH11 digestion of *esk1-5:kak-8* and AtGUX over-expressing *esk1-5:kak-8* AIR. Undigested biomass was analysed as a control. **B)** Average degree of xylan glucuronidation across three replicates of *esk1-5:kak-8* and AtGUX over-expressing *esk1-5:kak-8* plants. Error bars represent standard deviation, **p value ≤0.01; ***p value ≤0.001 when compared with *esk1-5* in Student's t-test.

6.4 Impact of AtGUX1 and AtGUX2 over-expression on interaction of *esk1-5:kak-8* xylan with the cellulose fibril

The presented results indicate that expression of both AtGUX1 and AtGUX2 can substantially increase the degree of *esk1-5* and *esk1-5:kak-8* xylan glucuronidation. Importantly, out of the two AtGUXs, only expression of AtGUX1 can rescue the growth phenotype of *esk1-5* plants (Figure 6.1). It is possible this may be caused by the differences in the specificity of AtGUX enzymes towards generating different patterns of [Me]GlcA branches (Bromley et al., 2013). In order to test this hypothesis, AIR isolated from WT, *esk1-5:kak-8* mutant plants and AtGUX1 and AtGUX2 overexpressing mutants was deacetylated and digested with xylanase GH30. The released oligosaccharides were analysed with PACE (Figure 6.3). As previously described (Chapter 5 and Grantham et al., 2017), the reduction of xylan acetylation in the *esk1-5:kak-8* background led to an alteration of [Me]GlcA pattern when compared to WT material. Most notably, the regular even spacing of [Me]GlcA branches observed in WT plants was absent in *esk1-5:kak-8* where the pattern was more dense and the preference for even spacing was lost. Interestingly, expression of AtGUX1 led to a change in *esk1-5:kak-8* [Me]GlcA patterning. Specifically, the proportion of long distance spacing between the acidic branches appeared to be decreased and the X6 became the dominant distance between the decorations. In contrast, expression of AtGUX2 appeared to have little impact on *esk1-5:kak-8* [Me]GlcA patterning, with the only noticeable difference being a larger proportion of some longer (DP>12) fragments being released from these lines. Interestingly, AIR isolated form *esk1-5:kak-8* and AtGUX1 and AtGUX2 over-expressing double mutant plants released an oligosaccharide with a migration profile similar to the putative U_2X_5 doubly glucuronidated structure which was produced in high amounts by hydrolysis of Pt2*esk-kak* AIR (Figure 5.7). This may indicate some consecutive glucuronidation is present on xylan of these plants or may be a result of GH30 side activity.

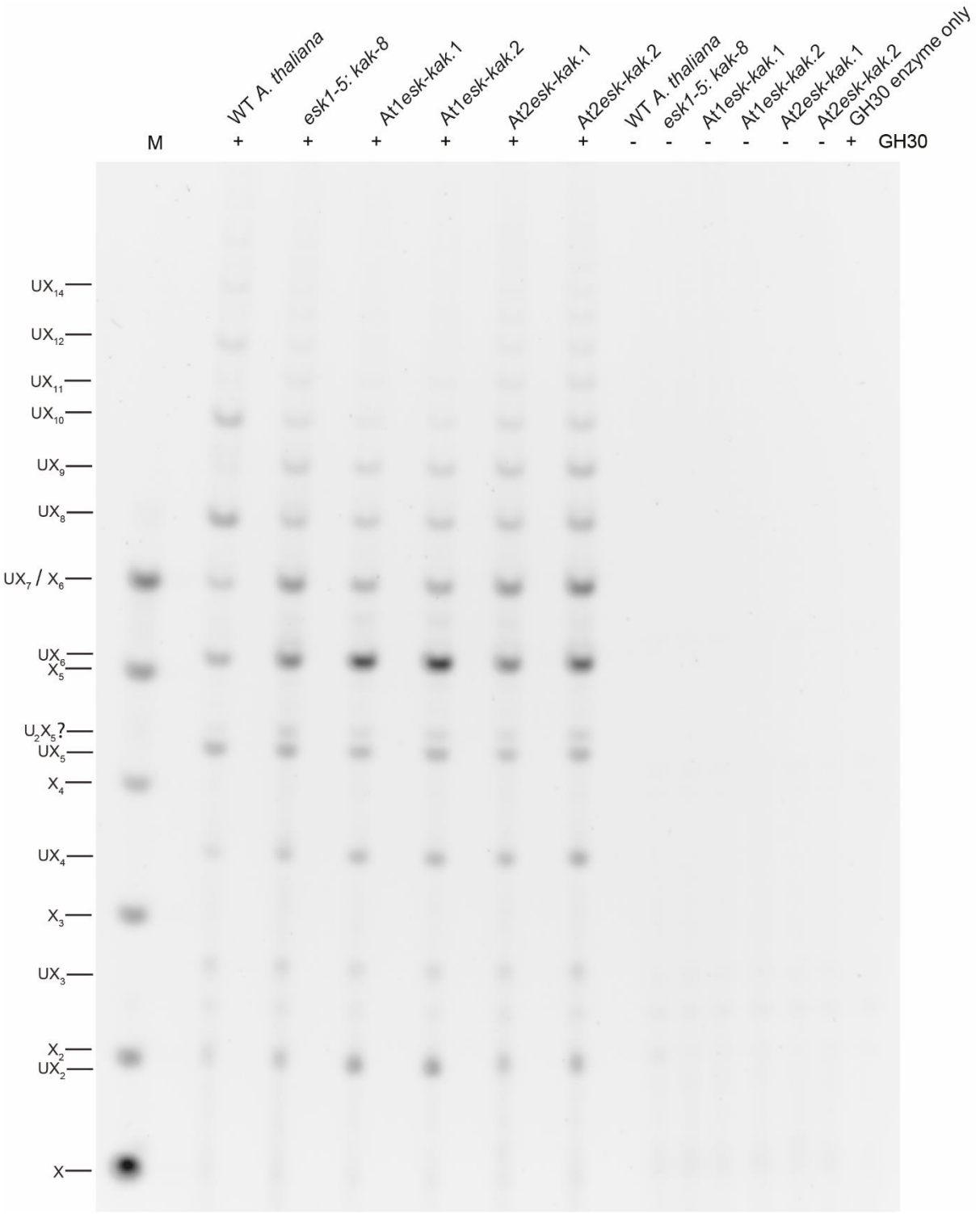


Figure 6.3 [Me]GlcA distribution on xylan in WT, *esk1-5:kak-8* and AtGUX over-expressing Arabidopsis. AIR of all plants was hydrolyzed with glucuronoxylanase GH30 and analysed by PACE. Undigested material was used as a control. Putative U_2X_6 oligosaccharide is marked with a question mark.

To further evaluate changes in *esk1-5:kak-8* [Me]GlcA patterning associated with AtGUX1 and AtGUX2 overexpression, the GH30 digestion and PACE analysis were repeated for three biological replicates of the material analysed in Figure 6.3. The sum of intensities of bands corresponding to [Me]UX_5 to [Me]UX_{12} oligosaccharides was quantified and the intensity of each band was expressed as a percentage of this total (Figure 6.4). In WT, spacing of 8 was the most frequent, as previously observed (Bromley et al., 2013). In line with the PACE analysis and Grantham et al. 2017, the *esk1-5:kak-8* pattern lacked the preference for even spacing of [Me]GlcA decorations observed in WT. For *esk1-5:kak-8* plants over-expressing AtGUX1 a clear preference for spacing of six xylose units was observed, with nearly 60% of all [Me]GlcA branches being placed with this pattern. This was associated with a decrease in other spacing. Notably, a spacing of 7 xylose units decreased by ~50% in these lines. For plants in which the AtGUX2 enzyme was over-expressed in the *esk1-5:kak-8* background, no significant change in the pattern of GlcA spacings was detected.

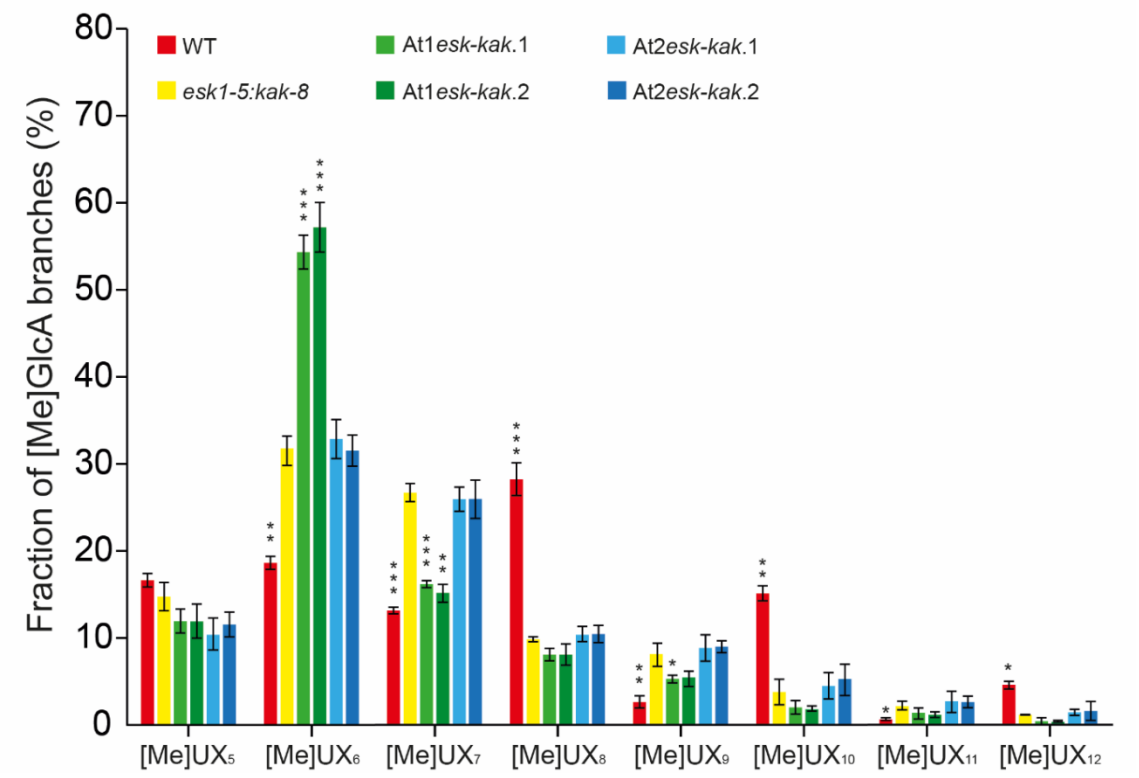


Figure 6.4 [Me]GlcA pattern quantitation on xylan in WT, *esk1-5:kak-8,* AtGUX over-expressing *A. thaliana*. Error bars show standard deviation of three replicates. Asterisk indicate significant difference to *esk1-5:kak-8* in for an oligosaccharide. *p value ≤0.05, **p value ≤0.01, ***p value ≤0.001 in Tukey test after ANOVA.

The results of this analysis complements data on PgGUX1 and PtGUX2 over-expressing *esk1-5:kak-8* plants described in Chapter 5 of this thesis. Together, this data indicates that the over-expression of either AtGUX1 or PgGUX1 alters the structure of *esk1-5:kak-8* xylan molecules to generate more compatible [Me]GlcA patterning. In the case of AtGUX2 and PtGUX2 over-expression, the additional [Me]GlcA added onto xylan is likely to maintain any lack of xylan compatibility. When analysed together with growth data presented on Figure 6.1, this may suggest that only generation of compatible xylan molecules restores the growth of *esk1-5* plants. Therefore, this may indicate a possible role for patterned xylan regions interacting with the hydrophilic surface of the cellulose fibril.

6.5 Changes in *esk1-5:kak-8* xylan patterning may be restoring its interaction with the cellulose fibril

As outlined in section 6.1 the structure of plant cell walls can be studied using ssNMR. Importantly, xylan in CP-INADEQUATE ssNMR experiments exhibits two distinct peak positions associated with backbone in two-fold (2f) and three-fold (3f) screw conformation (Simmons et al., 2016). The presence of secondary cell wall cellulose is necessary for the 2f xylan peak to be detected (Simmons et al., 2016). Thus, the 82.2ppm 2f xylan carbon 4 ($Xn4^{2f}$) signal is likely to be a hallmark for cellulose bound two-fold screw xylan. Interestingly, the change in xylan structure observed in *esk1-5:kak-8* plants leads to the loss of the $Xn4^{2f}$ peak and an increase in the 77.4 ppm 3f signal strength. Therefore, it is clear that much of the xylan from *esk1-5:kak-8* does not interact with the cellulose fibril and any that does may be doing so in a manner different to the one observed in WT plants (Grantham et al., 2017).

To analyse the ratio between the 2f and 3f xylan in *esk1-5:kak-8* plants overexpressing AtGUX1 and AtGUX2 enzymes two replicates of ^{13}C-labelled biomass were generated. CP-INADEQUATE experiments were performed for both replicates of biomass from AtGUX1 (Figure 6.5) and AtGUX2 (Figure 6.6) over-expressing *esk1-5:kak-8* plants by Prof. Ray Dupree at the University of Warwick. Assignment and data analysis were performed in collaboration with Prof. Dupree and Mr Oliver Terrett.

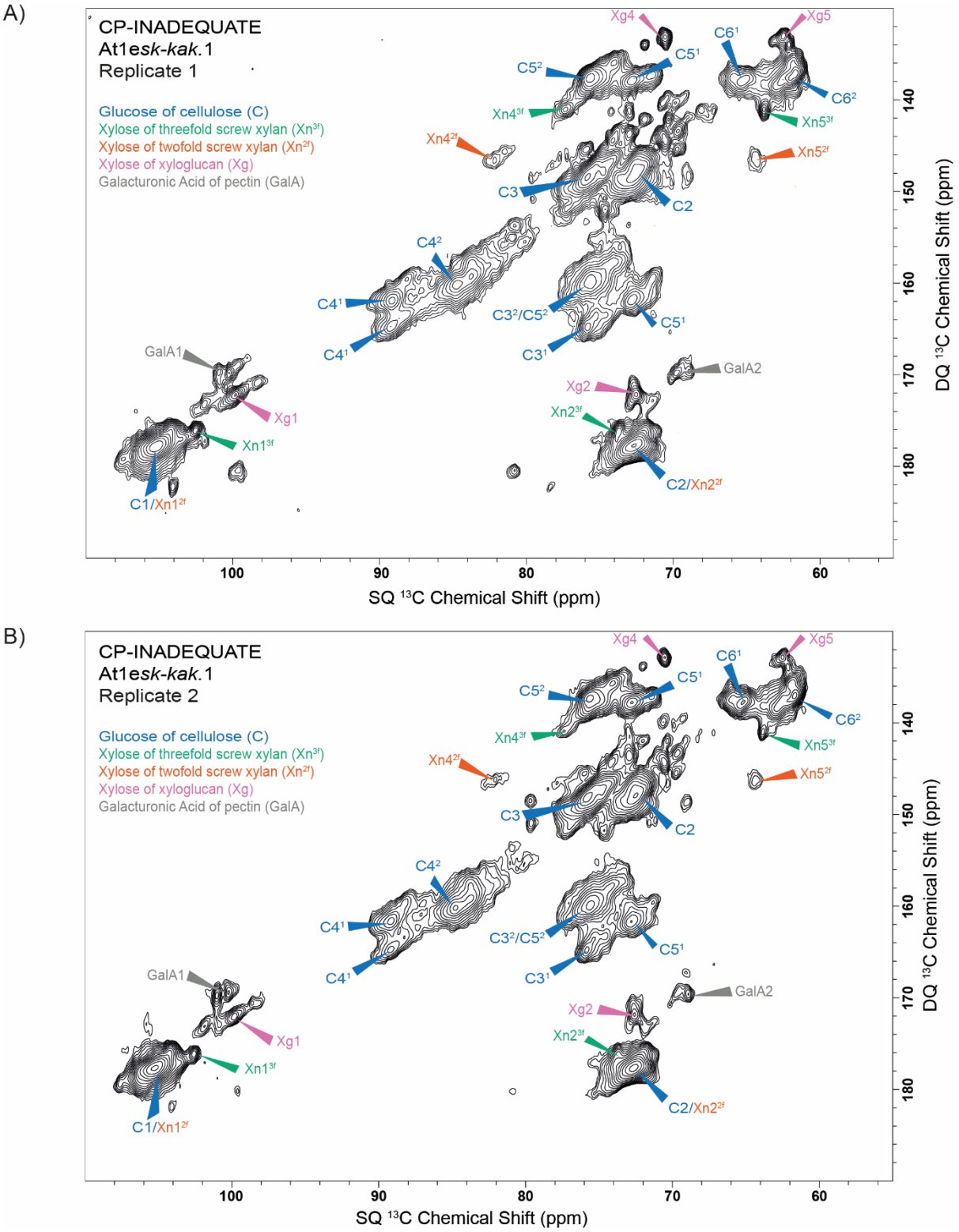


Figure 6.5 CP-INADEQUATE spectra of AtGUX1 over-expressing *esk1-5:kak-8* plants. A) Replicate 1 of the CP-INADEQUATE analysis B) Replicate 2 of CP-INADEQUATE analysis. Colour-coded legend is provided for cellulose, xylan, xyloglucan and galacturonic acid carbons. Two cellulose environment are annotated with a number in superscript.

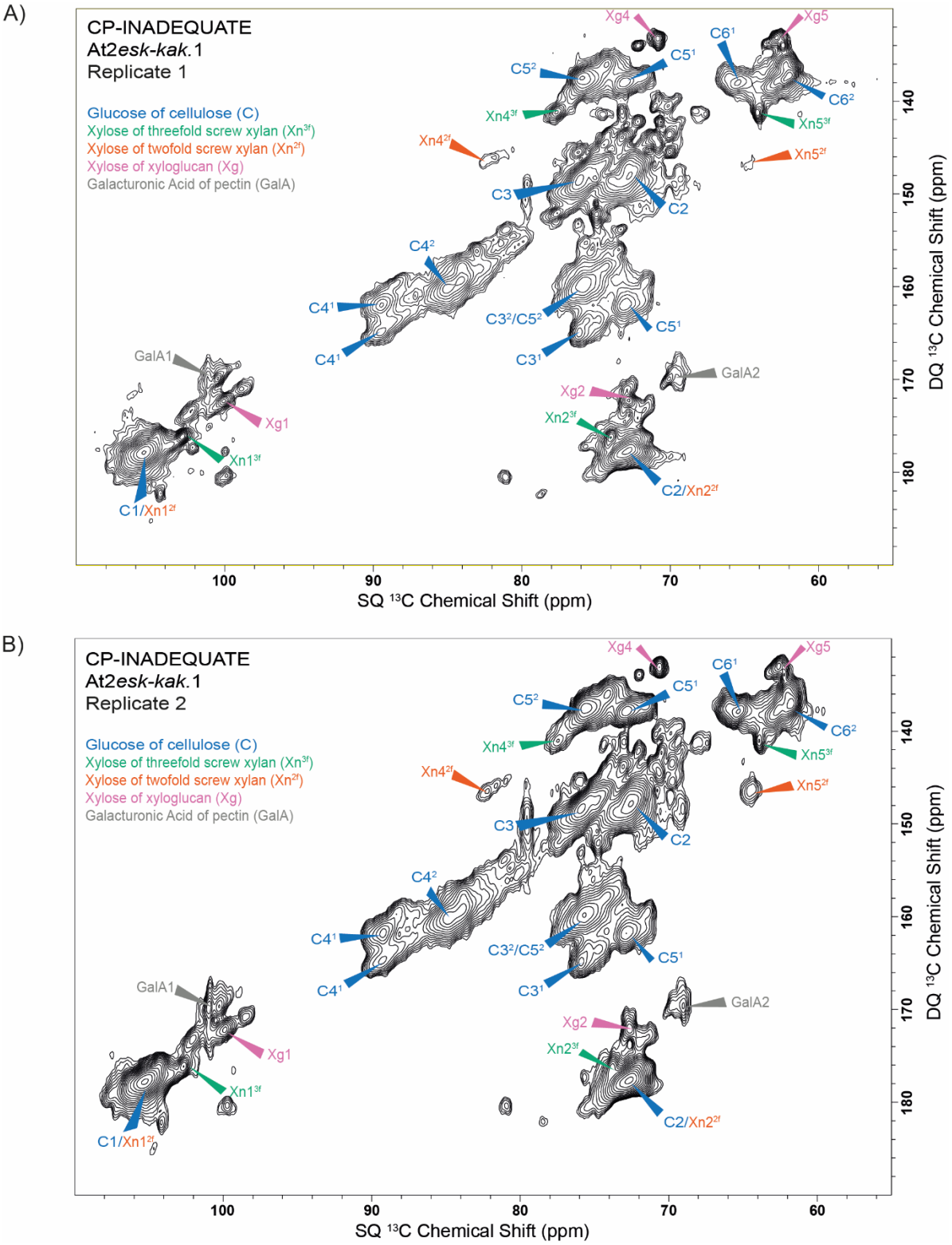


Figure 6.5 CP-INADEQUATE spectra of AtGUX2 over-expressing *esk1-5:kak-8* plants. A) Replicate 1 of the CP-INADEQUATE analysis B) Replicate 2 of CP-INADEQUATE analysis. Colour-coded legend is provided for cellulose, xylan, xyloglucan and galacturonic acid carbons. Two cellulose environment are annotated with a number in superscript.

To start with, material from all analysed plants produced spectra similar to these previously reported in published datasets (Dupree et al., 2015, Simmons et al., 2016, Grantham et al., 2017). This allowed annotation of different environments detected for cellulose glucose and xylan xylose. In addition to previous publications, the assignment of the spectra was guided by the rule in which the two carbons are covalently linked if the sum of their single quantum (SQ) chemical shift values matches their double quantum (DQ) chemical shift. For example, carbon 1 of cellulose has a SQ chemical shift of 105.2ppm and it is covalently linked to glucose carbon two which has a shift of 72.3ppm. Since their DQ chemical shift position was detected to be equal to 177.5ppm, it can be deduced that they are covalently linked. To provide a better overview of the samples, the galacturonic acid (GalA) and Xyloglucan xylose (Xg) signals were also annotated following the same set of rules. Signals from these polysaccharides are associated with primary cell wall and thus their intensity can be used to estimate its abundance and standardise the analysis. To achieve this, the signals for Xg carbon 4 and 5 were matched to have a similar intensity across the spectra by adjustment of contour levels.

In all four CP-INAD spectra, the strongest signals were associated with the most abundant cell wall polysaccharide – cellulose. In line with previous reports, the signals for cellulose glucose form two distinct sub-groups known and annotated as Domain 1 and Domain 2. These are likely to arise from differences in the C6 hydroxymethyl conformation linked to the presence of glucose either on the surface or within the cellulose crystal (Phyo et al., 2018, Simmons et al., 2016). As annotated on the spectra, the differences in the chemical shift of glucose carbons associated with the two domains was most clear for carbons 4, 5 and 6 of the monosaccharide. This ability to annotate the distinct cellulose domains indicates that the spectra are of quality good enough to distinguish local cellulose environments. Therefore, they are also likely to provide accurate data on the conformation of xylan xylose.

In order to evaluate these xylose environments, the spectra were studied for the presence of peaks with chemical shifts associated with carbon 4 and 5 of xylose in 2f and 3f conformation. In contrast with *esk1-5:kak-8* analysis, in which no 2f xylose carbon 4 signal at 82.2ppm and no paired xylose carbon 5 signal at 64.3ppm were detected (Grantham et al., 2017), all four spectra produced peaks with chemical shifts characteristic of both 2f and 3f xylose conformations. This indicates that over-

expression of both AtGUX1 and AtGUX2 leads to formation of larger amounts of 2f xylan than detected in the *esk1-5:kak-8* transformation background alone.

To further analyse the effect of overexpressing the AtGUX1 and AtGUX2 enzymes in the *esk1-5:kak-8* background the volume of peaks corresponding to 2f and 3f xylan carbons 4 and 5 was integrated across the four spectra. The intensity of each species was expressed as a proportion of the total xylan signal for a given carbon 4 or carbon 5 (Table 6.2). The data for both replicates of AtGUX1 and AtGUX2 over-expressing *esk1-5:kak-8* plants was analysed alongside intensities observed for WT *A. thaliana* material which were quantified using previously published spectra (Grantham et al., 2017). The intensity of signals was compared to these observed for WT material only and not for WT and *esk1* as on the spectrum from plants with *ESK1* gene mutated all detectable xylan signal can be attributed to the 3f conformation. In the WT material a great majority of xylan signal for both carbon 4 and 5 is coming from the 2f conformation of the polysaccharide. In contrast to that, in none of the four samples analysed as a part of this thesis the 2f xylan signal formed more than 50% of the total xylan intensity. However, when considering individual biological replicates clear differences between *esk1-5:kak-8* plants overexpressing AtGUX1 and these over-expressing AtGUX2 can be observed. Specifically, for biological replicate number one, expression of AtGUX1 enabled generation of more 2f xylan for both studied carbons than did the over-expression of AtGUX2. This indicates that, in this set of samples, expression of AtGUX1 enabled a larger proportion of xylan to form a fold associated with WT-like xylan-cellulose complex formation. For biological replicate two the data was somewhat different. Over-expression of both, AtGUX1 and AtGUX2 enabled generation of similar amounts of 2f xylan in the transgenic plants. This indicates possible inconsistency in sample quality or other experimental problems. For example, the growth of plants was not equally robust in both experiments- the At2*esk-kak*.1 plants might have been stronger in the 2nd replicate.

Table 6.2 Percentage of xylan xylose carbon four (Xn4) and carbon five (Xn5) signal intensity contributed by two-fold (2f) and three-fold (3f) signal in WT, AtGUX1 and AtGUX2 over-expressing *esk1-5:kak-8 A. thaliana.* In *esk1-5:kak-8* spectrum all the xylan signal on the CP-INADEQUATE spectrum can be attributed to the tree-fold screw backbone (Grantham et al., 2017).

Plant genotype	Xn4 2f	Xn4 3f	Xn5 2f	Xn5 3f
WT *A. thaliana*	70.2	29.8	77.1	22.9
At1*esk-kak*.1 replicate 1	42.1	57.9	39.4	60.6
At2*esk-kak*.1 replicate 1	32.8	67.2	20.4	79.6
At1*esk-kak*.1 replicate 2	47.3	52.7	46.5	53.5
At2*esk-kak*.1 replicate 2	39.6	60.4	46.3	53.7

Taken together, the results of this solid state NMR analysis provide some indication that expression of AtGUX1 enzyme may have a different effect on the fold of the *esk1-5:kak-8* xylan than the over-expression of AtGUX2. In the case of AtGUX1 over-expression a consistent restoration of some 2f xylan formation was observed across the two replicates. On the other hand, this was the case only for one replicate of *esk1-5:kak-8* material over-expressing AtGUX2. This variation in experimental results may indicate that factors other than xylan-cellulose interaction, such as the variation in plant growth, may influence the 2f to 3f xylan ratio. Alternatively, the observed result may indicate that formation of more compatible xylan molecules, associated with AtGUX1 overexpression in the acetylation mutant, could enable a more robust rescue of xylan-cellulose complex formation. Therefore, differences between the two over-expressing lines, such as those in stem height (Figure 6.1), may be somewhat linked to xylan-cellulose interaction and therefore they may be indicative of its biological role.

6.6 Complementation of *esk1-5* growth phenotype is correlated with the rescue of xylem vessel morphology

Reduction in xylan acetylation in *esk1* plants leads to vessel collapse (Lefebvre et al., 2011, Xiong et al., 2013, Yuan et al., 2013). It is possible that this collapse may be due to loss of xylan cellulose interaction in this biomass which may be associated with a reduction in mechanical strength of the vessels. Therefore, if the changes in *esk1* [Me]GlcA patterning associated with AtGUX1 over-expression enable formation of compatible xylan molecules and restore some xylan-cellulose interaction they may have a positive effect on vessel integrity. This might correlate with rescue of the *esk1* growth phenotype observed in these AtGUX1 over-expressing plants. The same might not be the case for *esk1* plants over-expressing AtGUX2, where the pattern of GlcA is not compatible and plant growth is not restored.

To evaluate the effect of altering the pattern of GlcA decorations on *esk1* plants by overexpressing AtGUX1 and AtGUX2, the vessel morphology of WT, mutant and transgenic mutant plants was evaluated. The integrity of vessels was studied using low temperature scanning electron microscopy (cryo-SEM). The technique requires little sample preparation and allows the maintenance of its hydration state. For the analysis, bottom 1 cm of stems of 7 week old *A. thaliana* plants was collected and placed in a cryo-SEM specimen stub. Liquid nitrogen frozen and sectioned WT stems were platinum coated and visualised with cryo-SEM (Figure 6.7A). Distinct vessel bundles were observed in each stem studied and use of increasing magnification allowed close examination of vessel morphology and possible collapse. Plants subjected to the cryo-SEM analysis were chosen at random from a pool of 60 plants of any given genotype. For the analysis of vessel collapse all visible vessel bundles and vessels were considered for sectioned stems of any given genotype.

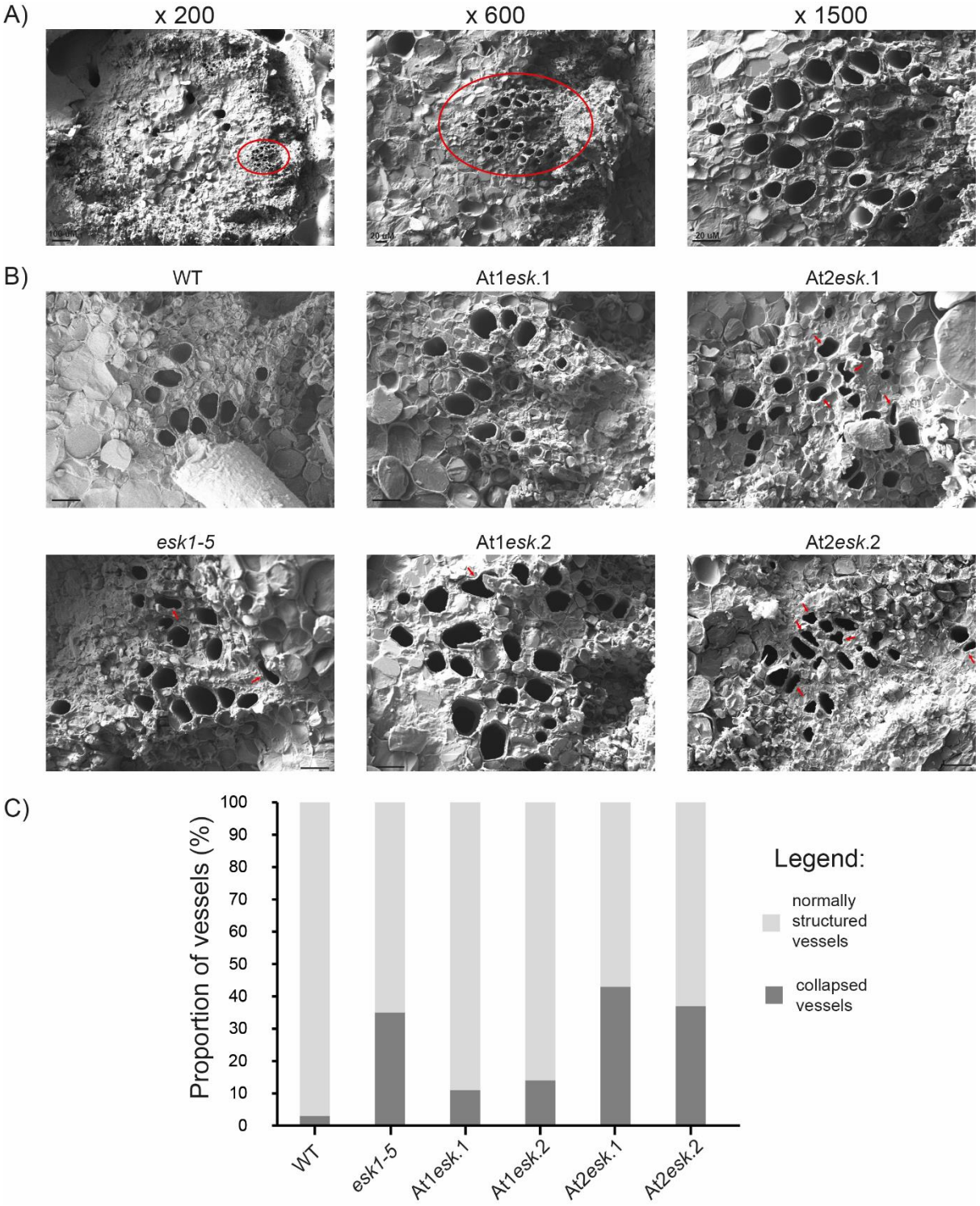


Figure 6.7 Vessel morphology in WT, *esk1-5* , AtGUX1 and AtGUX2 over-expressing *esk1-5 A. thaliana* A) Use of increasing magnification allows visualisation of vessel bundles (marked with a red circle) in WT stems. B) Representative images of vessels in WT, *esk1-5* and AtGUX1 and AtGUX2 over-expressing *esk1-5 A. thaliana*. Indented vessels, classified as collapsed, are marked with a red arrow. Size bar = 20 μm C) Proportion of normally structured and collapsed vessels across analysed genotypes. Data summarises analysis of between 130 and 150 vessels across 5 individual plants per genotype.

To study vessel morphology further, 5 individual plants were analysed for WT, *esk1-5* and two lines for each AtGUX1 and AtGUX2 over-expressing mutant *A. thaliana*. For each genotype, stem bundles were examined for the presence of collapsed vessels (Figure 6.7B). These were defined as vessels in which a clear indentation of the cell wall can be observed. Examples of such structures are marked on Figure 6.7B. To quantify the degree of vessel collapse between 130 and 150 vessels were examined across the studied genotypes. Each vessel was classified as either collapsed or normally structured and the proportion of each vessel type was quantified for different plant lines (Figure 6.7C). In WT plants only a small proportion of collapsed vessels was observed (<5%) while for the *esk1-5* a third of analysed vessels had cell wall indentations. This value was decreased for both lines of AtGUX1 over-expressing mutant and remained high for AtGUX2 over-expression.

This analysis indicates that the expression of AtGUX1 in the *esk1-5* plants enables rescue of vessel morphology while the over-expression of AtGUX2 does not. It is somewhat challenging to untangle cause from an effect in the rescue of this phenotype. It is possible that the improved growth of *At1esk* plants allows for formation of a larger number of more robust vessels. Alternatively, it is possible that this restoration of vessel morphology may be a direct reason for growth restoration observed in the At1*esk* plants (Figure 6.1).

6.7 Xylanocellulose fibrils are present in the cell walls of xylem vessels

Until now, the data presented has focused on the analysis of AtGUX1 and AtGUX2 overexpressing *esk1* plants. The current analysis demonstrates that AtGUX1 expression alters the pattern of *esk1* [Me]GlcA decorations, restores its vessel morphology and rescues the *esk1* dwarfing phenotype. These changes are not observed for the plants over-expressing AtGUX2. The NMR analysis indicated that AtGUX1 over-expressing *esk1-5:kak-8* plants may have a larger proportion of xylan bound to the cellulose fibril than these over-expressing AtGUX2. Together, this may indicate that the xylan-cellulose interaction may facilitate the maintenance of vessel strength and enable plant stem growth. However, the exact nanoscale structures in which this strength determining interaction may occur remain largely unknown.

To evaluate the nanoscale architecture of plant cell walls and identify possible structures containing these xylan-cellulose complexes high magnification cryo-SEM

was used to analyse WT *A. thaliana* cell walls (Figure 6.8). Vessels remained the main focus of the analysis as their strength and integrity may be influenced by the interaction between xylan and cellulose.

The initial analysis investigated the structure of WT vessels using up to x10000 magnification (Figure 6.8A). Again, sets of vessel bundles were detected and, using higher magnification, fibrous structures with bright ends were observed in the fractured cell wall material surrounding hollow vessels. These might arise from cellulose fibrils in the cell wall. To investigate the nature of these fibrils further, a comparative analysis was performed using magnification exceeding x30000 and analysing vessel cell walls and a commercially available fibrous cellulose standard extracted from cotton linters composed from 99% pure cellulose (Sczostak, 2009) (Figure 6.8B). In this experiment, clear individual fibres with distinct bright termini were observed in both samples indicating that the vessel wall fibrils have a similar appearance and dimensions to the cellulose fibrils present in the polysaccharide standard. To evaluate the composition of these fibres further, the morphology of WT *A. thaliana* vessel cell walls was compared to that of the *irx3* mutant. IRX3 is a member of the cellulose synthase complex and plants without its activity lack secondary cell wall cellulose (Ha et al., 2002). This is associated with extensive vessel collapse and dwarfing. Interestingly, when compared to the WT material, the *irx3* stems lacked fibrous structures in their vessel cell walls. Unlike in the WT material, the *irx3* cell walls appeared to be composed from a largely amorphous matrix (Figure 6.8C). Together with the analysis of the fibrous cellulose standard this data indicates strongly that fibres present in vessel secondary cell walls are likely to be composed, at least partially, from cellulose. To evaluate these fibres further, their diameter and that of fibres present in the cellulose standard was measured (Figure 6.8D). The measured fibril width showed high variation which may be associated with biological variation in the system or may be an artefact caused by difficulties in defining extremes of measured fibrils in the ImageJ software. Interestingly, the *A. thaliana* fibres appeared to be significantly wider than those detected in the polysaccharide standard. This may indicate different arrangement or number of glucan chains within the fibril or may be an indication of presence of other components within the structure studied in the Arabidopsis vessels.

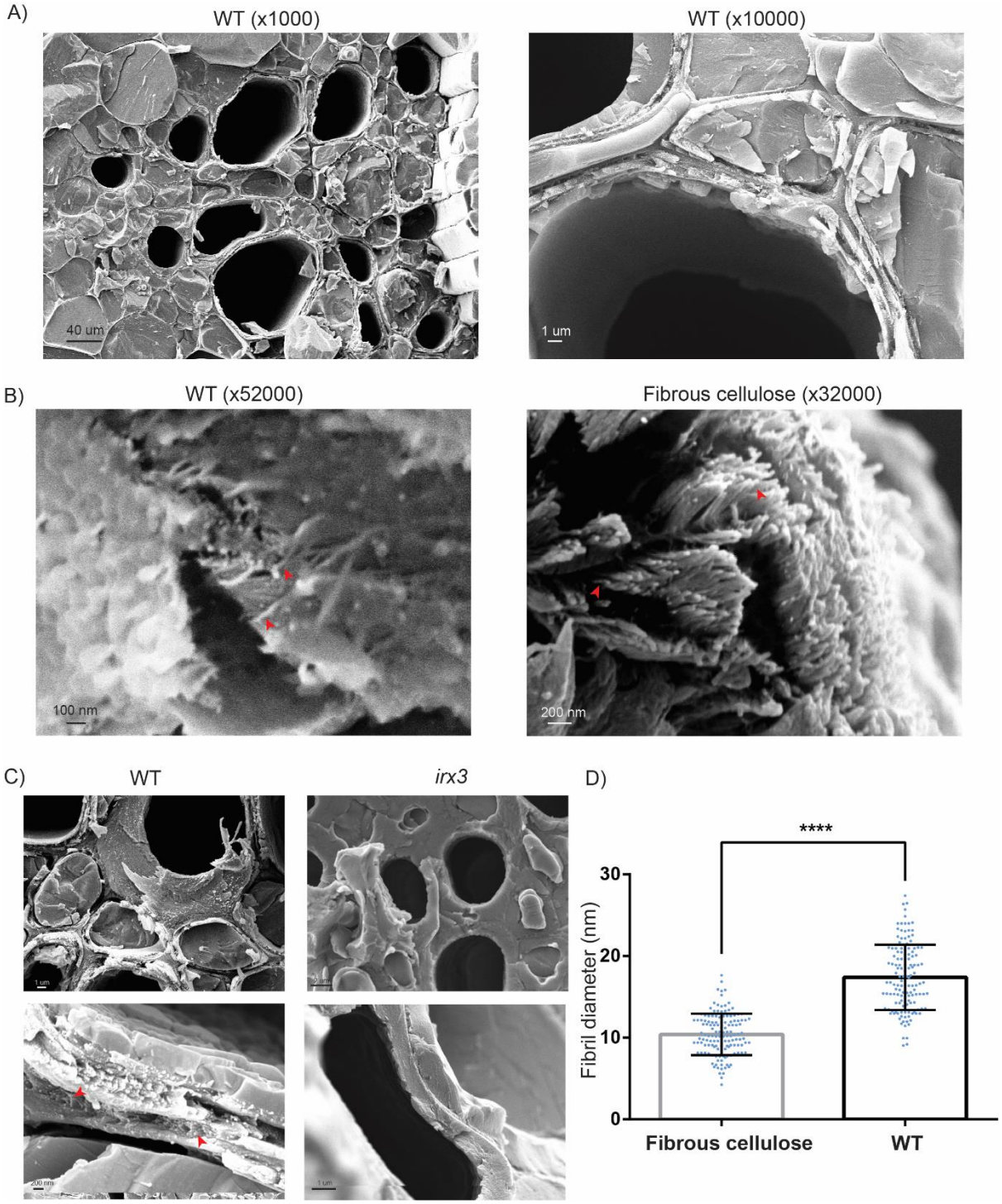


Figure 6.8 Identification of cell wall macrofibrils in *A. thaliana* vessel cell walls. A) Imaging of WT vessels at increasing magnification reveals presence of fibrillar structures in their cell walls. B) Cell wall fibrils have an appearance similar to fibrous cellulose standard C) Macrofibrils are not detectable in *irx3 A. thaliana* lacking secondary cell wall cellulose. D) Quantitation of macrofibril diameter in the cellulose standard and *in plantae.* N = 150 fibrils. Error bars represent standard deviation. **** denotes $p \leq 0.0001$ in Student's t-test. Red arrows mark macrofibrils.

To analyse the composition of the cell wall fibrils the diameter of WT structures was compared to these formed in *irx9 A. thaliana* plants (Figure 6.9A). IRX9 is a member of secondary cell wall xylan synthase complex (Zeng et al., 2016). Therefore, *irx9* plants have impaired xylan synthesis and as a result of that the content of this hemicellulose is decreased by ~ 50% (Brown et al., 2007). A comparative analysis of the diameter of WT and *irx9* cell wall fibres has indicated that they are ~30% thinner in the xylan mutant plants (Figure 6.9B). Interestingly, the diameter of *irx9* fibrils was not different to those observed in the fibrous cellulose standard from cotton linters. Again, high variation was observed in the measurement of the *irx9* fibril width. To overcome this, further experiments may need improved resolution which could be gained by decreasing the width of the coating or changing the metal used in this process. Despite the technical challenges, this analysis indicates that vessel cell walls contain fibrous structures composed from both cellulose and xylan. These structures have a diameter larger than a single cellulose microfibril and therefore may be described as macrofibrils. These macrofibrils may be the xylanocellulose complexes formed by the interaction between xylan and the hydrophilic surface of the cellulose macrofibril.

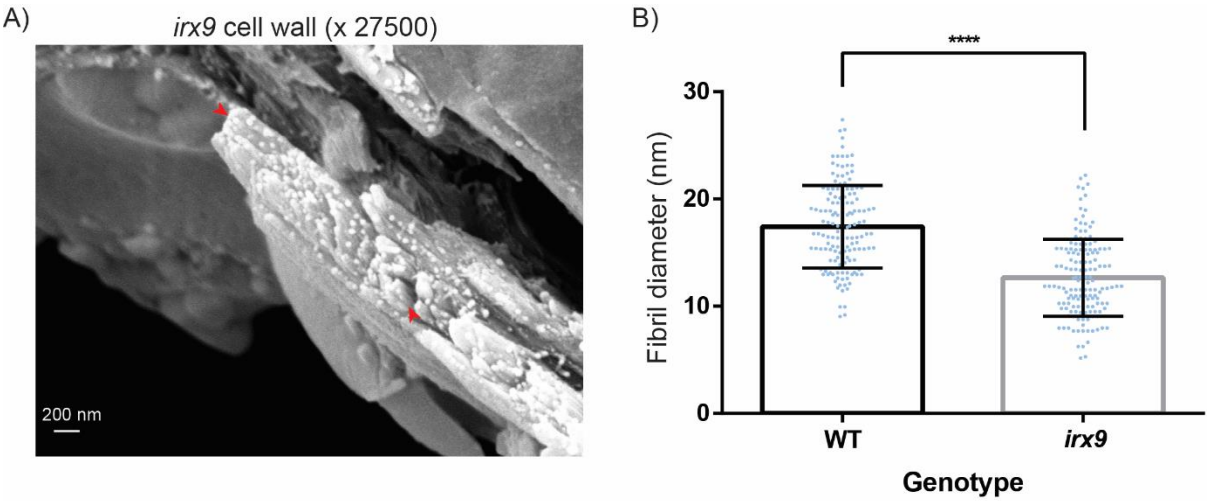


Figure 6.9 Cell wall macrofibrils contain xylan. A) Representative image of *irx9* macrofibrils **B)** Quantitation of macrofibril diameter in WT and *irx9* plants. N = 150. Error bars represent standard deviation. **** denotes p ≤ 0.0001 in Student's t- test. Red arrows mark macrofibrils.

6.8 Softwood tracheids contain fibrils similar to the structures observed in Arabidopsis cell walls

A significant part of this thesis was devoted to the analysis of softwood xylan biosynthesis. The main focus of this work was to identify conifer GUX enzymes responsible for the synthesis of compatible and incompatible xylan regions. Data presented in this chapter indicate that the xylan-cellulose interaction may be relevant to the strength of the cell wall. Thus, it may in turn influence properties of plant material including wood. Taking this into consideration it is relevant to investigate the nanoscale architecture of softwood cell walls. To evaluate the nature of softwood cell walls stem, sections of Norway spruce (*Picea abies*) were analysed with cryo-SEM.

Softwood has a morphology different to hardwood species. In contrast to the previously studied *A. thaliana* samples, the softwood water conducting tissue, xylem tracheids, were highly abundant and formed a solid cylinder over the majority of the stem cross section (Figures 6.10A and B). At higher magnification, the cell walls of tracheids did contain fibrous structures similar to those detected in *A. thaliana* (Figure 6.10C and D). Interestingly the softwood fibrils frequently assembled into larger cylindrical aggregates, a feature that was not observed in *A. thaliana* stems (Figure 6.10D and E). While the softwood fibrils did form aggregates, it was still possible to resolve individual units and determine their diameter using a magnification of x50000. The analysis of softwood fibril diameter indicated that they are on average nearly twice as wide as these detected in *A. thaliana* stems (Figure 6.10F).

This analysis indicates that structures morphologically similar to *A. thaliana* xylanocellulose fibrils can be detected in softwood tracheid cell walls. While their exact composition remains unknown, it is possible that, similarly to the *A. thaliana* fibrils, they might be involved in the maintenance of tracheid cell wall strength.

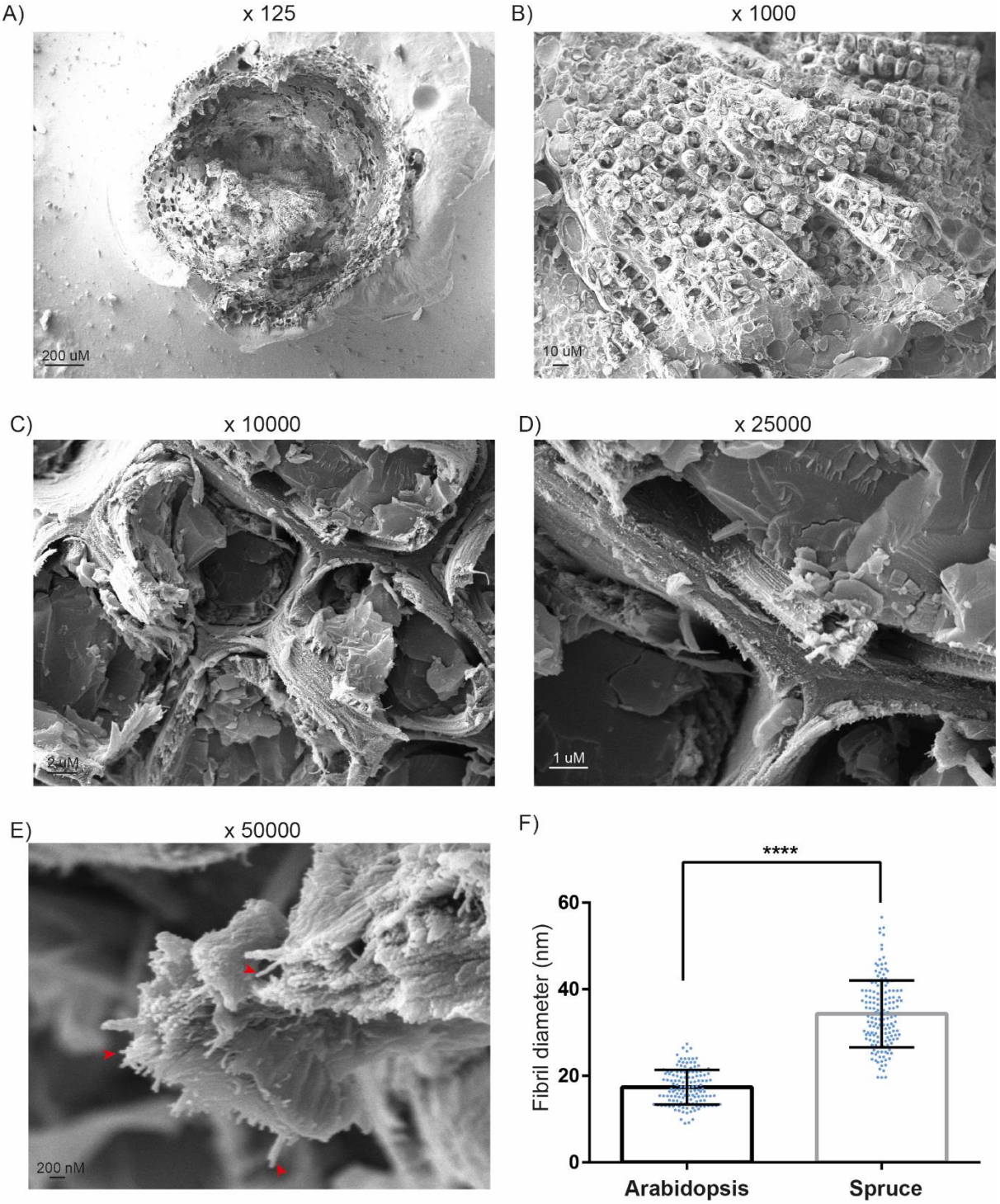


Figure 6.10 cryo-SEM analysis of *P. abies* stem sections. A) to E) Representative images of stem sections of one year old Spruce branch at different magnifications. F) Diameter of Spruce cell wall fibrils compared to these observed in WT *A. thaliana* material (same data as presented on Figure 6.8 and 6.9). For each bar 150 individual fibrils were measured. Error bars represent standard deviation. **** denotes $p \leq 0.0001$ in Student's t-test. Red arrows mark macrofibrils.

6.9 Discussion

The primary aim of the work described in this chapter was to use a range of plant phenotyping, carbohydrate analysis, ssNMR and microscopy techniques to evaluate a possible biological role for the xylan-cellulose interaction. Plants with no functional copy of the *ESK1* gene, which causes reduced xylan acetylation and disrupted [Me]GlcA patterning, were used as a model for a system in which this xylan-cellulose interaction is not normal. The results of the analysis indicate that the overexpression of AtGUX1 may restore compatible xylan patterning and is likely to contribute to an increase in the proportion of xylan that is bound to the hydrophilic surface of the cellulose fibril as a two-fold screw. This is correlated with a rescue in *esk1* plant growth and vessel morphology, providing a possible biological role for xylan-cellulose complex formation in the maintenance of wall strength.

6.9.1 Glucuronosyl residues on xylan are unlikely to be functionally equivalent to the acetyl substituents

Suppression of the *esk1* growth phenotype by AtGUX1 overexpression was previously reported in the literature (Xiong et al., 2015). The authors of this work performed experiments in which *esk1* plants were transformed with a construct in which AtGUX1 and two wheat xylan arabinosyl transferases (XATs) were expressed under the control of the *ESK1* promoter. Similarly to the results presented in this thesis, AtGUX1 expression resulted in an increase in the cell wall [Me]GlcA content and a full rescue of the *esk1* growth phenotype. No change in the plant phenotype was observed for *esk1* plants transformed with constructs to express XAT2 and XAT3. However, in the case of these transformants no increase in cell wall arabinose was detected over the *esk1* levels, suggesting possible issues with XAT expression or activity, which were not detected. The results led Xiong et al., to suggest that glucuronosyl and acetyl residues on xylan are functionally equivalent. It was proposed the rescue of the *esk1* phenotype may result from an increase in the total number of xylan branches in the plants overespressing AtGUX1 which may improve solubility of the xylan molecule.

In the introduction to their report Xiong et al., describe the xylan-cellulose interaction as an important factor for the maintenance of the plant cell wall network. At the time of publication, the authors were not aware of the loss of this xylan-cellulose interaction in *esk1* plants which was described by Grantham et al., only last year. Interestingly,

unlike Grantham et al. and the results of experiments presented in this thesis, Xiong et al., did not detect any significant increase in the degree of xylan glucuronidation in *esk1* plants over the WT biomass.

The results presented in this chapter indicate that the hypothesis suggesting functional equivalence of glucuronosyl and acetyl residues, proposed by Xiong et al., is unlikely to be fully correct. Expression of both conifer and angiosperm GUX1 and GUX2 enzymes led to an increase in the degree of *esk1* xylan glucuronidation (Figure 6.1). However, this increase in xylan glucuronidation did not always result in the rescue of the mutant plant growth phenotype, suggesting that only specific types of glucuronidation may restore xylan function. Indeed, the results of subsequent analyses indicate that expression of AtGUX1 has a different effect on the pattern of *esk1* [Me]GlcA decorations than the over-expression of AtGUX2. While AtGUX1 over-expression changed the pattern to become more compatible, the expression of AtGUX2 did not produce any clear changes in GlcA patterning. This observation has to be correlated with the results of PgGUX1 and PtGUX2 expression, where PgGUX1 increased the proportion of compatible xylan and suppressed the dwarfing, but PtGUX2 introduced consecutive patterning, which may result in the formation of xylan structures which cannot interact with cellulose, and did not suppress the dwarfing. Therefore these results may suggest that the formation of compatible xylan molecules may be a pre-requisite to restore *esk1* xylan function and plant growth.

6.9.2 Solid state NMR analysis does not provide conclusive information on the rescue of xylan-cellulose interaction in AtGUX over-expressing *esk1* plants

To better understand the possible reason for the compatible xylan requirement for *esk1* growth restoration, the interaction between xylan and cellulose was studied in *esk1-5:kak-8* plants over-expressing AtGUX1 and AtGUX2. Two replicates of the biomass were studied using CP- refocussed INADEQUATE ssNMR experiments which can be used to visualise different conformations of covalently linked carbons in polysaccharides. This has previously allowed the annotation of signals associated with twofold xylan molecules, which are likely to be formed by interaction of the hemicellulose with the hydrophilic surface of the cellulose microfibril (Simmons et al., 2016, Grantham et al., 2017).

The solid state NMR analysis provided inconsistent results across the two biological replicates of biomass. For replicate 1 over-expression of AtGUX1 in *esk1* resulted in formation of more two-fold xylan than AtGUX2 over-expression. The same was not the case for the second replicate of biomass studied. In the case of this second replicate, the *esk1* plants over-expressing AtGUX1 had two-fold xylan levels only marginally higher than these observed in *esk1* plants over-expressing AtGUX2. A third replicate of ^{13}C enriched biomass for both AtGUX1 and AtGUX2 over-expressing *esk1-5:kak-8* plants may need to be analysed with solid state NMR to resolve this issue.

The reasons for the differences between the results obtained for the two replicates is challenging to explain. Excluding simple mistakes in sample preparation, one possible explanation can be found in the exact solid state NMR experiment used. CP-refocussed INADEQUATE (CP-INAD) experiments are not strictly quantitative and this is primarily caused by the use of cross-polarisation to transfer magnetisation onto studied carbons (Simmons et al., 2016). This process happens more efficiently for less mobile molecules. Thus, CP spectra may accentuate any solid components of biomass and they may under-represent more mobile components of the cell wall. It is unclear how these non-mobile two-fold xylan molecules are formed in the cell walls. It is possible that the differences observed between the spectra might be associated with different growth of plants and that differences in maturity may therefore influence the amount of two-fold xylan. Alternatively, the two-fold xylan may form on different surfaces of the cellulose fibril in At1*esk* and At2*esk* plants (as seen on Figure 1.13). Therefore, the two-fold xylan observed in *esk1-5:kak-8* plants expressing both AtGUX1 and AtGUX2 may be not only bound to the hydrophilic surface of the cellulose fibril but it could also be binding onto the hydrophobic surface of the microfibril. This hydrophobic surface interaction may also involve two-fold xylan formation and could, due to the nature of the CP-INAD experiment, be causing strong two-fold xylan signals to be generated in both genotypes studied. The two-fold signals in the case of At1*esk* and At2*esk* would however be resulting from different binding modes. To resolve these issues, further experiments, such as CP-PDSD, will be required to investigate whether At1*esk-kak* and At2*esk-kak* xylans are in close proximity to different cellulose domains. In addition to that it will be interesting to study Pg1*esk-kak* and Pt2*esk-kak* cell walls with solid state NMR in order to provide additional information about the

importance of xylan patterning for association with different surfaces of the cellulose fibril.

As discussed the formation of two-fold xylan molecules may not require compatible patterns and may occur with any xylan molecule. However, if this is the case, some two-fold xylan should also be detected in *esk1* plants alone. Lack of this signal in the spectra of *esk1kak* plants (Grantham et al., 2017) may indicate possible issues with xylan solubility which prevent its proper deposition or delivery to the cell wall. This may be explored further by performing xylan microscopic localisation experiments in *esk1* plants using available anti-xylan immunoprobes (McCartney et al., 2005).

6.9.3 Cell wall macrofibrils may be formed from xylanocellulose complexes

Results presented in the first part of this Chapter indicate that the interaction between xylan and the hydrophilic surface of the cellulose fibril, which is likely to be enabled by the compatible pattern of GlcA decorations, can be important for vessel cell wall strength and plant growth. To investigate the structures in which this interaction may occur, the results presented in the latter part of the chapter employ cryo-SEM to analyse the nanoscale architecture of vessel secondary cell walls. As indicated in the next paragraph of this discussion, SEM was previously used to study the size of the cell wall macrofibrils (Donaldson, 2007). However, work presented in this chapter is the first to employ low temperature SEM to study plant cell walls. Due to cryopreservation this technique allows for the maintenance of the native hydration and increases the imaging resolution. The results of the cryo-SEM analysis presented in this chapter indicate that *A. thaliana* cell walls contain macrofibrils, which are likely to be composed from both cellulose and xylan. These complexes may contain xylan bound to the hydrophilic surface of the cellulose fibril. Results of this study and previously published work indicate that *esk1* plants, where xylan does not interact with cellulose, do show vessel collapse (Lefebvre et al., 2011, Xiong et al., 2013).

The *esk1* mutation is also associated with lower breaking force of stems when compared with WT (Yuan et al., 2013), suggesting that loss of the xylan-cellulose interaction influences stem mechanical properties. This indicates that xylanocellulose macrofibrils could be participating in the maintenance of vessel cell wall strength and may influence mechanical properties of biomass. Therefore, it would be very interesting to evaluate the strength of At1*esk* and At2*esk* stems and compare it to this of *esk1-5* stems.

The existence of the macrofibrils in *A. thaliana* cell walls is a somewhat surprising discovery of this study. Previous reports do not indicate presence of such structures in *A. thaliana* (Jarvis, 2018) but the macrofibrils were detected in another hardwood, poplar (Donaldson, 2007). Interestingly, Donaldson reported that poplar macrofibrils, depending on their position in cell wall, have an average diameter of between 14 and 18 nm, which is similar to what was measured for *A. thaliana* as a part of the current chapter. In addition to analysing poplar hardwood, Donaldson also investigated the presence and diameter of macrofibrils in softwood. In line with results presented in this chapter, he did observe macrofibrils in cell walls of *Pinus radiata* tracheids. Moreover, they were much larger than those seen in hardwoods, replicating the results presented here in Arabidopsis. However, with an average diameter ranging between 20 and 34 nm, the size of pine macrofibrils was somewhat smaller than what was measured in spruce wood as a part of this chapter. However, these observations are not necessarily inconsistent. Unlike in the case of experiments performed as a part of this thesis, Donaldson dehydrated wood samples prior to SEM imaging. As the spacing between softwood cellulose microfibrils is sensitive to hydration (Fernandes et al., 2011), at least part of the difference in macrofibril diameter might be due to changes in the water content within the structure.

The spatial distribution of hemicelluloses in softwood secondary cell walls remains largely unknown. In addition to xylan structures compatible (Busse-Wicher et al., 2016b) and incompatible (Martinez-Abad et al., 2017) with binding to the hydrophilic surface of the cellulose fibril, softwood cell walls contain large quantities of acetylated and unacetylated galactoglucomannans (Scheller and Ulvskov, 2010).

A model has been proposed for the localisation of these gymnosperm hemicelluloses in *Gingko biloba* tracheid cell walls (Terashima et al., 2009). The hypothesis presented by Terashima et al., is based on a combination of SEM analysis with density calculations of different cell wall components. This speculative model indicates that both mannan and xylan may be coating cellulose microfibrils. More specifically, the mannan was hypothesised to coat the cellulose fibril directly while xylan would coat this complex and, via the MeGlcA branch, cross-link it to lignin. However, given the structural similarities in compatible xylan molecules between angiosperms and gymnosperms (Busse-Wicher et al., 2016b) which were not known by Terashima et al., it is likely that at least some proportion of conifer xylan could interact directly with the hydrophilic surface of the cellulose fibril. If this is the case, similarly to biomass from angiosperm plants, at least part of softwood mechanical properties may be maintained by formation of xylanocellulose complexes within cell wall macrofibrils.

Chapter 7: Concluding remarks and future work

7.1 Results summary

Work presented in this thesis evaluated the impact of xylan branches on biomass recalcitrance and investigated the specificity of enzymes likely to be responsible for the addition of GlcA branches onto softwood xylan. The impact of xylan [Me]GlcA patterning on the interaction between xylan and cellulose and the importance of this association was also evaluated. Three main outcomes were fully or partially achieved throughout this project:

The presence of GlcA branches on xylan is likely to have a significant contribution towards the maintenance of *A. thaliana* biomass recalcitrance. The analysis of Arabidopsis plants mutated in enzymes responsible for xylan acetylation, glucuronidation and methylation of the GlcA in secondary cell walls revealed that, in the conditions used in the saccharification assay, the removal of [Me]GlcA branches results in near doubling of the total monosaccharide yield. The effect was particularly significant for the release of xylose from xylan, for which the yield obtained from *gux1/2* Arabidopsis, lacking secondary cell wall glucuronidation, was seven times greater than that from the WT plants. This suggests that [Me]GlcA branches are important for the maintenance of xylan recalcitrance and that *GUX* genes may be good mutagenesis targets to improve digestibility of biomass.

GUX enzymes from two distinct clades are likely to be responsible for the glucuronidation of softwood xylan. Due to the economic importance of conifer plants and the initial discovery of the impact the absence of [Me]GlcA branches on xylan has on biomass recalcitrance, the process of softwood xylan glucuronidation was investigated. Two distinct clades of GUX enzymes were identified from conifer transcriptomes and the reads encoding the enzymes were observed to be enriched in wood forming cells. *In vitro* and *in vivo* assays were used to demonstrate that the enzymes are indeed active glucuronosyltransferases onto xylan. Interestingly, the two enzymes are likely to generate distinct patterns of GlcA branches. The PgGUX1 clade may synthesise the even pattern of acidic decorations, with a dominant spacing of six xylose monomers, and PtGUX2 clade is likely to add GlcA branches onto consecutive xylose monomers. These different structures were detected in wood and may be important for the control of the xylan-cellulose interaction with the former pattern of

[Me]GlcA branches possibly allowing for the interaction between xylan and the hydrophilic surface of the cellulose microfibril and the latter one likely to inhibit binding of xylan to some cellulose faces.

The xylan-cellulose interaction is likely to contribute towards the maintenance of secondary cell wall strength and some of this interaction may occur in the cell wall macrofibrils. The discovery of two distinct softwood enzymes which may be responsible for the synthesis of different patterns of [Me]GlcA decorations prompted investigation of the importance these different xylan structures may have for biomass properties. One hypothesis was, that the compatible, cellulose bound, xylan would contribute to the strength of the secondary cell walls. To investigate this *esk1* plants, where xylan-cellulose interaction is lost due to mis-patterning of the xylan molecule were transformed with constructs enabling over-expression of different GUX enzymes. The results indicated that the over-expression of some GUX enzymes can restore the compatible xylan patterning. This restoration of compatible patterning was correlated with improved plant growth, restoration of vessel shape and increased levels of two-fold xylan, likely to result from the xylan-cellulose interaction. Expression of GUX enzymes which lead to the maintenance of incompatible xylan patterning in *esk1* plants did not result in restoration of plant growth or vessel shape. However, it did lead to the formation of two-fold xylan molecules not observed in *esk1* plants. This suggests that two-fold xylan formation may not necessarily require compatible xylan patterning. It is possible that in *esk1* plants issues with xylan solubility or xylan aggregation prevented its deposition in the cell wall. This aggregation may be overcome by overexpression of GUX enzymes generating more branched but still incompatible xylan. Once deposited in the cell wall these incompatible xylan chains may be able to interact with the hydrophobic surface of the fibril. This interaction however may not be sufficient to maintain cell wall strength. Together, these results suggest that specific modes of the xylan-cellulose interaction may have some contribution to the maintenance of cell wall strength and may influence plant growth. Further work will be needed to investigate this relationship further. In addition to that, the cryoSEM analysis enabled identification of cell wall features which are likely to be composed from both cellulose and xylan. These macrofibrils were detected in both *A. thaliana* and in spruce wood suggesting that the xylan-cellulose interaction may occur in both hardwood and softwood.

7.2 Central role for xylan and xylan-cellulose interaction in the maintenance of cell wall cross-linking

Results presented in this thesis indicate that xylan may play a unique and central role in the maintenance of cell wall cross linking. Both the presence of [Me]GlcA branches on xylan and the interaction between xylan and the cellulose microfibril are likely to contribute to the maintenance of distinct cell wall properties. It is vital to distinguish the different roles the two features are likely to play. The presence of [Me]GlcA branches on xylan appears to be important for the maintenance of secondary cell wall recalcitrance in the *A. thaliana* model. While biomass from the *gux1/2* mutant, lacking secondary cell wall xylan glucuronidation, shows a significant increase in monosaccharide yield over the WT material in the saccharification assay, the same cannot be said for *esk1* biomass in which the reduction of xylan acetylation levels results in mis-patterning of branches and loss of the xylan-cellulose interaction (Lyczakowski et al., 2017, Grantham et al., 2017). On the other hand, *esk1* plants are dwarfed and have collapsed xylem vessels (Lefebvre et al., 2011, Xiong et al., 2013). The dwarfing and vessel collapse phenotypes are not observed for *gux1/2/3* plants lacking any [Me]GlcA branches on xylan (Mortimer et al., 2015). These observations suggest that the maintenance of biomass recalcitrance and vessel strength may be controlled by two distinct processes, both of which may have some contribution from xylan.

The mechanism by which the [Me]GlcA acid is involved in the maintenance of biomass recalcitrance is not clear. Evidence from extraction experiments and NMR analysis suggest that [Me]GlcA might form ester linkages with monolignols of lignin (Takahashi and Koshijima, 1988, Giummarella and Lawoko, 2016). The results presented in Chapter 3 of this thesis indicate that in addition to reduced recalcitrance, xylan from *gux1/2 A. thaliana* plants can be more easily extracted with DMSO from the cell wall matrix. This increase in xylan extractability suggests that absence of [Me]GlcA branches may lead to reduction of xylan cross-linking to other cell wall components which may include ester linkages to lignin. Alternative hypotheses are also possible. For example, [Me]GlcA on xylan was proposed to be involved in the maintenance of cellulose fibril alignment (Reis and Vian, 2004). The absence of xylan glucuronidation could therefore result in changes in relative positioning of fibrils which was proposed to lead to changes in lignin deposition patterns (Reis and Vian, 2004). Further work

will be required to evaluate the contribution from these or other processes to the *gux1/2* saccharification phenotype.

The possible mechanism by which xylan contributes to the maintenance of cell wall strength is also not clear. Results presented in this thesis indicate that the presence of compatible branching patterning on xylan molecules can be linked to the maintenance of Arabidopsis vessel cell wall shape and WT-like plant growth. In *esk1* plants over-expressing AtGUX2 and PtGUX2 the degree of xylan glucuronidation increases but the amount of compatible xylan patterning is similar, or lower, than the one observed in the *esk1* mutant. Similarly to the mutant, the *esk1* lines over-expressing AtGUX2 and PtGUX2 are dwarfed and have collapsed vessels. In *esk1* plants over-expressing AtGUX1 and PgGUX1 enzymes both the degree of xylan glucuronidation and the proportion of compatible xylan patterning are increased over what is detected in the *esk1* material. This coincides with an improvement in vessel shape and WT-like plant growth.

Compatible xylan patterning was proposed to allow for the interaction between xylan, in a two-fold screw conformation, and the hydrophilic surface of the cellulose microfibril (Busse-Wicher et al., 2014, Grantham et al., 2017). This hydrogen bonding based interaction could be facilitated by generation of a branch free surface in the compatible domain of a two-fold screw xylan molecule. Some two-fold screw xylan can be detected in both AtGUX1 and AtGUX2 over-expressing *esk1* plants but the surface on which this two-fold screw xylan is formed remains unknown. It is possible that the only *esk1* lines in which the binding between xylan and the hydrophilic surface of the cellulose fibril is restored are the ones in which the extent of compatible patterning is increased. Therefore, specific binding between xylan and the hydrophilic surface of the cellulose fibril may be required for the maintenance of secondary cell wall shape and normal plant growth. Xylan may be unique amongst secondary cell wall hemicelluloses of hardwood and softwood in being able to interact with the hydrophilic surface of the cellulose microfibril. The other abundant hemicellulose, galactoglucomannan, has a high mannose to glucose ratio (Scheller and Ulvskov, 2010) which is likely resulting in formation of a polymer with consecutive mannosyl residues in the backbone. Mannose is a C2 epimer of glucose and the C2 hydroxyl group of mannose may impede interaction of mannan with some hydrophilic surfaces of the cellulose fibril (Yu et al., 2018). This negative influence of the C2 hydroxyl group

of mannose may be enhanced in a polymer with consecutive mannosyl residues. Despite that, mannan polymers with different structures were demonstrated to associate with the bacterial cellulose fibril *in vitro* in an unknown manner (Whitney et al., 1998). Therefore, it will be important to investigate the exact structure of hardwood and softwood GGM to establish its role in the cell wall. It will also be very interesting to evaluate presence of any interaction of softwood xylan and GGM with cellulose. This can be achieved by performing solid state NMR experiments on ^{13}C labelled wood from conifer plants.

There are multiple possible reasons why the specific interaction between xylan and the hydrophilic surface of the cellulose microfibril may be important for the maintenance of cell wall molecular architecture. Xylan is likely to be highly cross-linked to other cell wall components. In addition to putative lignin esters (Giummarella and Lawoko, 2016) it may be linked directly to RG-I (Ralet et al., 2016) or other pectic polysaccharides via AGPs (Tan et al., 2013). It is possible that while the removal of any of these single linkages, which may occur in *gux1/2/3* or *apap1* plants, does not result in plant dwarfing the combined effect of dissociating all of the xylan-bound cell wall components from the hydrophilic surface of the cellulose microfibril may result in radical changes in the molecular architecture of both primary and secondary cell walls. This may result in vessel collapse and the dwarfing phenotype observed in *esk1* plants. Moreover, the interaction between xylan and the hydrophilic surface of the cellulose was proposed to prevent inappropriate aggregation of the microfibrils (Busse-Wicher et al., 2014). While the results presented in Chapter 6 of this thesis indicate that some aggregation of cellulose fibrils does occur in both hardwood and softwood the exact extent of it may be altered in plants where xylan does not coat the hydrophilic surface of the microfibril. This may be investigated further using some of the techniques developed as a part of this project.

7.3 Potential for application of research to industrially relevant conifer plants

One of the outcomes of this project is the discovery of the saccharification phenotype associated with the absence of [Me]GlcA branches on xylan and the identification of two clades of conifer enzymes likely to be responsible for glucuronidation of softwood xylan. Importantly, Arabidopsis plants lacking xylan glucuronidation have no adverse growth phenotypes in laboratory conditions suggesting that the improvement in

monosaccharide release and the ease of biomass processing could be transferred to industrially relevant species without a significant yield penalty. Complete removal of [Me]GlcA branches may be required to achieve full benefit of reduced biomass recalcitrance. Thus, as the conifer GUXs may have distinct activities, it is likely that representatives of both conifer *GUX* clades will need to be deactivated in any one species to achieve an improvement in softwood processing.

Transgenic technology for conifer species was developed in the early 1990s (Huang et al., 1991) and includes both *Agrobacterium tumefaciens* (Huang et al., 1991) and particle bombardment (Ellis et al., 1993) mediated transfer of the foreign genetic material. Since then, multiple conifer species have been genetically modified including commercially relevant *Picea abies, Picea glauca, Pinus taeda, Pinus radiata* and *Pseudotsuga menziesii* (Tang and Newton, 2003). The existence of transgenic technology is a prerequisite for generation of genetically engineered forestry crops. To guide these attempts both this and other projects needed to identify suitable mutagenesis or expression targets. In conifer plants this is greatly facilitated by the release of *Picea abies* (Nystedt et al., 2013) and *Pinus taeda* (Zimin et al., 2014) genomic information and a wide array of conifer transcriptomes available via the OneKP project database (Matasci et al., 2014).

Targeted engineering of wood traits has been performed mostly using different hardwood species of the *Populus* genus. These include primarily RNAi techniques which were, for example, used to engineer lignin composition in hybrid poplar by down-regulation of cinnamoyl-CoA reductase (CCR) (Van Acker et al., 2014). Trees with strongly downregulated *CCR* expression showed a reduction in biomass recalcitrance but did grow significantly worse than WT plants in field trials. In addition to an array of RNAi approaches (Mottiar et al., 2016), complete knockdown of a gene encoding another lignin biosynthesis enzyme: 4-coumarate:CoA ligase was also achieved in hybrid poplar using CRISPR-Cas9 technology (Zhou et al., 2015). Wood engineering attempts include not only knockdown or knockout approaches but also gene overexpression. In this respect, the Zip-Lignin^{TM} approach in which overexpression of a specific monolignol ferulate transferase leads to introduction of a liable ester linkages (Wilkerson et al., 2014) into the lignin polymer has great potential to decrease wood recalcitrance and improve pulping efficiency (Zhou et al., 2017). In addition to modification of hardwood properties, the downregulation of *CCR* was also

demonstrated to alter cell wall composition in cultured *Pinus radiata* tracheary elements (Wagner et al., 2013). These and other reports demonstrate that genetic engineering can be used to alter wood properties. Therefore, it is technically feasible that the discovery of the saccharification phenotype associated with absence of [Me]GlcA branches on xylan and identification of conifer GUX enzymes likely to add this branching onto softwood xylan may enable formation of transgenic conifers with less recalcitrant softwood. To avoid any possible, not yet detected, negative phenotypes associated with absence of GlcA branches of xylan, GUX activity may be reconstituted in specific conifer tissues using targeted expression. This approach was demonstrated for Arabidopsis xylan mutants using a vessel specific complementation approach (Petersen et al., 2012).

7.4 Future work

Results and analysis presented in this thesis require additional work in order to strengthen some of the conclusions discussed throughout the text. This future work may include, but will very likely not be limited to:

Further analysis of phenotypes associated with the absence of [Me]GlcA branches on xylan. One of the hypotheses proposed as a part of this thesis assumes functional equivalence between the [Me]GlcA branches on hardwood and softwood xylan. To test this hypothesis, the identified gymnosperm *GUX* genes would need to be mutagenised in conifer plants and wood properties, such as recalcitrance, would need to be studied. In addition to that, it is vital to assess the impact that the absence of [Me]GlcA branches on xylan has on plant growth in field trial conditions. This long term goal could be achieved using transgenic conifer trees or modified poplar which may be easier to generate. If generated and viable in field trial conditions the transgenic conifer plants may perform better in pulping what will be important for the paper industry. In addition to that, the modification of xylan patterning in trees may allow for generation of large quantities of material with specific polysaccharide structures. Overexpression of conifer GUX clade 2 enzymes may enable formation of large amounts of xylan with GlcA branches on consecutive monomers. Such xylans may bind metal ions and may be suitable as filtering agents in bioremediation of industrial waste. Finally by increasing the amount of compatible xylan it may be possible to generate stronger timber which can perform better in building construction.

Analysis of specific activities of conifer GUX enzymes and function of different MeGlcA patterns on softwood xylan. Results presented in this thesis indicate that the two clades of conifer GUX enzymes may have distinct activities, with one generating a compatible pattern and the other one possibly involved in addition of GlcA branches onto consecutive xylose monomers. To strengthen this conclusion it might be necessary to investigate other representatives from each clade from different species using *in vitro* and *in vivo* assay techniques described in this thesis. In addition to that it would be very interesting to evaluate the function that the two distinct patterns may have in softwood. By analysing the structure of softwood xylan further one might evaluate if these are present on the same xylan molecule or if they are enriched in certain tissues. In the long term, using transgenic conifer plants, it may be possible to establish if specific patterns of GlcA may have similar importance for maintenance of the xylan-cellulose interaction and biomass properties as proposed for Arabidopsis.

Further investigation into the role of the xylan-cellulose interaction. Results presented in this thesis indicate that the specific interaction between xylan and the hydrophilic surface of the cellulose fibril may be important for maintenance of cell wall properties and plant growth. To investigate this hypothesis it is necessary to conduct additional solid state NMR experiments on biomass from some of the transgenic lines generated as a part of this project. In particular, it is essential to establish if the two-fold screw xylan observed in AtGUX1 and AtGUX2 over-expressing *esk1* plants interacts with the same or different surface of the cellulose microfibril. In addition to that, it may be interesting to use solid state NMR to study the biomass of *esk1* plants expressing the PtGUX2 enzyme which, according to the current model, has fully incompatible xylan patterning. Finally, it would be very interesting to use the cryoSEM imaging techniques to investigate the nanostructures present in the cell walls of *esk1* and generated GUX overexpressing *esk1* plants. The appearance and size of cell wall macrofibrils may be modified in these plants, which would provide further evidence that interaction between xylan and cellulose is important for cell wall properties. This in turn, may initiate research into possible ways of controlling this interaction in plants. Such modifications may enable fine-tuning of wood mechanical properties for specific applications in different industries. In addition to that, control of the xylan cellulose interaction may be part of the plant response to different environmental stimuli and therefore it may contribute to plant development or resistance to pathogens.

References

Anders, N., Wilkinson, M. D., Lovegrove, A., Freeman, J., Tryfona, T., Pellny, T. K., Weimar, T., Mortimer, J. C., Stott, K., Baker, J. M., Defoin-Platel, M., Shewry, P. R., Dupree, P. and Mitchell, R. A. C. (2012) 'Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses', *Proceedings of the National Academy of Sciences of the United States of America,* 109(3), pp. 989-993.

Andersson, S. I., Samuelson, O., Ishihara, M. and Shimizu, K. (1983) 'Structure of the reducing-end groups of xylan', *Carbohydrate Research,* 111(2), pp. 283-288.

Andrew, E. R. (2010) 'Magic Angle Spinning', in McDermott, A.E. & Polenova, T. (eds.) *Solid State NMR Studies of Biopolymers*. Hoboken, New Jersey, United States: John Wiley & Sons.

Arabidopsis-Genome-Initiative (2000) 'Analysis of the genome sequence of the flowering plant Arabidopsis thaliana', *Nature,* 408(6814), pp. 796-815.

Atanassov, II, Etchells, J. P. and Turner, S. R. (2009) 'A simple, flexible and efficient PCR-fusion/Gateway cloning procedure for gene fusion, site-directed mutagenesis, short sequence insertion and domain deletions and swaps', *Plant Methods,* 5.

Bajpai, P. (2015) 'Basic Overview of Pulp and Paper Manufacturing Process', in Bajpai, P. (ed.) *Green Chemistry and Sustainability in Pulp and Paper Industry*. Cham, Switzerland: Springer.

Baker, A. A., Helbert, W., Sugiyama, J. and Miles, M. J. (1997) 'High-resolution atomic force microscopy of native Valonia cellulose I microcrystals', *Journal of Structural Biology,* 119(2), pp. 129-138.

Bensussan, M., Lefebvre, V., Ducamp, A., Trouverie, J., Gineau, E., Fortabat, M. N., Guillebaux, A., Baldy, A., Naquin, D., Herbette, S., Lapierre, C., Mouille, G., Horlow, C. and Durand-Tardif, M. (2015) 'Suppression of Dwarf and irregular xylem Phenotypes Generates Low-Acetylated Biomass Lines in Arabidopsis', *Plant Physiology,* 168(2), pp. 452-+.

Bernhardt, C. and Tierney, M. L. (2000) 'Expression of AtPRP3, a proline-rich structural cell wall protein from arabidopsis, is regulated by cell-type-specific developmental pathways involved in root hair formation', *Plant Physiology,* 122(3), pp. 705-714.

Beylot, M. H., McKie, V. A., Voragen, A. G. J., Doeswijk-Voragen, C. H. L. and Gilbert, H. J. (2001) 'The Pseudomonas cellulosa glycoside hydrolase family 51 arabinofuranosidase exhibits wide substrate specificity', *Biochemical Journal,* 358, pp. 607-614.

Biely, P., Malovikova, A., Uhliarikova, I., Li, X. L. and Wong, D. W. S. (2015) 'Glucuronoyl esterases are active on the polymeric substrate methyl esterified glucuronoxylan', *Febs Letters,* 589(18), pp. 2334-2339.

Bischoff, V., Nita, S., Neumetzler, L., Schindelasch, D., Urbain, A., Eshed, R., Persson, S., Delmer, D. and Scheible, W. R. (2010) 'TRICHOME BIREFRINGENCE and Its Homolog AT5G01360 Encode Plant-Specific DUF231 Proteins Required for Cellulose Biosynthesis in Arabidopsis', *Plant Physiology,* 153(2), pp. 590-602.

Biswal, A. K., Atmodjo, M. A., Li, M., Baxter, H. L., Yoo, C. G., Pu, Y., Lee, Y.-C., Mazarei, M., Black, I. M., Zhang, J.-Y., Ramanna, H., Bray, A. L., King, Z. R., LaFayette, P. R., Pattathil, S., Donohoe, B. S., Mohanty, S. S., Ryno, D., Yee, K., Thompson, O. A., Rodriguez, M., Jr., Dumitrache, A., Natzke, J., Winkeler, K., Collins, C., Yang, X., Tan, L., Sykes, R. W., Gjersing, E. L., Ziebell, A., Turner, G. B., Decker, S. R., Hahn, M. G., Davison, B. H., Udvardi, M. K., Mielenz, J. R., Davis, M. F., Nelson, R. S., Parrott, W. A., Ragauskas, A. J., Stewart, C. N., Jr. and Mohnen, D. (2018) 'Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis', *Nature Biotechnology,* 36(3), pp. 249-+.

Bligh, E. G. and Dyer, W. J. (1959) 'A rapid method for total lipid extraction and purification', *Canadian Journal of Biochemistry and Physiology,* 37(8), pp. 911-917.

Boerjan, W., Ralph, J. and Baucher, M. (2003) 'Lignin biosynthesis', *Annual Review of Plant Biology,* 54, pp. 519-546.

Bonawitz, N. D. and Chapple, C. (2010) 'The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype', *Annual Review of Genetics, Vol 44,* 44, pp. 337-363.

Bosmans, T. J., Stepan, A. M., Toriz, G., Renneckar, S., Karabulut, E., Wagberg, L. and Gatenholm, P. (2014) 'Assembly of Debranched Xylan from Solution and on Nanocellulosic Surfaces', *Biomacromolecules,* 15(3), pp. 924-930.

Braidwood, L., Breuer, C. and Sugimoto, K. (2014) 'My body is a cage: mechanisms and modulation of plant cell growth', *New Phytologist,* 201(2), pp. 388-402.

Brodribb, T. J., McAdam, S. A. M., Jordan, G. J. and Martins, S. C. V. (2014) 'Conifer species adapt to low-rainfall climates by following one of two divergent pathways', *Proceedings of the National Academy of Sciences of the United States of America,* 111(40), pp. 14489-14493.

Bromley, J. R., Busse-Wicher, M., Tryfona, T., Mortimer, J. C., Zhang, Z., Brown, D. M. and Dupree, P. (2013) 'GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns', *Plant Journal,* 74(3), pp. 423-434.

Brown, D., Wightman, R., Zhang, Z. N., Gomez, L. D., Atanassov, I., Bukowski, J. P., Tryfona, T., McQueen-Mason, S. J., Dupree, P. and Turner, S. (2011) 'Arabidopsis genes IRREGULAR XYLEM (IRX15) and IRX15L encode DUF579-containing proteins that are essential for normal xylan deposition in the secondary cell wall', *Plant Journal,* 66(3), pp. 401-413.

Brown, D. M., Goubet, F., Vicky, W. W. A., Goodacre, R., Stephens, E., Dupree, P. and Turner, S. R. (2007) 'Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis', *Plant Journal,* 52(6), pp. 1154-1168.

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(8), pp. 2281-2295.

Brown, D. M., Zhang, Z. N., 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 Journal,* 57(4), pp. 732-746.

Buschiazzo, E., Ritland, C., Bohlmann, J. and Ritland, K. (2012) 'Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms', *Bmc Evolutionary Biology,* 12.

Busse-Wicher, M., Gomes, T. C. F., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N. J., Bolam, D. N., Skaf, M. S. and Dupree, P. (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 Journal,* 79(3), pp. 492-506.

Busse-Wicher, M., Grantham, N. J., Lyczakowski, J. J., Nikolovski, N. and Dupree, P. (2016a) 'Xylan decoration patterns and the plant secondary cell wall molecular architecture', *Biochemical Society Transactions,* 44(1), pp. 74-78.

Busse-Wicher, M., Li, A., Silveira, R. L., Pereira, C. S., Tryfona, T., Gomes, T. C. F., Skaf, M. S. and Dupree, P. (2016b) 'Evolution of Xylan Substitution Patterns in Gymnosperms and Angiosperms: Implications for Xylan Interaction with Cellulose', *Plant Physiology,* 171(4), pp. 2418-2431.

Cave, I. D. and Walker, J. C. F. (1994) ' Stiffness of wood in fast-grown plantation softwoods - the influence of microfibril angle ', *Forest Products Journal,* 44(5), pp. 43-48.

Chaikuad, A., Froese, D. S., Berridge, G., von Delft, F., Oppermann, U. and Yue, W. W. (2011) 'Conformational plasticity of glycogenin and its maltosaccharide substrate during glycogen biogenesis', *Proceedings of the National Academy of Sciences of the United States of America,* 108(52), pp. 21028-21033.

Chiniquy, D., Sharma, V., Schultink, A., Baidoo, E. E., Rautengarten, C., Cheng, K., Carroll, A., Ulvskov, P., Harholt, J., Keasling, J. D., Pauly, M., Scheller, H. V. and Ronald, P. C. (2012) 'XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan', *Proceedings of the National Academy of Sciences of the United States of America,* 109(42), pp. 17117-17122.

Chong, S. L., Derba-Maceluch, M., Koutaniemi, S., Gomez, L. D., McQueen-Mason, S. J., Tenkanen, M. and Mellerowicz, E. J. (2015) 'Active fungal GH115 alpha-glucuronidase produced in Arabidopsis thaliana affects only the UX1-reactive glucuronate decorations on native glucuronoxylans', *Bmc Biotechnology,* 15.

Chou, E. Y., Schuetz, M., Hoffmann, N., Watanabe, Y., Sibout, R. and Samuels, A. L. (2018) 'Distribution, mobility, and anchoring of lignin-related oxidative enzymes in Arabidopsis secondary cell walls', *Journal of Experimental Botany,* 69(8), pp. 1849-1859.

Christenhusz, M. J. M. and Byng, J. W. (2016) 'The number of known plants species in the world and its annual increase', *Phytotaxa,* 261(3), pp. 201-217.

Clough, S. J. and Bent, A. F. (1998) 'Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana', *Plant Journal,* 16(6), pp. 735-743.

Cocuron, J. C., Lerouxel, O., Drakakaki, G., Alonso, A. P., Liepman, A. H., Keegstra, K., Raikhel, N. and Wilkerson, C. G. (2007) 'A gene from the cellulose synthase-like C family encodes a beta-1,4 glucan synthase', *Proceedings of the National Academy of Sciences of the United States of America,* 104(20), pp. 8550-8555.

Cornuault, V., Buffetto, F., Rydahl, M. G., Marcus, S. E., Torode, T. A., Xue, J., Crepeau, M. J., Faria-Blanc, N., Willats, W. G. T., Dupree, P., Ralet, M. C. and Knox, J. P. (2015) 'Monoclonal antibodies indicate low-abundance links between heteroxylan and other glycans of plant cell walls', *Planta,* 242(6), pp. 1321-1334.

Cosgrove, D. J. (2000) 'Loosening of plant cell walls by expansins', *Nature,* 407(6802), pp. 321-326.

Cosgrove, D. J. (2005) 'Growth of the plant cell wall', *Nature Reviews Molecular Cell Biology,* 6(11), pp. 850-861.

Cosgrove, D. J. and Jarvis, M. C. (2012) 'Comparative structure and biomechanics of plant primary and secondary cell walls', *Frontiers in Plant Science,* 3.

Culbertson, A. T., Ehrlich, J. J., Choe, J. Y., Honzatko, R. B. and Zabotina, O. A. (2018) 'Structure of xyloglucan xylosyltransferase 1 reveals simple steric rules that define biological patterns of xyloglucan polymers', *Proceedings of the National Academy of Sciences of the United States of America,* 115(23), pp. 6064-6069.

d'Errico, C., Borjesson, J., Ding, H. S., Krogh, K., Spodsberg, N., Madsen, R. and Monrad, R. N. (2016) 'Improved biomass degradation using fungal glucuronoyl-esterases-hydrolysis of natural corn fiber substrate', *Journal of Biotechnology,* 219, pp. 117-123.

de Souza, W. R., Martins, P. K., Freeman, J., Pellny, T. K., Michaelson, L. V., Sampaio, B. L., Vinecky, F., Ribeiro, A. P., da Cunha, B., Kobayashi, A. K., de Oliveira, P. A., Campanha, R. B., Pacheco, T. F., Martarello, D. C. I., Marchiosi, R., Ferrarese, O., dos Santos, W. D., Tramontina, R., Squina, F. M., Centeno, D. C., Gaspar, M., Braga, M. R., Tine, M. A. S., Ralph, J., Mitchell, R. A. C. and Molinari, H. B. C. (2018) 'Suppression of a single BAHD gene in Setaria viridis causes large, stable decreases in cell wall feruloylation and increases biomass digestibility', *New Phytologist,* 218(1), pp. 81-93.

Dick-Perez, M., Zhang, Y. A., Hayes, J., Salazar, A., Zabotina, O. A. and Hong, M. (2011) 'Structure and Interactions of Plant Cell-Wall Polysaccharides by Two- and Three-Dimensional Magic-Angle-Spinning Solid-State NMR', *Biochemistry,* 50(6), pp. 989-1000.

Dinwoodie, J. M. (2000) *Timber: Its Nature and Behaviour.* 2 edn. Abingdon, Unite Kingdom: Taylor and Francis.

Doblin, M. S., Kurek, I., Jacob-Wilk, D. and Delmer, D. P. (2002) 'Cellulose biosynthesis in plants: from genes to rosettes', *Plant and Cell Physiology,* 43(12), pp. 1407-1420.

Dodd, D. and Cann, I. K. O. (2009) 'Enzymatic deconstruction of xylan for biofuel production', *Global Change Biology Bioenergy,* 1(1), pp. 2-17.

Donaldson, L. (2007) 'Cellulose microfibril aggregates and their size variation with cell wall type', *Wood Science and Technology,* 41(5), pp. 443-460.

Du, X. Y., Perez-Boada, M., Fernandez, C., Rencoret, J., del Rio, J. C., Jimenez-Barbero, J., Li, J. B., Gutierrez, A. and Martinez, A. T. (2014) 'Analysis of lignin-carbohydrate and lignin-lignin

linkages after hydrolase treatment of xylan-lignin, glucomannan-lignin and glucan-lignin complexes from spruce wood', *Planta,* 239(5), pp. 1079-1090.
Dupree, R., Simmons, T. J., Mortimer, J. C., Patel, D., Iuga, D., Brown, S. P. and Dupree, P. (2015) 'Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional C-13 Solid State Nuclear Magnetic Resonance Spectroscopy', *Biochemistry,* 54(14), pp. 2335-2345.
Ebringerova, A., Hromadkova, Z. and Heinze, T. (2005) 'Hemicellulose', *Polysaccharides 1: Structure, Characterization and Use,* 186, pp. 1-67.
Edwards, M. E., Dickson, C. A., Chengappa, S., Sidebottom, C., Gidley, M. J. and Reid, J. S. G. (1999) 'Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis', *Plant Journal,* 19(6), pp. 691-697.
Ellis, D. D., McCabe, D. E., McInnis, S., Ramachandran, R., Russell, D. R., Wallace, K. M., Martinell, B. J., Roberts, D. R., Raffa, K. F. and McCown, B. H. (1993) 'Stable transformation of Picea glacua by particle acceleration', *Bio-Technology,* 11(1), pp. 84-89.
Fagard, M., Desnos, T., Desprez, T., Goubet, F., Refregier, G., Mouille, G., McCann, M., Rayon, C., Vernhettes, S. and Hofte, H. (2000) 'PROCUSTE1 encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of arabidopsis', *Plant Cell,* 12(12), pp. 2409-2423.
Faik, A. (2010) 'Xylan Biosynthesis: News from the Grass', *Plant Physiology,* 153(2), pp. 396-402.
Faria-Blanc, N., Mortimer, J. C. and Dupree, P. (2018) 'A Transcriptomic Analysis of Xylan Mutants Does Not Support the Existence of a Secondary Cell Wall Integrity System in Arabidopsis', *Frontiers in Plant Science,* 9.
Fernandes, A. N., Thomas, L. H., Altaner, C. M., Callow, P., Forsyth, V. T., Apperley, D. C., Kennedy, C. J. and Jarvis, M. C. (2011) 'Nanostructure of cellulose microfibrils in spruce wood', *Proceedings of the National Academy of Sciences of the United States of America,* 108(47), pp. E1195-E1203.
Fry, S. C. (1988) *The growing plant cell wall: chemical and metabolic analysis.* Harlow, United Kingdom: Longman Group Limited.
Fry, S. C. (1989) 'The structure and function of xyloglucan', *Journal of Experimental Botany,* 40(210), pp. 1-11.
Fu, C. X., Xiao, X. R., Xi, Y. J., Ge, Y. X., Chen, F., Bouton, J., Dixon, R. A. and Wang, Z. Y. (2011) 'Downregulation of Cinnamyl Alcohol Dehydrogenase (CAD) Leads to Improved Saccharification Efficiency in Switchgrass', *Bioenergy Research,* 4(3), pp. 153-164.
Gille, S., de Souza, A., Xiong, G. Y., Benz, M., Cheng, K., Schultink, A., Reca, I. B. and Pauly, M. (2011) 'O-Acetylation of Arabidopsis Hemicellulose Xyloglucan Requires AXY4 or AXY4L, Proteins with a TBL and DUF231 Domain', *Plant Cell,* 23(11), pp. 4041-4053.
Giummarella, N. and Lawoko, M. (2016) 'Structural Basis for the Formation and Regulation of Lignin-Xylan Bonds in Birch', *Acs Sustainable Chemistry & Engineering,* 4(10), pp. 5319-5326.
Goellner, E. M., Utermoehlen, J., Kramer, R. and Classen, B. (2011) 'Structure of arabinogalactan from Larix laricina and its reactivity with antibodies directed against type-II-arabinogalactans', *Carbohydrate Polymers,* 86(4), pp. 1739-1744.
Gonneau, M., Desprez, T., Guillot, A., Vernhettes, S. and Hofte, H. (2014) 'Catalytic Subunit Stoichiometry within the Cellulose Synthase Complex', *Plant Physiology,* 166(4), pp. 1709-1712.
Goubet, F., Barton, C. J., Mortimer, J. C., Yu, X. L., Zhang, Z. N., Miles, G. P., Richens, J., Liepman, A. H., Seffen, K. and Dupree, P. (2009) 'Cell wall glucomannan in Arabidopsis is synthesised by CSLA glycosyltransferases, and influences the progression of embryogenesis', *Plant Journal,* 60(3), pp. 527-538.
Goubet, F., Dupree, P. and Johansen, K. S. (2011) 'Carbohydrate Gel Electrophoresis', *Plant Cell Wall: Methods and Protocols,* 715, pp. 81-92.

Goubet, F., Jackson, P., Deery, M. J. and Dupree, P. (2002) 'Polysaccharide analysis using carbohydrate gel electrophoresis: A method to study plant cell wall polysaccharides and polysaccharide hydrolases', *Analytical Biochemistry,* 300(1), pp. 53-68.

Goubet, F. and Mohnen, D. (1999) 'Solubilization and partial characterization of homogalacturonan-methyltransferase from microsomal membranes of suspension-cultured tobacco cells', *Plant Physiology,* 121(1), pp. 281-290.

Grabber, J. H. (2005) 'How do lignin composition, structure, and cross-linking affect degradability? A review of cell wall model studies', *Crop Science,* 45(3), pp. 820-831.

Grantham, N. J., Wurman-Rodrich, J., Terrett, O. M., Lyczakowski, J. J., Stott, K., Iuga, D., Simmons, T. J., Durand-Tardif, M., Brown, S. P., Dupree, R., Busse-Wicher, M. and Dupree, P. (2017) 'An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls', *Nature Plants,* 3(11), pp. 859-865.

Griffiths, J. S. and North, H. M. (2017) 'Sticking to cellulose: exploiting Arabidopsis seed coat mucilage to understand cellulose biosynthesis and cell wall polysaccharide interactions', *New Phytologist,* 214(3), pp. 959-966.

Gross, A. S. and Chu, J.-W. (2010) 'On the Molecular Origins of Biomass Recalcitrance: The Interaction Network and Solvation Structures of Cellulose Microfibrils', *Journal of Physical Chemistry B,* 114(42), pp. 13333-13341.

Guan, R., Zhao, Y. P., Zhang, H., Fan, G. Y., Liu, X., Zhou, W. B., Shi, C. C., Wang, J. H., Liu, W. Q., Liang, X. M., Fu, Y. Y., Ma, K. L., Zhao, L. J., Zhang, F. M., Lu, Z. H., Lee, S. M. Y., Xu, X., Wang, J., Yang, H. M., Fu, C. X., Ge, S. and Chen, W. B. (2016) 'Draft genome of the living fossil Ginkgo biloba', *Gigascience,* 5.

Ha, M. A., MacKinnon, I. M., Sturcova, A., Apperley, D. C., McCann, M. C., Turner, S. R. and Jarvis, M. C. (2002) 'Structure of cellulose-deficient secondary cell walls from the irx3 mutant of Arabidopsis thaliana', *Phytochemistry,* 61(1), pp. 7-14.

Hallac, B. B. and Ragauskas, A. J. (2011) 'Analyzing cellulose degree of polymerization and its relevancy to cellulosic ethanol', *Biofuels Bioproducts & Biorefining-Biofpr,* 5(2), pp. 215-225.

Harholt, J., Suttangkakul, A. and Scheller, H. V. (2010) 'Biosynthesis of Pectin', *Plant Physiology,* 153(2), pp. 384-395.

Harris, D., Stork, J. and Debolt, S. (2009) 'Genetic modification in cellulose-synthase reduces crystallinity and improves biochemical conversion to fermentable sugar', *Global Change Biology Bioenergy,* 1(1), pp. 51-61.

Hocq, L., Pelloux, J. and Lefebvre, V. (2017) 'Connecting Homogalacturonan-Type Pectin Remodeling to Acid Growth', *Trends in Plant Science,* 22(1), pp. 20-29.

Hofte, H. and Voxeur, A. (2017) 'Plant cell walls', *Current Biology,* 27(17), pp. R865-R870.

Holland, N., Holland, D., Helentjaris, T., Dhugga, K. S., Xoconostle-Cazares, B. and Delmer, D. P. (2000) 'A comparative analysis of the plant cellulose synthase (CesA) gene family', *Plant Physiology,* 123(4), pp. 1313-1323.

Hood, E. E. (2016) 'Plant-based biofuels', *F1000Research,* 5(185).

Huang, C. L., Lindstrom, H., Nakada, R. and Ralston, J. (2003) 'Cell wall structure and wood properties determined by acoustics - a selective review', *Holz Als Roh-Und Werkstoff,* 61(5), pp. 321-335.

Huang, Y., Diner, A. M. and Karnosky, D. F. (1991) 'Agrobacterium rhizogenes-mediated genetic transformation and regeneration of a conifer:Larix decidua ', *In Vitro Cellular & Developmental Biology - Plant,* 27(4), pp. 201–207.

Hult, E. L., Larsson, P. T. and Iversen, T. (2000) 'A comparative CP/MAS C-13-NMR study of cellulose structure in spruce wood and kraft pulp', *Cellulose,* 7(1), pp. 35-55.

Jarvis, M. C. (2013) 'Cellulose Biosynthesis: Counting the Chains', *Plant Physiology,* 163(4), pp. 1485-1486.

Jarvis, M. C. (2018) 'Structure of native cellulose microfibrils, the starting point for nanocellulose manufacture', *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences,* 376(2112).

Jensen, J. K., Johnson, N. and Wilkerson, C. G. (2013) 'Discovery of diversity in xylan biosynthetic genes by transcriptional profiling of a heteroxylan containing mucilaginous tissue', *Frontiers in Plant Science,* 4.

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 Journal,* 66(3), pp. 387-400.

Jensen, J. K., Sorensen, S. O., Harholt, J., Geshi, N., Sakuragi, Y., Moller, I., Zandleven, J., Bernal, A. J., Jensen, N. B., Sorensen, C., Pauly, M., Beldman, G., Willats, W. G. T. and Scheller, H. V. (2008) 'Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in Arabidopsis', *Plant Cell,* 20(5), pp. 1289-1302.

Johnson, K. L., Cassin, A. M., Lonsdale, A., Wong, G. K. S., Soltis, D. E., Miles, N. W., Melkonian, M., Melkonian, B., Deyholos, M. K., Leebens-Mack, J., Rothfels, C. J., Stevenson, D. W., Graham, S. W., Wang, X. M., Wu, S. X., Pires, J. C., Edger, P. P., Carpenter, E. J., Bacic, A., Doblin, M. S. and Schultz, C. J. (2017) 'Insights into the Evolution of Hydroxyproline-Rich Glycoproteins from 1000 Plant Transcriptomes', *Plant Physiology,* 174(2), pp. 904-921.

Jokipii-Lukkari, S., Sundell, D., Nilsson, O., Hvidsten, T. R., Street, N. R. and Tuominen, H. (2017) 'NorWood: a gene expression resource for evo-devo studies of conifer wood development', *New Phytologist,* 216(2), pp. 482-494.

Jones, R., Ougham, H., Thomas, H. and Waaland, S. (2013) *The Molecular Life of Plants.* Oxford, UK: Wiley-Blackwell and The American Society of Plant Biologists.

Kabel, M. A., van den Borne, H., Vincken, J. P., Voragen, A. G. J. and Schols, H. A. (2007) 'Structural differences of xylans affect their interaction with cellulose', *Carbohydrate Polymers,* 69(1), pp. 94-105.

Karol, K. G., McCourt, R. M., Cimino, M. T. and Delwiche, C. F. (2001) 'The closest living relatives of land plants', *Science,* 294(5550), pp. 2351-2353.

Katzen, F. (2007) 'Gateway (R) recombinational cloning: a biological operating system', *Expert Opinion on Drug Discovery,* 2(4), pp. 571-589.

Keller, B. and Baumgartner, C. (1991) 'VASCULAR-SPECIFIC EXPRESSION OF THE BEAN GRP-1.8 GENE IS NEGATIVELY REGULATED', *Plant Cell,* 3(10), pp. 1051-1061.

Kenrick, P. and Crane, P. R. (1997) 'The origin and early evolution of plants on land', *Nature,* 389(6646), pp. 33-39.

Kieliszewski, M. J. (2001) 'The latest hype on Hyp-O-glycosylation codes', *Phytochemistry,* 57(3), pp. 319-323.

Kimura, S., Laosinchai, W., Itoh, T., Cui, X. J., Linder, C. R. and Brown, R. M. (1999) 'Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant Vigna angularis', *Plant Cell,* 11(11), pp. 2075-2085.

Kulkarni, A. R., Pena, M. J., Avci, U., Mazumder, K., Urbanowicz, B. R., Pattathil, S., Yin, Y., O'Neill, M. A., Roberts, A. W., Hahn, M. G., Xu, Y., Darvill, A. G. and York, W. S. (2012) 'The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes', *Glycobiology,* 22(3), pp. 439-451.

Kumar, M. and Turner, S. (2015) 'Plant cellulose synthesis: CESA proteins crossing kingdoms', *Phytochemistry,* 112, pp. 91-99.

Lamport, D. T. A., Kieliszewski, M. J., Chen, Y. N. and Cannon, M. C. (2011) 'Role of the Extensin Superfamily in Primary Cell Wall Architecture', *Plant Physiology,* 156(1), pp. 11-19.

Larsson, P. T., Wickholm, K. and Iversen, T. (1997) 'A CP/MAS C-13 NMR investigation of molecular ordering in celluloses', *Carbohydrate Research,* 302(1-2), pp. 19-25.

Lee, C., Teng, Q., Zhong, R. Q. and Ye, Z. H. (2012) 'Arabidopsis GUX Proteins Are Glucuronyltransferases Responsible for the Addition of Glucuronic Acid Side Chains onto Xylan', *Plant and Cell Physiology,* 53(7), pp. 1204-1216.

Lefebvre, V., Fortabat, M. N., Ducamp, A., North, H. M., Maia-Grondard, A., Trouverie, J., Boursiac, Y., Mouille, G. and Durand-Tardif, M. (2011) 'ESKIMO1 Disruption in Arabidopsis Alters Vascular Tissue and Impairs Water Transport', *Plos One,* 6(2).

Lewicka, A., Lyczakowski, J., Blackhurst, G., Pashkuleva, C., Rothschild-Mancinelli, K., Tautvaišas, D., Thornton, H., Villanueva, H., Xiao, W., Slikas, J., Horsfall, L., Elfick, A. and French, C. (2014a) 'Fusion of Pyruvate Decarboxylase and Alcohol Dehydrogenase Increases Ethanol Production in Escherichia coli', *ACS Synthetic Biology,* 3(12), pp. 976 - 978.

Li, X., Jackson, P., Rubtsov, D. V., Faria-Blanc, N., Mortimer, J. C., Turner, S. R., Krogh, K. B., Johansen, K. S. and Dupree, P. (2013) 'Development and application of a high throughput carbohydrate profiling technique for analyzing plant cell wall polysaccharides and carbohydrate active enzymes', *Biotechnology for Biofuels,* 6.

Liepman, A. H., Wightman, R., Geshi, N., Turner, S. R. and Scheller, H. V. (2010) 'Arabidopsis - a powerful model system for plant cell wall research', *Plant Journal,* 61(6), pp. 1107-1121.

Liepman, A. H., Wilkerson, C. G. and Keegstra, K. (2005) 'Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases', *Proceedings of the National Academy of Sciences of the United States of America,* 102(6), pp. 2221-2226.

Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M. and Henrissat, B. (2014) 'The carbohydrate-active enzymes database (CAZy) in 2013', *Nucleic Acids Research,* 42(D1), pp. D490-D495.

Loque, D., Scheller, H. V. and Pauly, M. (2015) 'Engineering of plant cell walls for enhanced biofuel production', *Current Opinion in Plant Biology,* 25, pp. 151-161.

Lyczakowski, J. J., Wicher, K. B., Terrett, O. M., Faria-Blanc, N., Yu, X. L., Brown, D., Krogh, K., Dupree, P. and Busse-Wicher, M. (2017) 'Removal of glucuronic acid from xylan is a strategy to improve the conversion of plant biomass to sugars for bioenergy', *Biotechnology for Biofuels,* 10.

Macquet, A., Ralet, M. C., Kronenberger, J., Marion-Poll, A. and North, H. M. (2007) 'In situ, chemical and macromolecular study of the composition of Arabidopsis thaliana seed coat mucilage', *Plant and Cell Physiology,* 48(7), pp. 984-999.

Manabe, Y., Verhertbruggen, Y., Gille, S., Harholt, J., Chong, S. L., Pawar, P. M. A., Mellerowicz, E. J., Tenkanen, M., Cheng, K., Pauly, M. and Scheller, H. V. (2013) 'Reduced Wall Acetylation Proteins Play Vital and Distinct Roles in Cell Wall O-Acetylation in Arabidopsis', *Plant Physiology,* 163(3), pp. 1107-1117.

Mangeon, A., Junqueira, R. M. and Sachetto-Martins, G. (2010) 'Functional diversity of the plant glycine-rich proteins superfamily.', *Plant Signaling & Behavior,* 5(2), pp. 99 - 104.

Mapes, G. and Rothwell, G. (1991) 'Structure and relationships of primitive conifers', *Neues Jahrb. Geol. Palaontol. Abh.,* 183(1), pp. 269-287.

Martinez-Abad, A., Berglund, J., Toriz, G., Gatenholm, P., Henriksson, G., Lindstrom, M., Wohlert, J. and Vilaplana, F. (2017) 'Regular Motifs in Xylan Modulate Molecular Flexibility and Interactions with Cellulose Surfaces', *Plant Physiology,* 175(4), pp. 1579-1592.

Matasci, N., Hung, L. H., Yan, Z. X., Carpenter, E. J., Wickett, N. J., Mirarab, S., Nguyen, N., Warnow, T., Ayyampalayam, S., Barker, M., Burleigh, J. G., Gitzendanner, M. A., Wafula, E., Der, J. P., dePamphilis, C. W., Roure, B., Philippe, H., Ruhfel, B. R., Miles, N. W., Graham, S. W., Mathews, S., Surek, B., Melkonian, M., Soltis, D. E., Soltis, P. S., Rothfels, C., Pokorny, L., Shaw, J. A., DeGironimo, L., Stevenson, D. W., Villarreal, J. C., Chen, T., Kutchan, T. M., Rolf, M., Baucom, R. S., Deyholos, M. K., Samudrala, R., Tian, Z. J., Wu, X. L., Sun, X., Zhang, Y., Wang, J., Leebens-Mack, J. and Wong, G. K. S. (2014) 'Data access for the 1,000 Plants (1KP) project', *Gigascience,* 3.

McCann, M. C. and Carpita, N. C. (2015) 'Biomass recalcitrance: a multi-scale, multi-factor, and conversion-specific property', *Journal of Experimental Botany,* 66(14), pp. 4109-4118.
McCartney, L., Marcus, S. E. and Knox, J. P. (2005) 'Monoclonal antibodies to plant cell wall xylans and arabinoxylans', *Journal of Histochemistry & Cytochemistry,* 53(4), pp. 543-546.
McKinley, D. C., Ryan, M. G., Birdsey, R. A., Giardina, C. P., Harmon, M. E., Heath, L. S., Houghton, R. A., Jackson, R. B., Morrison, J. F., Murray, B. C., Pataki, D. E. and Skog, K. E. (2011) 'A synthesis of current knowledge on forests and carbon storage in the United States', *Ecological Applications,* 21(6), pp. 1902-1924.
Meents, M. J., Watanabe, Y. and Samuels, A. L. (2018) 'The cell biology of secondary cell wall biosynthesis', *Annals of Botany,* 121(6), pp. 1107-1125.
Meier, H. and Reid, J. (1982) 'Reserve polysaccharides other than starch in higher plants', in Loewus , F.A. & Tanner , W. (eds.) *Plant carbohydrates I*. Berlin and Heidelberg, Germany: Springer-Verlag
Micheli, F. (2001) 'Pectin methylesterases: cell wall enzymes with important roles in plant physiology', *Trends in Plant Science,* 6(9), pp. 414-419.
Mikkelsen, D., Flanagan, B. M., Wilson, S. M., Bacic, A. and Gidley, M. J. (2015) 'Interactions of Arabinoxylan and (1,3)(1,4)-beta-Glucan with Cellulose Networks', *Biomacromolecules,* 16(4), pp. 1232-1239.
Mikkelsen, M. D., Harholt, J., Ulvskov, P., Johansen, I. E., Fangel, J. U., Doblin, M. S., Bacic, A. and Willats, W. G. T. (2014) 'Evidence for land plant cell wall biosynthetic mechanisms in charophyte green algae', *Annals of Botany,* 114(6), pp. 1217-1236.
Mohnen, D. (2008) 'Pectin structure and biosynthesis', *Current Opinion in Plant Biology,* 11(3), pp. 266-277.
Moller, I., Sorensen, I., Bernal, A. J., Blaukopf, C., Lee, K., Obro, J., Pettolino, F., Roberts, A., Mikkelsen, J. D., Knox, J. P., Bacic, A. and Willats, W. G. T. (2007) 'High-throughput mapping of cell-wall polymers within and between plants using novel microarrays', *Plant Journal,* 50(6), pp. 1118-1128.
Monrad, R. N., Eklöf, J., Krogh, K. B. R. M. and Biely, P. (2018) 'Glucuronoyl esterases: diversity, properties and biotechnological potential. A review', *Critical Reviews in Biotechnology*.
Moore, J. (2011) *Wood properties and uses of Sitka spruce in Britain*, Edinburgh: Forestry Comission.
Moreira, L. R. S. and Filho, E. X. F. (2008) 'An overview of mannan structure and mannan-degrading enzyme systems', *Applied Microbiology and Biotechnology,* 79(2), pp. 165-178.
Mortimer, J. C., Faria-Blanc, N., Yu, X., Tryfona, T., Sorieul, M., Ng, Y. Z., Zhang, Z., Stott, K., Anders, N. and Dupree, P. (2015) 'An unusual xylan in Arabidopsis primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14', *Plant Journal,* 83(3), pp. 413-527.
Mortimer, J. C., Miles, G. P., Brown, D. M., Zhang, Z., Segura, M. P., Weimar, T., Yu, X., Seffen, K. A., Stephens, E., Turner, S. R. and Dupree, P. (2010) 'Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass', *Proceedings of the National Academy of Sciences of the United States of America,* 107(40), pp. 17409-17414.
Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J. and Mansfield, S. D. (2016) 'Designer lignins: harnessing the plasticity of lignification', *Current Opinion in Biotechnology,* 37, pp. 190-200.
Mutte, S. K., Kato, H., Rothfels, C., Melkonian, M., Wong, G. K. S. and Weijers, D. (2018) 'Origin and evolution of the nuclear auxin response system', *Elife,* 7.
Myburg, A., Lev-Yadun, S. and Sederoff, R. 2013. Xylem Structure and Function. *Encyclopedia of Life Sciences (eLS).* Chichester: John Wiley & Sons.
Nabuurs, G. J., O., Masera, K., Andrasko, P., Benitez-Ponce, R., Boer, M., Dutschke, E., Elsiddig, J., Ford-Robertson, P., Frumhoff, T., Karjalainen, O., Krankina, W. A., Kurz, M., Matsumoto, W., Oyhantcabal, N. H., Ravindranath, M. J. and Sanz Sanchez, X. (2007) 'Forestry', in Metz, B., Davidson, O.R., Bosch, P.R., Dave, R. & Meyer, L.A. (eds.) *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental*

Panel on Climate Change. Cambridge, United Kingdom and New York, NY, USA.: Cambridge University Press.

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A. S., Basle, A., Gray, J., Venditto, I., Briggs, J., Zhang, X. Y., Labourel, A., Terrapon, N., Buffetto, F., Nepogodiev, S., Xiao, Y., Field, R. A., Zhu, Y. P., O'Neill, M. A., Urbanowicz, B. R., York, W. S., Davies, G. J., Abbott, D. W., Ralet, M. C., Martens, E. C., Henrissat, B. and Gilbert, H. J. (2017) 'Complex pectin metabolism by gut bacteria reveals novel catalytic functions', *Nature,* 544(7648), pp. 65-+.

Nieduszynski, I. and Marchessault, R. H. (1971) 'STRUCTURE OF BETA-D-(1- 4')XYLAN HYDRATE', *Nature,* 232(5305), pp. 46-+.

Nishikubo, N., Takahashi, J., Roos, A. A., Derba-Maceluch, M., Piens, K., Brumer, H., Teeri, T. T., Stalbrand, H. and Mellerowicz, E. J. (2011) 'Xyloglucan endo-Transglycosylase-Mediated Xyloglucan Rearrangements in Developing Wood of Hybrid Aspen', *Plant Physiology,* 155(1), pp. 399-413.

Nishimura, H., Kamiya, A., Nagata, T., Katahira, M. and Watanabe, T. (2018) 'Direct evidence for alpha ether linkage between lignin and carbohydrates in wood cell walls', *Scientific Reports,* 8.

Nishiyama, Y., Langan, P. and Chanzy, H. (2002) 'Crystal structure and hydrogen-bonding system in cellulose 1 beta from synchrotron X-ray and neutron fiber diffraction', *Journal of the American Chemical Society,* 124(31), pp. 9074-9082.

North, H. M., Berger, A., Saez-Aguayo, S. and Ralet, M. C. (2014) 'Understanding polysaccharide production and properties using seed coat mutants: future perspectives for the exploitation of natural variants', *Annals of Botany,* 114(6), pp. 1251-1263.

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., Vezzi, F., Delhomme, N., Giacomello, S., Alexeyenko, A., Vicedomini, R., Sahlin, K., Sherwood, E., Elfstrand, M., Gramzow, L., Holmberg, K., Hallman, J., Keech, O., Klasson, L., Koriabine, M., Kucukoglu, M., Kaller, M., Luthman, J., Lysholm, F., Niittyla, T., Olson, A., Rilakovic, N., Ritland, C., Rossello, J. A., Sena, J., Svensson, T., Talavera-Lopez, C., Theissen, G., Tuominen, H., Vanneste, K., Wu, Z. Q., Zhang, B., Zerbe, P., Arvestad, L., Bhalerao, R., Bohlmann, J., Bousquet, J., Gil, R. G., Hvidsten, T. R., de Jong, P., MacKay, J., Morgante, M., Ritland, K., Sundberg, B., Thompson, S. L., Van de Peer, Y., Andersson, B., Nilsson, O., Ingvarsson, P. K., Lundeberg, J. and Jansson, S. (2013) 'The Norway spruce genome sequence and conifer genome evolution', *Nature,* 497(7451), pp. 579-584.

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

Pan, Y. D., Birdsey, R. A., Fang, J. Y., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S. L., Rautiainen, A., Sitch, S. and Hayes, D. (2011) 'A Large and Persistent Carbon Sink in the World's Forests', *Science,* 333(6045), pp. 988-993.

Patron, N., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., Raitskin, O., Leveau, A., Farre, G., Rogers, C., Smith, A., Hibberd, J., Webb, A. A. R., Locke, J., Schornack, S., Ajioka, J., Baulcombe, D. C., Zipfel, C., Kamoun, S., Jones, J. D. G., Kuhn, H., Robatzek, S., Van Esse, H. P., Sanders, D., Oldroyd, G., Martin, C., Field, R., O'Connor, S., Fox, S., Wulff, B., Miller, B., Breakspear, A., Radhakrishnan, G., Delaux, P. M., Loque, D., Granell, A., Tissier, A., Shih, P., Brutnell, T. P., Quick, W. P., Rischer, H., Fraser, P. D., Aharoni, A., Raines, C., South, P. F., Ane, J. M., Hamberger, B. R., Langdale, J., Stougaard, J., Bouwmeester, H., Udvardi, M., Murray, J. A. H., Ntoukakis, V., Schafer, P., Denby, K., Edwards, K. J., Osbourn, A. and Haseloff, J. (2015) 'Standards for plant synthetic biology: a common syntax for exchange of DNA parts', *New Phytologist,* 208(1), pp. 13-19.

Pauly, M. and Keegstra, K. (2008) 'Cell-wall carbohydrates and their modification as a resource for biofuels', *Plant Journal,* 54(4), pp. 559-568.

Pauly, M. and Keegstra, K. (2016) 'Biosynthesis of the Plant Cell Wall Matrix Polysaccharide Xyloglucan', *Annual Review of Plant Biology, Vol 67,* 67, pp. 235-259.
Pavy, N., Pelgas, B., Laroche, J., Rigault, P., Isabel, N. and Bousquet, J. (2012) 'A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers', *Bmc Biology,* 10.
Pawar, P. M. A., Koutaniemi, S., Tenkanen, M. and Mellerowicz, E. J. (2013) 'Acetylation of woody lignocellulose: significance and regulation', *Frontiers in Plant Science,* 4.
Pawar, P. M. A., Ratke, C., Balasubramanian, V. K., Chong, S. L., Gandla, M. L., Adriasola, M., Sparrman, T., Hedenstrom, M., Szwaj, K., Derba-Maceluch, M., Gaertner, C., Mouille, G., Ezcurra, I., Tenkanen, M., Jonsson, L. J. and Mellerowicz, E. J. (2017) 'Downregulation of RWA genes in hybrid aspen affects xylan acetylation and wood saccharification', *New Phytologist,* 214(4), pp. 1491-1505.
Pellerin, P., Doco, T., Vidal, S., Williams, P., Brillouet, J. M. and Oneill, M. A. (1996) 'Structural characterization of red wine rhamnogalacturonan II', *Carbohydrate Research,* 290(2), pp. 183-197.
Pena, M. J., Kulkarni, A. R., Backe, J., Boyd, M., O'Neill, M. A. and York, W. S. (2016) 'Structural diversity of xylans in the cell walls of monocots', *Planta,* 244(3), pp. 589-606.
Pena, M. J., Zhong, R. Q., Zhou, G. K., Richardson, E. A., O'Neill, M. A., Darvill, A. G., York, W. S. and Ye, Z. H. (2007) 'Arabidopsis irregular xylem8 and irregular xylem9: Implications for the complexity of glucuronoxylan biosynthesis', *Plant Cell,* 19(2), pp. 549-563.
Pereira, C. S., Silveira, R. L., Dupree, P. and Skaf, M. S. (2017) 'Effects of Xylan Side-Chain Substitutions on Xylan-Cellulose Interactions and Implications for Thermal Pretreatment of Cellulosic Biomass', *Biomacromolecules,* 18(4), pp. 1311-1321.
Perrin, R. M., DeRocher, A. E., Bar-Peled, M., Zeng, W. Q., Norambuena, L., Orellana, A., Raikhel, N. V. and Keegstra, K. (1999) 'Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis', *Science,* 284(5422), pp. 1976-1979.
Petersen, P. D., Lau, J., Ebert, B., Yang, F., Verhertbruggen, Y., Kim, J. S., Varanasi, P., Suttangkakul, A., Auer, M., Loque, D. and Scheller, H. V. (2012) 'Engineering of plants with improved properties as biofuels feedstocks by vessel-specific complementation of xylan biosynthesis mutants', *Biotechnology for Biofuels,* 5.
Phyo, P., Wang, T., Yang, Y., O'Neil, H. and Hong, M. (2018) 'Direct Determination of Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using H-1 Polarization Transfer Solid-State NMR', *Biomacromolecules,* 19(5), pp. 1485-1497.
Popper, Z. A. (2008) 'Evolution and diversity of green plant cell walls', *Current Opinion in Plant Biology,* 11(3), pp. 286-292.
Poutanen, K., Ratto, M., Puls, J. and Viikari, L. (1987) 'Evaluation of different microbial xylanolytic systems', *Journal of Biotechnology,* 6(1), pp. 49-60.
Proost, S., Van Bel, M., Vaneechoutte, D., Van de Peer, Y., Inze, D., Mueller-Roeber, B. and Vandepoele, K. (2015) 'PLAZA 3.0: an access point for plant comparative genomics', *Nucleic Acids Research,* 43(D1), pp. D974-D981.
Ralet, M.-C., Crepeau, M.-J., Vigouroux, J., Tran, J., Berger, A., Salle, C., Granier, F., Botran, L. and North, H. M. (2016) 'Xylans Provide the Structural Driving Force for Mucilage Adhesion to the Arabidopsis Seed Coat', *Plant Physiology,* 171(1), pp. 165-178.
Ralph, J., Grabber, J. H. and Hatfield, R. D. (1995) ' Lignin-ferulate cross-links in grasses - active incorporation of ferulate polysaccharide esters into ryegrass lignins ', *Carbohydrate Research,* 275(1), pp. 167-178.
Ramage, M. H., Burridge, H., Busse-Wicher, M., Fereday, G., Reynolds, T., Shah, D. U., Wu, G. L., Yu, L., Fleming, P., Densley-Tingley, D., Allwood, J., Dupree, P., Linden, P. F. and Scherman, O. (2017) 'The wood from the trees: The use of timber in construction', *Renewable & Sustainable Energy Reviews,* 68, pp. 333-359.

Ramirez, V., Guangyan, X., Mashigushi, K., Yamaguchi, S. and Pauly, M. (2018) 'Growth- and stress related defects associated to wall hypoacetylation are strigolactone dependent.', *Plant Direct,* 2, pp. 1-11.

Razeq, F. M., Jurak, E., Stogios, P. J., Yan, R. Y., Tenkanen, M., Kabel, M. A., Wang, W. J. and Master, E. R. (2018) 'A novel acetyl xylan esterase enabling complete deacetylation of substituted xylans', *Biotechnology for Biofuels,* 11.

Reed, J., Stephenson, M. J., Miettinen, K., Brouwer, B., Leveau, A., Brett, P., Goss, R. J. M., Goossens, A., O'Connell, M. A. and Osbourn, A. (2017) 'A translational synthetic biology platform for rapid access to gram-scale quantities of novel drug-like molecules', *Metabolic Engineering,* 42, pp. 185-193.

Reis, D. and Vian, B. (2004) 'Helicoidal pattern in secondary cell walls and possible role of xylans in their construction', *Comptes Rendus Biologies,* 327(9-10), pp. 785-790.

Ren, Y. F., Hansen, S. F., Ebert, B., Lau, J. and Scheller, H. V. (2014) 'Site-Directed Mutagenesis of IRX9, IRX9L and IRX14 Proteins Involved in Xylan Biosynthesis: Glycosyltransferase Activity Is Not Required for IRX9 Function in Arabidopsis', *Plos One,* 9(8).

Rennie, E. A., Hansen, S. F., Baidoo, E. E. K., Hadi, M. Z., Keasling, J. D. and Scheller, H. V. (2012) 'Three Members of the Arabidopsis Glycosyltransferase Family 8 Are Xylan Glucuronosyltransferases', *Plant Physiology,* 159(4), pp. 1408-1417.

Richmond, T. A. and Somerville, C. R. (2000) 'The cellulose synthase superfamily', *Plant Physiology,* 124(2), pp. 495-498.

Ridley, B. L., O'Neill, M. A. and Mohnen, D. A. (2001) 'Pectins: structure, biosynthesis, and oligogalacturonide-related signaling', *Phytochemistry,* 57(6), pp. 929-967.

Rogowski, A., Basle, A., Farinas, C. S., Solovyova, A., Mortimer, J. C., Dupree, P., Gilbert, H. J. and Bolam, D. N. (2014) 'Evidence That GH115 alpha-Glucuronidase Activity, Which Is Required to Degrade Plant Biomass, Is Dependent on Conformational Flexibility', *Journal of Biological Chemistry,* 289(1), pp. 53-64.

Ryser, U., Schorderet, M., Guyot, R. and Keller, B. (2004) 'A new structural element containing glycine-rich proteins and rhamnogalacturonan I in the protoxylem of seed plants', *Journal of Cell Science,* 117(7), pp. 1179-1190.

Sainsbury, F., Thuenemann, E. C. and Lomonossoff, G. P. (2009) 'pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants', *Plant Biotechnology Journal,* 7(7), pp. 682-693.

Scheller, H. V. and Ulvskov, P. (2010) 'Hemicelluloses', *Annual Review of Plant Biology, Vol 61,* 61, pp. 263-289.

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D. and Lohmann, J. U. (2005) 'A gene expression map of Arabidopsis thaliana development', *Nature Genetics,* 37(5), pp. 501-506.

Schroder, R., Nicolas, P., Vincent, S. J. F., Fischer, M., Reymond, S. and Redgwell, R. J. (2001) 'Purification and characterisation of a galactoglucomannan from kiwifruit (Actinidia deliciosa)', *Carbohydrate Research,* 331(3), pp. 291-306.

Schultink, A., Naylor, D., Dama, M. and Pauly, M. (2015) 'The Role of the Plant-Specific ALTERED XYLOGLUCAN9 Protein in Arabidopsis Cell Wall Polysaccharide O-Acetylation', *Plant Physiology,* 167(4), pp. 1271-U243.

Schweingruber, F. H. (2007) *Wood structure and environment.* Berlin: Springer.

Sczostak, A. (2009) 'Cotton Linters: An Alternative Cellulosic Raw Material', *Macromolecular Symposia,* 280, pp. 45-53.

Sedjo, R. (2001) 'Biotechnology's Potential Contribution to Global Wood Supply and Forest Conservation ', Available: Resources for the Future.

Seymour, G. B., Colquhoun, I. J., Dupont, M. S., Parsley, K. R. and Selvendran, R. R. (1990) ' Composition and structural features of cell-wall polysaccharides from tomato fruits ', *Phytochemistry,* 29(3), pp. 725-731.

Shatalov, A. A., Evtuguin, D. V. and Neto, C. P. (1999) '(2-O-alpha-D-galactopyranosyl-4-O-methyl-alpha-D-glucurono)-D-xylan from Eucalyptus globulus Labill', *Carbohydrate Research,* 320(1-2), pp. 93-99.

Shimada, T. L., Shimada, T. and Hara-Nishimura, I. (2010) 'A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana', *Plant Journal,* 61(3), pp. 519-528.

Shimizu, K., Hashi, M. and Sakurai, K. (1978) ' Isolation from a softwood xylan of oligosaccharides containing 2 4-o-methyl-d-glucuronic acid residues ', *Carbohydrate Research,* 62(1), pp. 117-126.

Silveira, M. H. L., Morais, A. R. C., Lopes, A. M. D., Olekszyszen, D. N., Bogel-Lukasik, R., Andreaus, J. and Ramos, L. P. (2015) 'Current Pretreatment Technologies for the Development of Cellulosic Ethanol and Biorefineries', *Chemsuschem,* 8(20), pp. 3366-3390.

Simmons, T. J., Mortimer, J. C., Bernardinelli, O. D., Poppler, A. C., Brown, S. P., Deazevedo, E. R., Dupree, R. and Dupree, P. (2016) 'Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR', *Nature Communications,* 7.

Sjostrom, E. (2013) *Wood Chemistry - Fundamentals and Applications.* 2 edn. Cambridge , Massachusetts, USA: Academic Press.

Sjöström, E. and Westermark, U. (1999) 'Chemical Composition of Wood and Pulps: Basic Constituents and Their Distribution', in Sjöström, E. & Alén, R. (eds.) *Analytical Methods in Wood Chemistry, Pulping, and Papermaking. Springer Series in Wood Science*. Berlin and Heidelberg, Germany: Springer.

Somerville, C. (2006) 'Cellulose synthesis in higher plants', *Annual Review of Cell and Developmental Biology,* 22, pp. 53-78.

Sorensen, I., Domozych, D. and Willats, W. G. T. (2010) 'How Have Plant Cell Walls Evolved?', *Plant Physiology,* 153(2), pp. 366-372.

Spanikova, S. and Biely, P. (2006) 'Glucuronoyl esterase - Novel carbohydrate esterase produced by Schizophyllum commune', *Febs Letters,* 580(19), pp. 4597-4601.

Sparkes, I. A., Runions, J., Kearns, A. and Hawes, C. (2006) 'Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants', *Nature Protocols,* 1(4), pp. 2019-2025.

Sun, G., Dilcher, D. L., Wang, H. S. and Chen, Z. D. (2011) 'A eudicot from the Early Cretaceous of China', *Nature,* 471(7340), pp. 625-628.

Sundberg, A., Sundberg, K., Lillandt, C. and Holmbom, B. (1996) 'Determination of hemicelluloses and pectins in wood and pulp fibres by acid methanolysis and gas chromatography ', *Nordic Pulp and Paper Research Journal,* 226(4), pp. 216-219.

Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y. C., Sjodin, A., Van de Peer, Y., Jansson, S., Hvidsten, T. R. and Street, N. R. (2015) 'The Plant Genome Integrative Explorer Resource: PlantGenIE.org', *New Phytologist,* 208(4), pp. 1149-1156.

Takahashi, N. and Koshijima, T. (1988) ' Ester linkages between lignin and glucuronoxylan in a lignin-carbohydrate complex from beech (Fagus-crenata) wood ', *Wood Science and Technology,* 22(3), pp. 231-241.

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

Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C. H., Hao, Z. Y., Zhu, X., Avci, U., Miller, J. S., Baldwin, D., Pham, C., Orlando, R., Darvill, A., Hahn, M. G., Kieliszewski, M. J. and Mohnen, D. (2013) 'An Arabidopsis Cell Wall Proteoglycan Consists of Pectin and Arabinoxylan Covalently Linked to an Arabinogalactan Protein', *Plant Cell,* 25(1), pp. 270-287.

Tang, W. and Newton, R. J. (2003) 'Genetic transformation of conifers and its application in forest biotechnology', *Plant Cell Reports,* 22(1), pp. 1-15.

Taylor, E. J., Gloster, T. M., Turkenburg, J. P., Vincent, F., Brzozowski, A. M., Dupont, C., Shareck, F., Centeno, M. S. J., Prates, J. A. M., Puchart, V., Ferreira, L. M. A., Fontes, C., Biely, P. and Davies, G. J. (2006) 'Structure and activity of two metal ion-dependent acetylxylan esterases involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases', *Journal of Biological Chemistry,* 281(16), pp. 10968-10975.

Taylor, N. G., Howells, R. M., Huttly, A. K., Vickers, K. and Turner, S. R. (2003) 'Interactions among three distinct CesA proteins essential for cellulose synthesis', *Proceedings of the National Academy of Sciences of the United States of America,* 100(3), pp. 1450-1455.

Tenkanen, M. and Siika-aho, M. (2000) 'An alpha-glucuronidase of Schizophyllum commune acting on polymeric xylan', *Journal of Biotechnology,* 78(2), pp. 149-161.

Terashima, N., Kitano, K., Kojima, M., Yoshida, M., Yamamoto, H. and Westermark, U. (2009) 'Nanostructural assembly of cellulose, hemicellulose, and lignin in the middle layer of secondary wall of ginkgo tracheid', *Journal of Wood Science,* 55(6), pp. 409-416.

Terrett, O. M., Lyczakowski, J. J., Yu, L., Dupree, R. and Dupree, P. in preparation. Galactoglucomannan and xylan bind to the surface of cellulose microfibrils in spruce wood.

Thomas, L. H., Altaner, C. M. and Jarvis, M. C. (2013) 'Identifying multiple forms of lateral disorder in cellulose fibres', *Journal of Applied Crystallography,* 46, pp. 972-979.

Thuenemann, E. C., Meyers, A. E., Verwey, J., Rybicki, E. P. and Lomonossoff, G. P. (2013) 'A method for rapid production of heteromultimeric protein complexes in plants: assembly of protective bluetongue virus-like particles', *Plant Biotechnology Journal,* 11(7), pp. 839-846.

Timell, T. E. and Syracuse, N. Y. (1967) 'Recent progress in the chemistry of wood hemicelluloses', *Wood Science and Technology,* 1(1).

Tryfona, T., Liang, H. C., Kotake, T., Tsumuraya, Y., Stephens, E. and Dupree, P. (2012) 'Structural Characterization of Arabidopsis Leaf Arabinogalactan Polysaccharides', *Plant Physiology,* 160(2), pp. 653-666.

Turner, S. R. and Somerville, C. R. (1997) 'Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall', *Plant Cell,* 9(5), pp. 689-701.

Urbanikova, L., Vrsanska, M., Krogh, K., Hoff, T. and Biely, P. (2011) 'Structural basis for substrate recognition by Erwinia chrysanthemi GH30 glucuronoxylanase', *Febs Journal,* 278(12), pp. 2105-2116.

Urbanowicz, B. R., Pena, M. J., Moniz, H. A., Moremen, K. W. and York, W. S. (2014) 'Two Arabidopsis proteins synthesize acetylated xylan in vitro', *Plant Journal,* 80(2), pp. 197-206.

Urbanowicz, B. R., Pena, M. J., Ratnaparkhe, S., Avci, U., Backe, J., Steet, H. F., Foston, M., Li, H., O'Neill, M. A., Ragauskas, A. J., Darvill, A. G., Wyman, C., Gilbert, H. J. and York, W. S. (2012) '4-O-methylation of glucuronic acid in Arabidopsis glucuronoxylan is catalyzed by a domain of unknown function family 579 protein', *Proceedings of the National Academy of Sciences of the United States of America,* 109(35), pp. 14253-14258.

Vain, T., Crowell, E. F., Timpano, H., Biot, E., Desprez, T., Mansoori, N., Trindade, L. M., Pagant, S., Robert, S., Hofte, H., Gonneau, M. and Vernhettes, S. (2014) 'The Cellulase KORRIGAN Is Part of the Cellulose Synthase Complex', *Plant Physiology,* 165(4), pp. 1521-1532.

Van Acker, R., Leple, J. C., Aerts, D., Storme, V., Goeminne, G., Ivens, B., Legee, F., Lapierre, C., Piens, K., Van Montagu, M. C. E., Santoro, N., Foster, C. E., Ralph, J., Soetaert, W., Pilate, G. and Boerjan, W. (2014) 'Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase', *Proceedings of the National Academy of Sciences of the United States of America,* 111(2), pp. 845-850.

Van Acker, R., Vanholme, R., Storme, V., Mortimer, J. C., Dupree, P. and Boerjan, W. (2013) 'Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in Arabidopsis thaliana', *Biotechnology for Biofuels,* 6.

Vanholme, R., Demedts, B., Morreel, K., Ralph, J. and Boerjan, W. (2010) 'Lignin Biosynthesis and Structure', *Plant Physiology,* 153(3), pp. 895-905.

Vincken, J. P., Schols, H. A., Oomen, R., McCann, M. C., Ulvskov, P., Voragen, A. G. J. and Visser, R. G. F. (2003) 'If homogalacturonan were a side chain of rhamnogalacturonan I. Implications for cell wall architecture', *Plant Physiology,* 132(4), pp. 1781-1789.

Voiniciuc, C., Dean, G. H., Griffiths, J. S., Kirchsteiger, K., Hwang, Y. T., Gillett, A., Dow, G., Western, T. L., Estelle, M. and Haughn, G. W. (2013) 'FLYING SAUCER1 Is a Transmembrane RING E3 Ubiquitin Ligase That Regulates the Degree of Pectin Methylesterification in Arabidopsis Seed Mucilage', *Plant Cell,* 25(3), pp. 944-959.

Voiniciuc, C., Gunl, M., Schmidt, M. H. W. and Usadel, B. (2015a) 'Highly Branched Xylan Made by IRREGULAR XYLEM14 and MUCILAGE-RELATED21 Links Mucilage to Arabidopsis Seeds', *Plant Physiology,* 169(4), pp. 2481-2495.

Voiniciuc, C., Schmidt, M. H. W., Berger, A., Yang, B., Ebert, B., Scheller, H. V., North, H. M., Usadel, B. and Gunl, M. (2015b) 'MUCILAGE-RELATED10 Produces Galactoglucomannan That Maintains Pectin and Cellulose Architecture in Arabidopsis Seed Mucilage', *Plant Physiology,* 169(1), pp. 403-+.

Voragen, A. G. J., Coenen, G.-J., Verhoef, R. P. and Schols, H. A. (2009) 'Pectin, a versatile polysaccharide present in plant cell walls', *Structural Chemistry,* 20(2), pp. 263-275.

Wagner, A., Tobimatsu, Y., Goeminne, G., Phillips, L., Flint, H., Steward, D., Torr, K., Donaldson, L., Boerjan, W. and Ralph, J. (2013) 'Suppression of CCR impacts metabolite profile and cell wall composition in Pinus radiata tracheary elements', *Plant Molecular Biology,* 81(1-2), pp. 105-117.

Watanabe, T. and Koshijima, T. (1988) ' Evidence for an ester linkage between lignin and glucuronic-acid in lignin carbohydrate complexes by ddq-oxidation ', *Agricultural and Biological Chemistry,* 52(11), pp. 2953-2955.

Weber, E., Engler, C., Gruetzner, R., Werner, S. and Marillonnet, S. (2011) 'A Modular Cloning System for Standardized Assembly of Multigene Constructs', *Plos One,* 6(2).

Western, T. L. (2012) 'The sticky tale of seed coat mucilages: production, genetics, and role in seed germination and dispersal', *Seed Science Research,* 22(1), pp. 1-25.

Whitehead, D. (2011) 'Forests as carbon sinks-benefits and consequences', *Tree Physiology,* 31(9), pp. 893-902.

Whitney, S. E. C., Brigham, J. E., Darke, A. H., Reid, J. S. G. and Gidley, M. J. (1998) 'Structural aspects of the interaction of mannan-based polysaccharides with bacterial cellulose', *Carbohydrate Research,* 307(3-4), pp. 299-309.

Wilkerson, C. G., Mansfield, S. D., Lu, F., Withers, S., Park, J. Y., Karlen, S. D., Gonzales-Vigil, E., Padmakshan, D., Unda, F., Rencoret, J. and Ralph, J. (2014) 'Monolignol Ferulate Transferase Introduces Chemically Labile Linkages into the Lignin Backbone', *Science,* 344(6179), pp. 90-93.

Willfor, S., Sundberg, A., Hemming, J. and Holmbom, B. (2005a) 'Polysaccharides in some industrially important softwood species', *Wood Science and Technology,* 39(4), pp. 245-258.

Willfor, S., Sundberg, A., Pranovich, A. and Holmbom, B. (2005b) 'Polysaccharides in some industrially important hardwood species', *Wood Science and Technology,* 39(8), pp. 601-617.

Willfor, S., Sundberg, K., Tenkanen, M. and Holmbom, B. (2008) 'Spruce-derived mannans - A potential raw material for hydrocolloids and novel advanced natural materials', *Carbohydrate Polymers,* 72(2), pp. 197-210.

Wu, A.-M., Hornblad, E., Voxeur, A., Gerber, L., Rihouey, C., Lerouge, P. and Marchant, A. (2010) 'Analysis of the Arabidopsis IRX9/IRX9-L and IRX14/IRX14-L Pairs of Glycosyltransferase Genes Reveals Critical Contributions to Biosynthesis of the Hemicellulose Glucuronoxylan', *Plant Physiology,* 153(2), pp. 542-554.

Xiong, G., Dama, M. and Pauly, M. (2015) 'Glucuronic Acid Moieties on Xylan Are Functionally Equivalent to O-Acetyl-Substituents', *Molecular Plant,* 8(7), pp. 1119-1121.

Xiong, G. Y., Cheng, K. and Pauly, M. (2013) 'Xylan O-Acetylation Impacts Xylem Development and Enzymatic Recalcitrance as Indicated by the Arabidopsis Mutant tbl29', *Molecular Plant,* 6(4), pp. 1373-1375.

Yamasaki, T., Enomoto, A., Kato, A., Ishii, T. and Shimizu, K. (2011) 'Structural unit of xylans from sugi (Cryptomeria japonica) and hinoki (Chamaecyparis obtusa)', *Journal of Wood Science,* 57(1), pp. 76-84.

Yang, W. B., Schuster, C., Beahan, C. T., Charoensawan, V., Peaucelle, A., Bacic, A., Doblin, M. S., Wightman, R. and Meyerowitz, E. M. (2016) 'Regulation of Meristem Morphogenesis by Cell Wall Synthases in Arabidopsis', *Current Biology,* 26(11), pp. 1404-1415.

Yu, L., Lyczakowski, J. J., Pereira, C. S., Kotake, T., Yu, X., Li, A., Mogelsvang, S., Skaf, M. S. and Dupree, P. (2018) 'The patterned structure of galactoglucomannan suggests it may bind to cellulose in seed mucilage', *Plant Physiology,* in press.

Yu, L., Shi, D. C., Li, J. L., Kong, Y. Z., Yu, Y. C., Chai, G. H., Hu, R. B., Wang, J., Hahn, M. G. and Zhou, G. K. (2014) 'CELLULOSE SYNTHASE-LIKE A2, a Glucomannan Synthase, Is Involved in Maintaining Adherent Mucilage Structure in Arabidopsis Seed(1 C W)', *Plant Physiology,* 164(4), pp. 1842-1856.

Yuan, Y. X., Teng, Q., Zhong, R. Q., Haghighat, M., Richardson, E. A. and Ye, Z. H. (2016) 'Mutations of Arabidopsis TBL32 and TBL33 Affect Xylan Acetylation and Secondary Wall Deposition', *Plos One,* 11(1).

Yuan, Y. X., Teng, Q., Zhong, R. Q. and Ye, Z. H. (2013) 'The Arabidopsis DUF231 Domain-Containing Protein ESK1 Mediates 2-O- and 3-O-Acetylation of Xylosyl Residues in Xylan', *Plant and Cell Physiology,* 54(7), pp. 1186-1199.

Zandleven, J., Sorensen, S. O., Harholt, J., Beldman, G., Schols, H. A., Scheller, H. V. and Voragen, A. J. (2007) 'Xylogalacturonan exists in cell walls from various tissues of Arabidopsis thaliana', *Phytochemistry,* 68(8), pp. 1219-1226.

Zeng, W., Lampugnani, E. R., Picard, K. L., Song, L. L., Wu, A. M., Farion, I. M., Zhao, J., Ford, K., Doblin, M. S. and Bacic, A. (2016) 'Asparagus IRX9, IRX10, and IRX14A Are Components of an Active Xylan Backbone Synthase Complex that Forms in the Golgi Apparatus', *Plant Physiology,* 171(1), pp. 93-109.

Zhang, B. C., Zhang, L. J., Li, F., Zhang, D. M., Liu, X. L., Wang, H., Xu, Z. P., Chu, C. C. and Zhou, Y. H. (2017) 'Control of secondary cell wall patterning involves xylan deacetylation', *Nature Plants,* 3(3).

Zhang, J., Nieminen, K., Serra, J. A. A. and Helariutta, Y. (2014) 'The formation of wood and its control', *Current Opinion in Plant Biology,* 17, pp. 56-63.

Zhang, Y., Nikolovski, N., Sorieul, M., Vellosillo, T., McFarlane, H. E., Dupree, R., Kesten, C., Schneider, R., Driemeier, C., Lathe, R., Lampugnani, E., Yu, X. L., Ivakov, A., Doblin, M. S., Mortimer, J. C., Brown, S. P., Persson, S. and Dupree, P. (2016) 'Golgi-localized STELLO proteins regulate the assembly and trafficking of cellulose synthase complexes in Arabidopsis', *Nature Communications,* 7.

Zhao, Q., Nakashima, J., Chen, F., Yin, Y. B., Fu, C. X., Yun, J. F., Shao, H., Wang, X. Q., Wang, Z. Y. and Dixon, R. A. (2013) 'LACCASE Is Necessary and Nonredundant with PEROXIDASE for Lignin Polymerization during Vascular Development in Arabidopsis', *Plant Cell,* 25(10), pp. 3976-3987.

Zhong, R. Q., Cui, D. T., Phillips, D. R. and Ye, Z. H. (2018) 'A Novel Rice Xylosyltransferase Catalyzes the Addition of 2-O-Xylosyl Side Chains onto the Xylan Backbone', *Plant and Cell Physiology,* 59(3), pp. 554-565.

Zhong, R. Q., Cui, D. T. and Ye, Z. H. (2017) 'Regiospecific Acetylation of Xylan is Mediated by a Group of DUF231-Containing O-Acetyltransferases', *Plant and Cell Physiology,* 58(12), pp. 2126-2138.

Zhou, S. F., Runge, T., Karlen, S. D., Ralph, J., Gonzales-Vigil, E. and Mansfield, S. D. (2017) 'Chemical Pulping Advantages of Zip-lignin Hybrid Poplar', *Chemsuschem,* 10(18), pp. 3565-3573.

Zhou, X. H., Jacobs, T. B., Xue, L. J., Harding, S. A. and Tsai, C. J. (2015) 'Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate: CoA ligase specificity and redundancy', *New Phytologist,* 208(2), pp. 298-301.

Zhou, Z.-Y. (2009) 'An overview of fossil Ginkgoales', *Palaeoworld,* 18(1), pp. 1-22.

Zimin, A., Stevens, K. A., Crepeau, M., Holtz-Morris, A., Koriabine, M., Marcais, G., Puiu, D., Roberts, M., Wegrzyn, J. L., de Jong, P. J., Neale, D. B., Salzberg, S. L., Yorke, J. A. and Langley, C. H. (2014) 'Sequencing and Assembly of the 22-Gb Loblolly Pine Genome', *Genetics,* 196(3), pp. 875-890.

Appendix – Peer-reviewed journal publications

Some of the results and analyses presented in this thesis were published in the following journal article:

Lyczakowski JJ., Wicher K., Terrett OM., Faria-Blanc N., Yu X., Brown D., Krogh KBRM., Dupree P., Busse-Wicher M. **Removal of glucuronic acid from xylan is a strategy to improve the conversion of plant biomass to sugars for bioenergy.** *Biotechnology for Biofuels***.** 2017 10:224

This article is attached to the thesis in this Appendix.

In addition to that, work associated with experiments performed in this thesis formed part of the following journal articles:

Yu L., Lyczakowski JJ., Pereira CS., Kotake T., Yu X., Li A., Mogelsvang S., Skaf MS., Dupree P. **The patterned structure of galactoglucomannan suggests it may bind to cellulose in seed mucilage.** *Plant Physiology* 2018 in press

Grantham NJ., Wurman-Rodrich J., Terrett OM., Lyczakowski JJ., Stott K., Iuga D., Simmons TJ., Durand-Tardif M., Brown SP., Dupree R., Dupree P. **An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls.** *Nature Plants* 2017 3:859-865

Busse-Wicher M., Granthem NJ., Lyczakowski JJ., Nikolovski N., Dupree P. **Xylan decoration patterns and the plant secondary cell wall molecular architecture.** *Biochemical Society Transactions* 2016 44:1

Lyczakowski *et al. Biotechnol Biofuels (2017) 10:224*
DOI 10.1186/s13068-017-0902-1

Biotechnology for Biofuels

RESEARCH

Open Access

Removal of glucuronic acid from xylan is a strategy to improve the conversion of plant biomass to sugars for bioenergy

Jan J. Lyczakowski[1,3,4], Krzysztof B. Wicher[2,6], Oliver M. Terrett[1], Nuno Faria-Blanc[1], Xiaolan Yu[1], David Brown[1,7], Kristian B. R. M. Krogh[5], Paul Dupree[1,3,4*] and Marta Busse-Wicher[1,3,4*]

Abstract

Background: Plant lignocellulosic biomass can be a source of fermentable sugars for the production of second generation biofuels and biochemicals. The recalcitrance of this plant material is one of the major obstacles in its conversion into sugars. Biomass is primarily composed of secondary cell walls, which is made of cellulose, hemicelluloses and lignin. Xylan, a hemicellulose, binds to the cellulose microfibril and is hypothesised to form an interface between lignin and cellulose. Both softwood and hardwood xylan carry glucuronic acid side branches. As xylan branching may be important for biomass recalcitrance and softwood is an abundant, non-food competing, source of biomass it is important to investigate how conifer xylan is synthesised.

Results: Here, we show using Arabidopsis *gux* mutant biomass that removal of glucuronosyl substitutions of xylan can allow 30% more glucose and over 700% more xylose to be released during saccharification. Ethanol yields obtained through enzymatic saccharification and fermentation of *gux* biomass were double those obtained for non-mutant material. Our analysis of additional xylan branching mutants demonstrates that absence of GlcA is unique in conferring the reduced recalcitrance phenotype. As in hardwoods, conifer xylan is branched with GlcA. We use transcriptomic analysis to identify conifer enzymes that might be responsible for addition of GlcA branches onto xylan in industrially important softwood. Using a combination of in vitro and in vivo activity assays, we demonstrate that a white spruce (*Picea glauca*) gene, *PgGUX*, encodes an active glucuronosyl transferase. Glucuronic acid introduced by PgGUX reduces the sugar release of Arabidopsis *gux* mutant biomass to wild-type levels indicating that it can fulfil the same biological function as native glucuronosylation.

Conclusion: Removal of glucuronic acid from xylan results in the largest increase in release of fermentable sugars from Arabidopsis plants that grow to the wild-type size. Additionally, plant material used in this work did not undergo any chemical pretreatment, and thus increased monosaccharide release from *gux* biomass can be achieved without the use of environmentally hazardous chemical pretreatment procedures. Therefore, the identification of a gymnosperm enzyme, likely to be responsible for softwood xylan glucuronosylation, provides a mutagenesis target for genetically improved forestry trees.

Keywords: Biofuels, Xylan, Glucuronic acid, Conifers, Softwood, GUX

*Correspondence: pd101@cam.ac.uk; mnb29@cam.ac.uk
[1] Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK
Full list of author information is available at the end of the article

Background

The growing population demands that plant biomass use becomes as efficient as possible, especially in large-scale applications: as a food and feed resource, and production of lignocellulosic biofuels and renewable materials [1–3]. Increasing the yield of sugars from both hemicelluloses and cellulose in the cell wall is important for the development of economic biorefineries and for use of improved plant biomass as an animal feed.

The intricate assembly and cross linking of lignin and polysaccharides within the cell wall renders the polysaccharides largely inaccessible to degradation [4]. Biomass digestion can be achieved by the use of expensive acid or alkali pretreatment, steam explosion or organic solvents [5], but these are not yet widely commercially viable processes. Consequently, there is considerable effort to increase the yield of sugar from biomass by breeding improved biomass crops, advancing pretreatment processes and improving enzyme cocktails. Despite substantial advances in all these areas, the lack of understanding of the molecular basis of recalcitrance prevents a targeted approach to improvement of saccharification. Cellulose is naturally resistant to enzymatic attack, but in the cell wall it is protected by hemicelluloses and lignin, which are removed in pretreatment processes to allow enzymatic saccharification of the cellulose [6]. Experiments using genetically modified plants and studies of genetic diversity have implicated lignin as one of the main cell wall components that influence digestibility [7–9].

In this work, we focussed on xylan, the major hemicellulose in secondary cell walls of eudicot angiosperms such as poplar, and an important hemicellulose in conifer cell walls, constituting up to 30 and 15% of dry material, respectively [10]. It is built of a β (1, 4)-linked xylose backbone that carries acetyl and [methyl]glucuronic acid ([Me]GlcA) branches in hardwoods and arabinose and MeGlcA substitutions in softwoods. In *Arabidopsis thaliana,* a model for hardwood secondary cell walls, xylan acetylation is believed to be catalysed by several acetyltransferases, of which TBL29/ESK is responsible for transfer of over 50% of acetyl groups [11]. The addition of α-1–2 linked GlcA branches to xylosyl residues is catalysed by GlucUronic acid substitution of Xylan (GUX) enzymes [12, 13]. The average degree of xylan glucuronosylation is 1 in every 8 xylosyl residues in angiosperms [14]. The frequency of [Me]GlcA in gymnosperms is higher, around 1 in every 6 xylose units [15]. GUX1/GUX2-deficient *Arabidopsis thaliana* plants (*gux1/2)* were shown to have no [Me]GlcA decorations on their secondary cell wall xylan [12]. GUX3 is responsible for glucuronosylation of xylan in primary walls [16]. GUX1, GUX2, and GUX4 enzymes show activity in vitro [13]. GlcA branches are 4-O-methylated by the activity of GXM enzymes [17, 18]. In our previous work, we demonstrated that extracted, deacetylated xylan from *gux1/2* plants can be completely digested by just two enzymes: a xylanase and a β-xylosidase [12]. This is due to simplification of this xylan substrate as xylosidases are inhibited by the presence of [Me]GlcA on wild-type xylan; complete digestion requires the additional action of α-glucuronidases [12, 19]. Commercial cocktails therefore contain xylan glucuronidases. Similarly, acetylation inhibits digestion of xylan, necessitating addition of acetyl esterases to cocktails [20].

Several attempts to modify xylan branching in vivo have had varying but limited success in improving cellulose and xylan digestibility. Reducing acetylation of xylan by about 50% caused strong dwarfing in the *tbl29*/*esk* mutant, and did not improve saccharification [11]. On the other hand, there was a small improvement in saccharification in the growth-suppressed *esk/kak* plants [21]. Similarly, partial in muro deacetylation of xylan with acetyl esterases in Arabidopsis showed modest increases in cellulose digestion [22]. Manipulation of GlcA methylation, by reducing it from 70 to 30% in the *gxm1* mutant, has also been reported to show a small increase in xylose release [17]. Attempts to remove [Me]GlcA in muro using a glucuronidase did not substantially change [Me]GlcA levels, or saccharification [23]. The *gux1/2* mutants show a small improvement in extractability of xylan with sodium hydroxide [12], but is not yet clear to what extent removal of [Me]GlcA branches could affect saccharification of cell walls.

We report here that both cellulose and xylan are much more easily digestible by the commercial enzymatic cocktail Cellic® CTec2 in non-pretreated cell wall biomass of *gux1/2* plants than those of WT plants. Simultaneous saccharification and fermentation experiments demonstrated that sugars in the *gux1/2* biomass can be more efficiently converted to ethanol. To transfer this knowledge to industrially relevant plants, we identify and characterise a conifer GUX enzyme.

Results

Biomass lacking GlcA branches on xylan has reduced recalcitrance

To investigate the effect of modifying xylan decorations on recalcitrance, we performed saccharification experiments of modified stem biomass from *A. thaliana*. We studied plants with alterations in all the secondary cell wall xylan decorations: reduced acetylation (*tbl29/esk, 11),* without methylation of GlcA (*gxm1/2/3*, [24]) and without any [Me]GlcA sugar decorations (*gux1/2,* [12]). The reduced acetylation mutant is severely dwarfed [11], but the *gxm1/2/3* and *gux1/2* mutants grow without yield penalty (Additional file 1: Figure S1, Additional

file 2: Figure S2). First, we measured sugar release from alcohol insoluble residues (AIR), a standard cell wall preparation involving milling in ethanol. Experiments used AIR without any further chemical pretreatment to avoid reducing any differences in biomass recalcitrance between plant genotypes. After enzymatic saccharification of AIR with the Cellic® CTec2 enzyme cocktail, we observed increased release of both glucose (Fig. 1a) and xylose (Fig. 1b) from the *gux1/2* biomass in comparison with wild-type (WT). Remarkably, over 700% more xylose was released from *gux1/2* AIR (Fig. 1b). On the other hand, no statistically significant difference in sugar release was observed from either *tbl29* or *gxm1/2/3* plants in these saccharification conditions. Use of a feedstock without chemical pretreatment could reduce the environmental impact of biofuel production. Therefore, we next evaluated if the saccharification phenotype is also observed for wet-milled stems without any chemical pretreatments. After saccharification of WT and *gux1/2* milled stems, we observed double the release of glucose from the mutant biomass (Fig. 1c). Similarly to AIR saccharification, xylose release from *gux1/2* milled stems was improved five-fold (Fig. 1d). Thus, the [Me]GlcA decorations are critical for recalcitrance of biomass, and have an exceptionally large impact on xylose release.

A possible explanation for the *gux* saccharification phenotype is an increased amount of xylan in the stems of these plants. Some literature sources suggest an increase in xylose measurement for *gux1/2* by monosaccharide analysis [23]. We were unable to observe the same phenotype in our previous work [12]. However, this potential increase cannot fully explain the several-fold

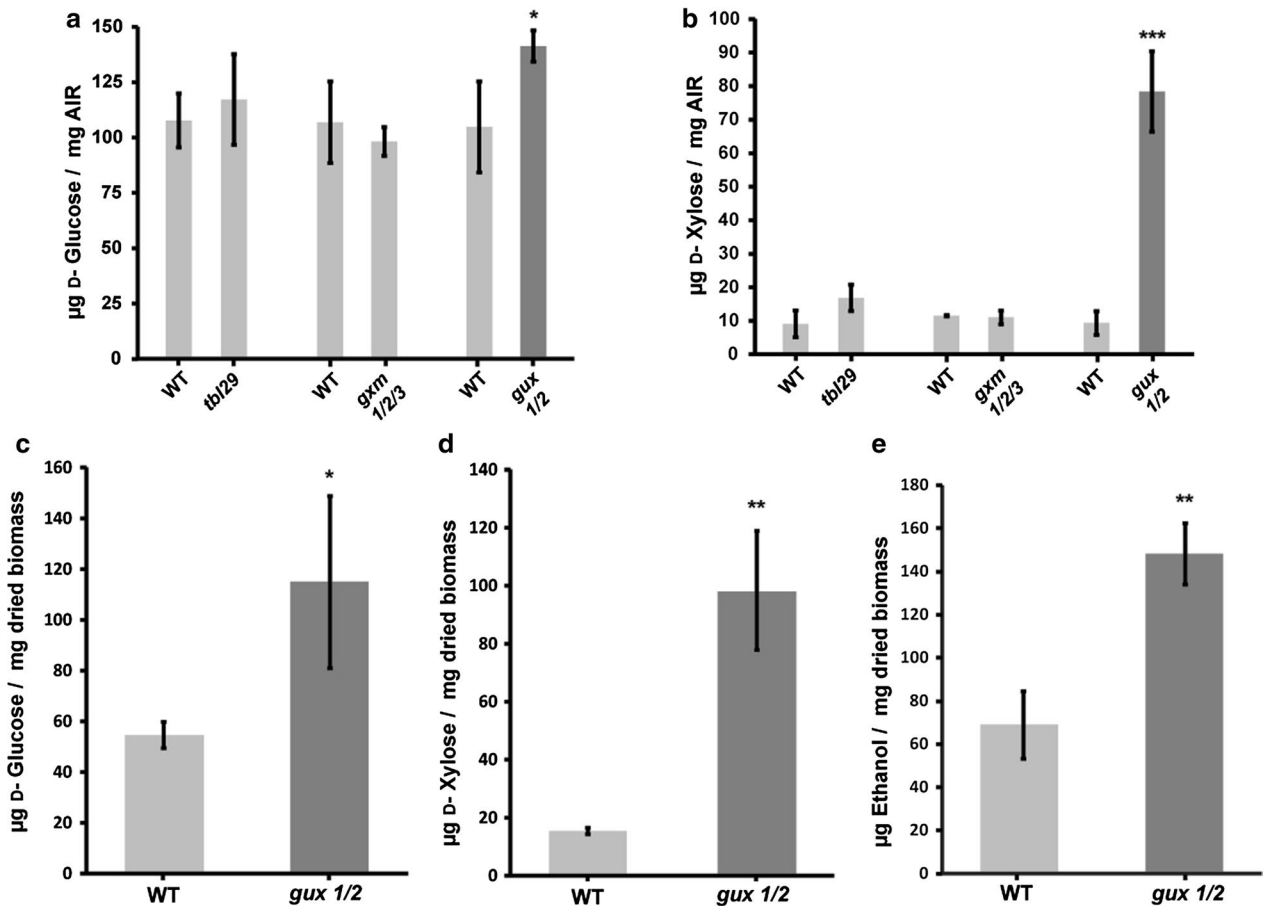


Fig. 1 Biomass lacking xylan–[Me]GlcA decorations has reduced recalcitrance and is a superior feedstock for bioethanol production. Average D-glucose (**a**) and D-xylose (**b**) release following saccharification of WT, *tbl29*, *gxm1/2/3* and *gux1/2* AIR. D-glucose (**c**) and D-xylose release (**d**) from WT and *gux1/2* milled dried stems. Ethanol concentration after 96 h of simultaneous saccharification and fermentation of WT and *gux1/2* biomass (**e**). Ethanol yields were standardised for the readings from a fermentation reaction which did not include saccharification enzymes (Additional file 7: Table S1). Error bars represent standard deviation of three matching WT and mutant biological replicates of biomass, **p* value ≤0.05; ***p* value ≤0.01; ****p* value ≤0.001

improvement in xylose yields observed in this work. Another possible explanation for increased xylose release from *gux1/2* might be insufficiency of glucuronidase activity in the enzyme mix used for saccharification. In this case, glucuronosylated oligosaccharides might accumulate in a reaction when WT feedstock is used, and such oligosaccharides would not be measured when xylose release is quantified. However, GH115 glucuronidase [19] treatment of Cellic® CTec2 digested biomass, followed by second saccharification step, did not increase xylose release (Additional file 3: Figure S3). Therefore, the major improvement in xylan digestion in *gux1/2* biomass is likely to reflect alterations to the molecular architecture of the secondary cell walls.

To test if the improved monosaccharide yield can enhance bioethanol production, we carried out simultaneous saccharification and fermentation (SSF) experiments using wet-milled stems as a feedstock and transgenic *Escherichia coli* capable of ethanol production [25]. We used an *E. coli* strain that can utilise pentoses and produce ethanol. Thus, it is well suited to metabolise xylan saccharification products. Fermentation of simultaneously saccharified *gux1/2* biomass had double the yield of ethanol compared to ethanol from WT biomass SSF (Fig. 1e).

Gymnosperm genomes encode putative GUX enzymes

Xylan in hardwoods and softwoods contains patterned [Me]GlcA branches [15]. Conservation of this decoration suggests it plays an important function in all vascular plants [15]. Therefore, removing glucuronosylation from commercially relevant plants such as trees, including conifers that produce softwoods, could facilitate processing of wood into materials and biofuels. To identify possible conifer GUX glycosyltransferases, we examined sequences from the *P. glauca* transcriptome and found a read (PgGUX, clone: GQ03239_L13, GeneBank: BT111578.1) encoding a protein with up to 73% similarity to Arabidopsis GUX enzymes. Analysis of transcriptomic data from the OneKP project [26–29] identified reads

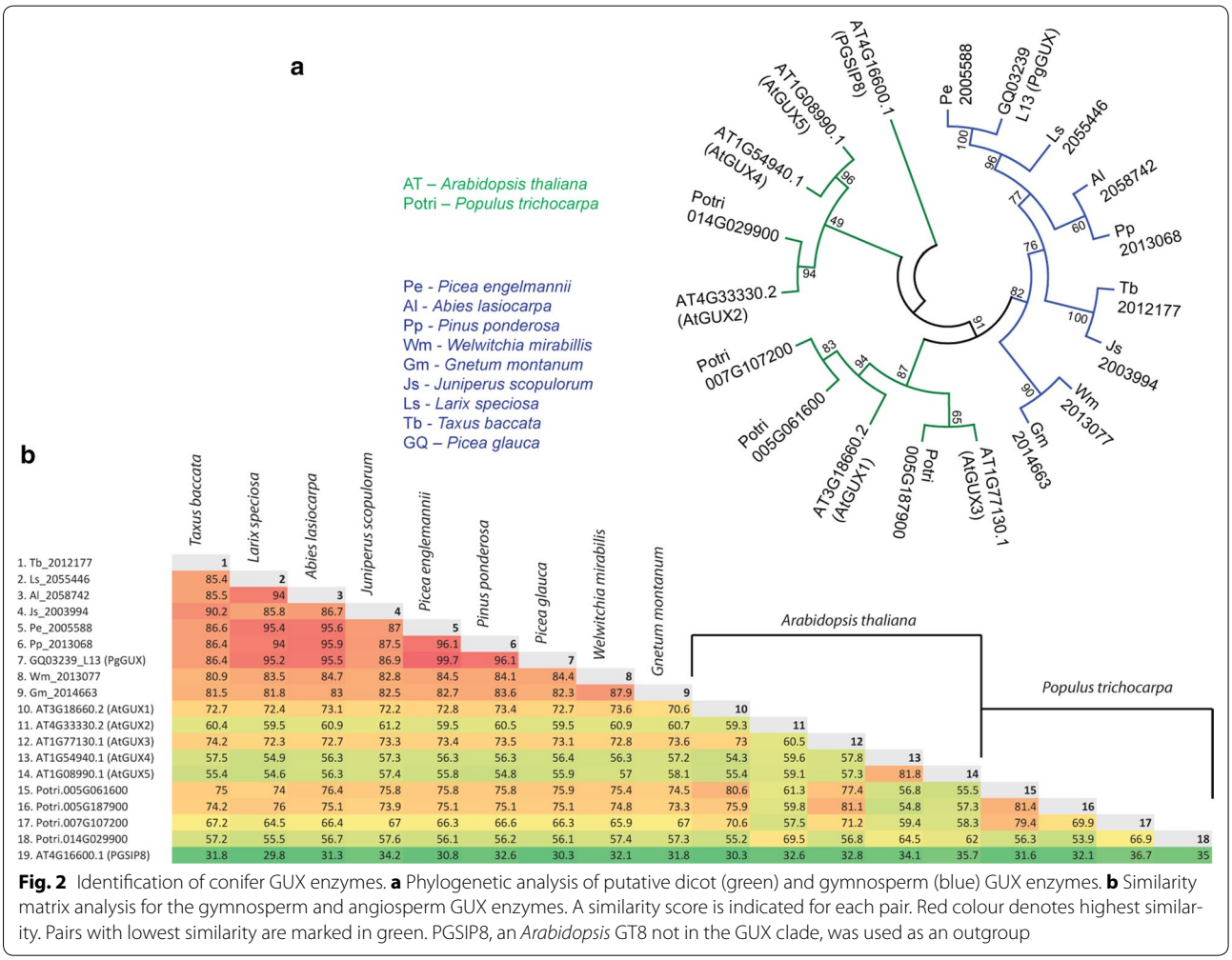


Fig. 2 Identification of conifer GUX enzymes. **a** Phylogenetic analysis of putative dicot (green) and gymnosperm (blue) GUX enzymes. **b** Similarity matrix analysis for the gymnosperm and angiosperm GUX enzymes. A similarity score is indicated for each pair. Red colour denotes highest similarity. Pairs with lowest similarity are marked in green. PGSIP8, an *Arabidopsis* GT8 not in the GUX clade, was used as an outgroup

encoding putative GUX enzymes of other gymnosperm species (Additional file 4: Table S2). A phylogenetic analysis of these newly identified gymnosperm putative GUX proteins showed that they are in a separate clade (Fig. 2a) and that they share a high degree of sequence similarity (Fig. 2b). Analysis of *Picea abies* transcriptome heat-map suggests that mRNA encoding a PgGUX homologue is enriched in wood supporting its function in xylan biosynthesis (Additional file 5: Figure S4) [30].

PgGUX can decorate acetylated xylan in vitro

To detect any PgGUX xylan glucuronosyltransferase activity, the enzyme was expressed as a *myc*-tagged protein in the tobacco *Nicotiana benthamiana* (Additional file 6: Figure S5). As a control for any endogenous tobacco glucuronosyltransferase activity, GTL6/MUCI10 (a Golgi-localised glucomannan galactosyltransferase, [31, 32]) was similarly expressed. Intact polymeric xylan from *gux1/2* lacking any [Me]GlcA decorations was used as an acceptor. Since this xylan is insoluble without acetylation, microsomes extracted from tobacco expressing PgGUX or the control GTL6 were incubated with acetylated *gux1/2* xylan and the reaction products were deacetylated and analysed by PACE using digestion with a GH11 xylanase. A GlcA-xylotetraose (UX_4) product, indicating xylan glucuronosyltransferase activity, was observed for microsomes from PgGUX overexpressing plants incubated with the xylan acceptor in the presence of UDP-GlcA (Fig. 3).

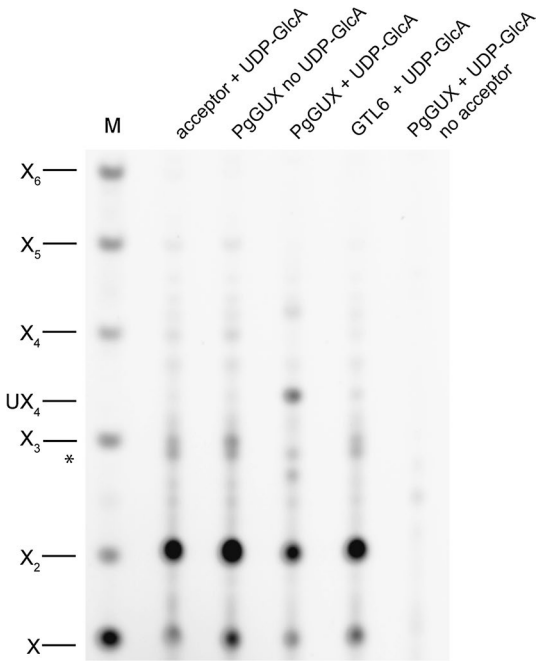


Fig. 3 PgGUX has xylan glucuronosyltransferase activity in vitro. The assay was performed with UDP-GlcA, acetylated xylan without [Me]GlcA decorations, and microsomes from *N. benthamiana* expressing PgGUX or the control GTL6 protein. Products of the in vitro glucuronosylation were digested with xylanase GH11 and analysed by PACE. The enzyme generates xylose, xylobiose, plus the UX_4 oligosaccharide if any GlcA is present on the xylan [12]. Average degree of glucuronisation was 10.9%. *Denote a background bands

PgGUX is an active conifer xylan glucuronosyltransferase in vivo

Having established that PgGUX is active on acetylated xylan in vitro, we tested whether we could use it to introduce xylan decorations in vivo. The *PgGUX* coding sequence was placed under the control of the secondary cell wall specific Arabidopsis *IRX3* promoter. The construct was transformed into Arabidopsis *gux1/2/3* plants which lack [Me]GlcA branches on both primary and secondary cell wall xylan [16]. Use of these mutant plants ensured that any [Me]GlcA detected in the transgenic plants would be generated by the PgGUX enzyme. The degree of xylan glucuronosylation was evaluated by analysis of xylanase GH11 digestion products by PACE (Fig. 4a), as described [33]. In all three homozygous transgenic *gux1/2/3* lines expressing PgGUX, [Me]GlcA decorations were reintroduced (Fig. 4a, b). Thus, the cloned PgGUX enzyme is an active conifer glucuronosyltransferase in vivo.

To evaluate the recalcitrance of stem biomass from plants expressing PgGUX, we performed saccharification of AIR. In PgGUX lines, the monosaccharide release was reduced to levels measured for WT (Fig. 4c), indicating that the reintroduced glucuronosylation of the xylan is able to restore recalcitrance to the cell walls as effectively as native decorations.

Discussion

Numerous studies have genetically modified plant cell walls to decrease recalcitrance but the improved saccharification has been moderate, or offset by yield penalties [34–36]. *Arabidopsis* without glucuronic acid decorations on xylan *(gux1/2)* shows no yield penalty (Additional file 2: Figure S2, [12, 16]). Here, we have demonstrated an exceptional increase in sugar release during enzymatic saccharification specifically for this xylan branching mutant. Moreover, there is a substantially improved SSF ethanol yield from chemically unpretreated *gux1/2* biomass.

In our work, we evaluated the impact of all modifications of dicot secondary cell wall xylan branching on biomass recalcitrance. No significant recalcitrance reduction was observed in plants with decreased xylan acetylation or methylation using our conditions, which is consistent with the no to moderate phenotypes previously reported with different saccharification conditions [11, 17].

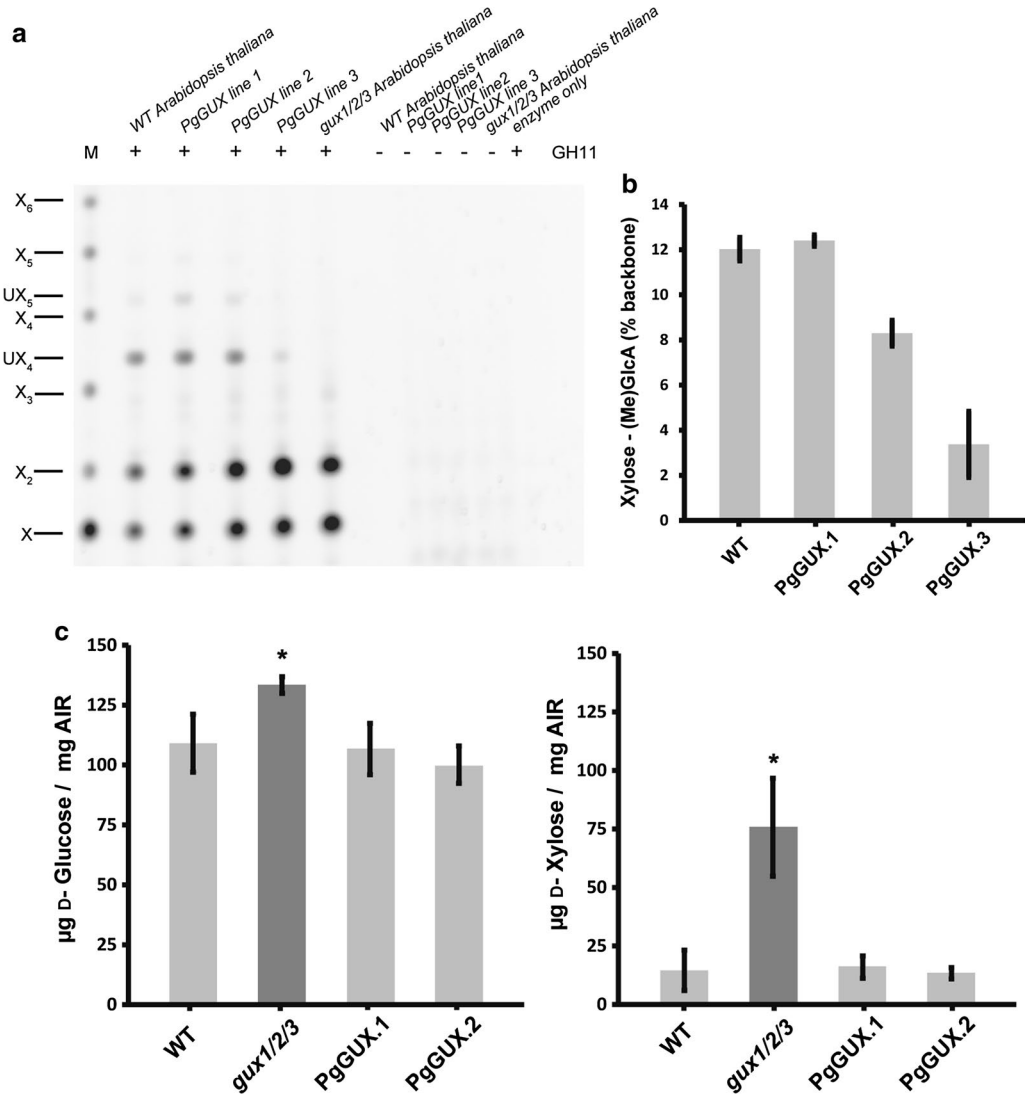


Fig. 4 PgGUX is a functional xylan glucuronosyltransferase. **a** PACE analysis of GH11 xylanase digests of WT, three independent transgenic lines of PgGUX in *gux1/2/3* and control *gux1/2/3* AIR. Undigested AIR controls (−). The [Me]GlcA-xylotetraose band (UX_4) was observed only in WT and PgGUX expressing lines. **b** Quantitation of degree of [Me]GlcA substitutions. **c** D-glucose and D-xylose release following saccharification of WT, *gux1/2/3* and two lines of PgGUX AIR. Error bars represent standard deviation of three biological replicates, **p* value ≤0.05

We found that *gux1/2* is unique among xylan decoration mutants in yielding a strong improvement in sugar release following enzymatic saccharification. The effect is especially marked for xylan, where most of the xylan becomes enzymatically digestible without any chemical cell wall pretreatments. These observations strongly indicate *GUX* genes as prime candidates for designing biomass for biorefining. The molecular origins of biomass recalcitrance are not fully understood, but are thought to involve protection of the cellulose from enzyme attack through sheathing in hemicellulose and embedding in lignin [37–40]. The molecular dynamic simulations and acetylation pattern of xylan of the *gux1/2* mutant suggest that xylan is able to interact with cellulose microfibrils [41]. The unexpected sensitivity of recalcitrance to the presence of [Me]GlcA on xylan could be due to other factors, such as alteration of xylan linkages to lignin via a proposed [Me]GlcA ester [42, 43]. Alternatively, lack of [Me]GlcA on xylan may impact the relative positioning of cellulose microfibrils and lead to changes in lignin deposition patterns [44] which may result in reduction of biomass recalcitrance.

Softwood, an abundant biomass resource, has a simple pattern of [Me]GlcA decorations [15]. This suggests a possible lower complexity of the xylan biosynthesis machinery than in eudicots, which could be an advantage

when engineering the conifer genome. We used a synthetic biology approach to characterise the xylan glucuronosyltransferase from *P. glauca* (PgGUX). PgGUX was able to glucuronosylate xylan in vitro and in the *gux1/2/3* mutant *Arabidopsis thaliana*. The efficiency of in vitro GlcA transfer by the PgGUX (Fig. 3) was comparable to degree of xylan glucuronisation *in planta* (Fig. 4b), showing unprecedented yield for a plant polysaccharide in vitro glycosylation reaction. Importantly, the [Me] GlcA decorations introduced by the PgGUX enzyme suppressed the improved saccharification phenotype observed in *gux1/2/3* plants, indicating that xylan glucuronosylation by PgGUX is functional.

Identification of a gymnosperm GUX enzyme is a major advance in the study of softwoods as a biomass resource and can contribute to the development of transgenic conifers more suited for biofuel production and biorefining. Not only the digestibility of cellulose can be improved in GUX-deficient Arabidopsis, but also their altered cell walls promote especially a more efficient utilisation of hemicelluloses, which form up to one-third of dry biomass. Additionally, many paper and pulp production processes remove hemicelluloses from wood using thermo-chemical treatment. It is likely that the efficiency of more environmentally benign pulping processes, such as enzymatic hemicellulose degradation, would be improved in GUX-deficient plants. Being a relatively simple alteration itself, *GUX* removal could be combined with other biomass modifications, further improving the quality of the lignocellulosic feedstock.

Conclusion

As the improvements in saccharification and fermentation are especially prominent in *gux1/2* biomass without chemical treatment, our findings will promote innovation in more environmentally and economically sustainable processes of biomass use. Moreover, efficient saccharification without chemical pretreatment opens the possibility of using sugars from recalcitrant biomass such as wood as a feedstock for in vitro protein production, since toxic inhibitors are avoided with this pretreatment. As digestibility is improved in GUX-deficient plants, it may become possible to broaden the biomass that can contribute to animal feed, for example by using wood. With the growing population and pressure on land utilisation, such alternative uses of otherwise non-digestible biomass will become more significant. It is important to appreciate that different factors may determine recalcitrance of hardwoods and softwoods. Thus, softwood derived from conifers without functional copies of *PgGUX* homologues will need to be evaluated. Nonetheless, the discovery of [Me]GlcA importance for *A. thaliana* biomass recalcitrance together with identification of the softwood glucuronosyltransferase should provide a strong incentive for industrial and governmental organisations to invest in work aiming at engineering conifer genomes.

Methods

Plant material used and AIR preparation

*A*rabidopsis *thaliana* plants of the Columbia-0 ecotype were grown in a cabinet maintained at 21 °C, with a 16-h light, 8-h dark photoperiod. Mutant insertion lines described in [12] (*gux1/2)* and [16] (*gux1/2/3)* were used for saccharification and transformation experiments, respectively. Alcohol insoluble residue (AIR) was prepared from 5-cm-long sections of mature *A. thaliana* stem. All AIR preparation was carried out as described in [12].

A. thaliana biomass saccharification using Cellic® CTec2

Novozymes Cellic® CTec2 (also available from Sigma-Aldrich/Merck) was used for all saccharification and fermentation experiments. Enzyme stock (35 µL) was diluted to a total volume of 2.5 mL with 0.1M ammonium acetate pH = 5.0 (AmAc) buffer. The enzyme sample was cleared from residual sugars using PD-10 desalting column (GE Healthcare) and eluted using 3.5 mL AmAc buffer, generating 1:00 (v/v) Cellic® CTec2 solution. AIR aliquots (1 mg) were homogenised in 1 mL of AmAc buffer. Homogenised AIR was amended with 25 µL 1:100 Cellic® CTec2 working solution.

For saccharification of dried stems 20 µL of 1:10 (v:v) Cellic® CTec2 solution was used. The enzyme solution was added to 1 mg stem material/ml 0.1 M AmAc buffer. Stem suspension was generated by ball milling 8 mg of the biomass in 8 mL buffer for three periods of 10 min at 25 Hz, with 10-min intervals between each ball milling.

For both AIR and dried stems, saccharification was carried out for 24 h at 45 °C with 1400 rpm applied for 30 s every 4 min. The reaction was terminated by heat-treating the suspension at 100 °C for 10 min. D-Glucose and D-Xylose release from the biomass was quantified using commercial kits (Megazyme, catalogue codes: K-XYLOSE and K-GLUHK-220A). Sugar concentration for each experiment was standardised with readings obtained from biomass and enzyme only controls.

Glucuronidase supplementation was performed by incubating the products of Cellic® CTec2 saccharification with 10 µL of 1 mg/mL *Bacteroides ovatus* GH115d (Bo_03449).

Simultaneous saccharification and fermentation (SSF) experiments

50 mg dried stems of WT and *gux1/2* plants were used for each fermentation reaction. Stems were ball milled in 7 mL LB medium for 4 periods of 5 min at 20 Hz, with 5-min intervals between each ball milling cycle. The

material was removed from ball milling vessel. To fully recover the biomass, the vessel was washed with further 2.5 mL LB. The stem suspension was sterilised by heat treatment at 85 °C for 10 min followed by cooling on ice. Each fermentation reaction was amended with 250 µL 1:10 Cellic® CTec2 solution, prepared as described in saccharification section, and 250 µL of TOP10 *E. coli* inoculum bearing the BBa_K1122676 BioBrick (OD_{600} of the inoculum was within 0.55–0.6 range). BBa_K1122676 encodes a Pyruvate decarboxylase and Alcohol dehydrogenase from *Zymonomas mobilis* which allow ethanol production in *E. coli* [25]. Biomass only reactions were supplemented with 250 µL of AmAc buffer and the bacterial inoculum. The plasmid was maintained by provision of 25 µg/mL Chloramphenicol (Duchefa Biochemie). The simultaneous saccharification and fermentation reactions were carried out for 96 h. at 37 °C and 200 rpm. Fermentation vessel was kept air-tight throughout the experiment. Ethanol levels were analysed using a commercial kit (Megazyme, catalogue code: K-ETOH).

Molecular phylogeny analysis of GUX amino acid sequences

The coding sequences of *A. thaliana GUX 1*, *2* and *3* were used to identify putative GUX encoding transcripts from *Populus trichocarpa* using data available via the NCBI BLAST service. The same *Arabidopsis* CDSs were used as a query to identify transcripts encoding putative GUX enzymes from Coniferophyta and Gnetophyta transcriptomic data available via 1000 Plant Genome BLAST service [26]. PgGUX transcript and amino acid sequences were retrieved from GeneBank. All amino acid sequences were reconstructed from transcripts with ExPASy translate tool and aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE) algorithm. A maximum likelihood phylogenetic tree was constructed using MEGA 6 software [45].

Molecular cloning and generation of transgenic *A. thaliana* lines

The gene encoding the *P. glauca* enzyme with a 3xMyc C-terminal tag was synthesised by GeneScript. Gateway cloning was used to insert the gene into the p3KC binary vector [46]. Protein expression was driven by a 1.7 kbp promoter sequence of *A. thaliana IRX3* gene. *A. thaliana gux1/2/3* plants were transformed using the floral dip method [47]. Kanamycin resistant plants were screened for the construct using PCR (Forward primer: 5′-ACTCC CAGTTGGATCCTGTG-3′, Reverse primer: 5′-TCCAT AAGCTGGAAGGT-3′). Three independent *gux1/2/3* lines homozygous for the pIRX3:PgGUX-Myc:NosT construct were derived and analysed in this study.

Expression of PgGUX in *Nicothiana benthamiana*

PgGUX-3xMyc was amplified from the synthetic construct using Q5 DNA Polymerase (NEB, Forward primer: 5′-AT GAGGCCCTCTTCAGGAGTTC-3′, Reverse primer: 5′-TCAAAGCAAATCCTCTTCTGAGATCAGT-3′). PCR product was ligated into NruI (NEB) digested pEAQ-HT *N. benthamiana* overexpression vector [48] using T4 DNA ligase (Thermo-Fisher Scientific). The construct was transformed into competent AGL-1 *Agrobacterium tumefaciens* and infiltrated into *N. benthamiana* leaves according to a published protocol [49]. Leaves were harvested 3 days following the infiltration and the membranes fraction enriched for PgGUX was collected as described in [13]. Same protocol was followed for GTL6-3xMyc overexpression.

Western blot analysis of PgGUX and GTL6 enriched membranes fraction

Protein concentration in the membranes fraction was quantified using modified Bradford reagent (Expedeon). Each well of SDS-PAGE (10–15% gradient, Bio-Rad) was loaded with 2.5 or 5 µg of PgGUX or GTL6 enriched *N. benthamiana* leaf membrane protein. Following the run, the gel was transferred onto nitrocellulose membrane using iBlot system (Life Technologies). The membrane was blocked o/n in 5% milk in TBS solution. The following day it was probed with 1:2000 anti-Myc primary antibody (rabbit polyclonal, Santa-Cruz, A14) and with 1:10,000 mouse anti-rabbit HRP linked secondary antibody (Bio-Rad, 170-6515). Amersham ECL prime HRP substrate (GE-Lifesciences) was used to obtain signal from membrane bound antibodies.

Glucuronosyltransferase activity assay

Acetylated heteroxylan lacking GlcA decorations was extracted from *gux1gux2 A. thaliana* as previously described [41]. Buffer exchange PD-10 columns (GE Lifesciences) were used to remove xylan from DMSO and elute it in water. Xylan aliquots were dried and used as an acceptor for in vitro GlcA transfer reaction. Each reaction mix was prepared as described in [16] with omission of UDP-Xylose. UDP-GlcA (5 mM) was replaced with water in certain reactions to control for non-specific glucuronosylation. Reaction was terminated with heat treatment (100 °C, 10 min) and the polysaccharides were extracted using methanol and chloroform as previously described [16]. Extracted polysaccharides were pelleted with 70% ethanol, washed in 100% ethanol and dried. Dry pellet was deacetylated with 4 M NaOH and digested with *Neocallimastix patriciarum* GH11 as previously described [12].

Polysaccharide analysis by carbohydrate gel electrophoresis (PACE) of xylanase GH11 digestion products

AIR material (0.5 mg) was digested with *N. patriciarum* GH11 enzyme overnight as described in [12]. Released oligosaccharides were dried and derivatised with 8-aminonapthalene-1,3,6-trisulphonic acid (ANTS; Invitrogen). ANTS derivitasation, PACE running and visualisation were performed as previously described [50, 51].

Statistical analysis and sampling

For the analysis of sugar release, fermentation efficiency and degree of GlcA substitution all plants were grown in three biological replicates. Each biological replicate consisted of a pooled sample of 36 plants. For each biological replicate, 3 technical replicates were analysed. Average sugar release/ethanol for each biological replicate was used for the statistical analysis with Student's *T* test. The variance between each WT and mutant pair was estimated to be similar with Levene's test.

Additional files

Additional file 1: Figure S1. Analysis of *gxm 1/2/3* plants lacking methylation of GlcA. The triple mutant plants grow to the same height (**A**). The xylan (**B**) and GlcA (**C**) content of *gxm1/2/3* plants is not different when compared to WT. Growth images and xylan quantitation are representable for 3 biological replicates.

Additional file 2: Figure S2. Analysis of *gux1/2* plant growth and biomass production. Plant growth is not affected by removal of GlcA branches from xylan. The graphs represent average height of 7 week old plants (**A**, n = 36 for WT and 34 for *gux1/2*), average total plant mass (**B**, n = 36 for WT and 34 for *gux1/2*) and average mass of 5 cm basal stem sections (**C**, n = 36 for WT and 33 for *gux1/2*). Error bars represent standard deviation. There is no statistically significant difference between the values measured for the WT and the mutant plant (Student's t test).

Additional file 3: Figure S3. Sugar release from WT and *gux1/2* biomass for saccharification reactions supplemented with glucuronidase GH115. Reaction scheme for GH115 supplementation (**A**) and D-Xylose release from Col0 (WT) and *gux1/2/3* biomass following supplementation (+) or not (-) with GH115 (**B**). Sterilisation steps were carried out between different stages of the experiment to avoid microbial growth in the reaction tubes. Reactions were performed in triplicate, and xylose measured after stage 1 or stage 3.

Additional file 4: Table S2. Gene Bank and OneKP transcript catalogue numbers. Those transcripts encode conifer GUX enzymes used to construct the maximum likelihood phylogeny presented in Fig. 2a.

Additional file 5: Figure S4. Expression heat-map for the putative conifer *GUX*. ExIMAGE feature of the Congenie datanase [30] was used to visualise the expression of *Picea abies* gene MA_84103g0010 which encodes a homologue of PgGUX. Reads encoding the enzyme are clearly enriched in both late and early wood.

Additional file 6: Figure S5. Western Blot analysis of *N. benthamiana* membrane fraction extracted from leaves enriched for GTL6 and PgGUX. PageRuler™ Prestained Protein Ladder, 10 to 180 kDa (Thermo-Fisher Scientific) was used as a molecular size marker.

Additional file 7: Table S1. Average ethanol production during Simultaneous saccharification and Fermentation experiments. This data was used to generate Fig. 1e.

Abbreviations

AIR: alcohol insoluble residues; AmAc: ammonium acetate; ANTS: 8-aminonapthalene-1,3,6-trisulphonic acid; GH: glycosyl hydrolase; GlcA: glucuronic acid; GUX: glucuronic acid substitution of xylan; GXM: glucuronoxylan methyltransferase; HRP: horseradish peroxidase; IRX: irregular xylem; MeGlcA: 4-O-methylglucuronic acid; OneKP: one thousand plants; PACE: polysaccharide analysis by carbohydrate gel electrophoresis; Pg: *Picea glauca*; SSF: simultaneous saccharification and fermentation; TBL: trichome birefringence-like; UDP-Xyl: uridine[5′]diphospho-α-d-xylopyranoside; UX4: glucurono-xylotetraose; WT: wild-type.

Authors' contributions

JJL and KBW and MBW designed and performed experiments and bioinformatics analysis, analysed data and wrote the paper; OMT performed experiments, analysed data and wrote the paper; XY and NF-B performed experiments and analysed data; DB designed and performed initial experiments and analysed data; KBK analysed data and wrote the paper; PD designed experiments, analysed data and wrote the paper. All authors read and approved the final manuscript.

Author details

[1] Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK. [2] The Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK. [3] Natural Material Innovation Centre, University of Cambridge, 1 Scroope Terrace, Cambridge CB2 1PX, UK. [4] OpenPlant Synthetic Biology Research Centre, Department of Plant Sciences, University of Cambridge, Downing Site, Cambridge CB2 3EA, UK. [5] Department of Protein Biochemistry and Stability, Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Denmark. [6] Present Address: Ossianix, Stevenage Bioscience Catalyst, Gunnels Wood Rd, Stevenage, Hertfordshire SG1 2FX, UK. [7] Present Address: Shell Global Solutions International BV, Lange Kleiweg 40, 2288 GK Rijswijk, The Netherlands.

Acknowledgements

We would like to thank Prof. Christopher French (University of Edinburgh, UK) for providing us with DNA constructs required for ethanol fermentation in *E. coli*. We would like to acknowledge Prof. George Lomonossoff (John Innes Centre, UK), who developed the pEAQ-HyperTrans expression system used in this study. Plant Bioscience Limited supplied the pEAQ-HT vector which was used in this work. The GH11 and GH115 enzymes used in this work were a kind gift from Dr. Dave Bolam and Prof. Harry Gilbert (University of Newcastle, UK). We would also like to thank Dr. Nadine Anders and Dr. Henry Temple for helpful discussions and critical reading of the manuscript.

Competing interests

PD has filed a related patent application on xylan modification (WO2009037502). The remaining authors have no competing interests.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. All material used in this study can be requested from the corresponding authors.

Funding

This work was supported by the Leverhulme Trust Centre for Natural Material Innovation and the OpenPlant Synthetic Biology Research Centre. J.J.L. was in receipt of a studentship from the Biotechnology and Biological Sciences Research Council (BBSRC) of the UK as part of the Cambridge BBSRC-DTP Programme (Reference BB/J014540/1). O.M.T was a recipient of an iCASE studentship from the BBSRC (Reference BB/M015432/1).

Publisher's Note

Received: 2 June 2017 Accepted: 7 September 2017
Published online: 19 September 2017

References

1. Albers SC, Berklund AM, Graff GD. The rise and fall of innovation in biofuels. Nat Biotechnol. 2016;34:814–21.
2. Cate HDJ, Ball AS. Editorial overview: energy biotechnology. Curr Opin Biotechnol. 2016;38:v–vii.
3. Hood EE. Plant-based biofuels. F1000Res. 2016;17;5. pii: F1000 Faculty Rev-185.
4. McCann MC, Carpita NC. Biomass recalcitrance: a multi-scale, multi-factor, and conversion-specific property. J Exp Bot. 2015;66(14):4109–18.
5. Silveira MH, Morais AR, da Costa Lopes AM, Olekszyszen DN, Bogel-Łukasik R, Andreaus J, Pereira Ramos L. Current pretreatment technologies for the development of cellulosic ethanol and biorefineries. Chem Sus Chem. 2015;268(20):3366–90.
6. Himmel ME. Biomass recalcitrance: deconstructing the plant cell wall for bioenergy. Oxford: Blackwell Publishing; 2008.
7. Vanholme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W. Metabolic engineering of novel lignin in biomass crops. New Phytol. 2012;196(4):978–1000.
8. Wang P, Dudareva N, Morgan JA, Chapple C. Genetic manipulation of lignocellulosic biomass for bioenergy. Curr Opin Chem Biol. 2015;29:32–9.
9. Mottiar Y, Vanholme R, Boerjan W, Ralph J, Mansfield SD. Designer lignins: harnessing the plasticity of lignification. Curr Opin Biotechnol. 2016;37:190–200.
10. Scheller HV, Ulvskov P. Hemicelluloses. Annu Rev Plant Biol. 2010;61:263–89.
11. Xiong G, Cheng K, Pauly M. Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the *Arabidopsis* mutant tbl29. Mol Plant. 2013;6(4):1373–5.
12. Mortimer JC, Miles GP, Brown DM, Zhang Z, Segura MP, Weimar T, Yu X, Seffen KA, Stephens E, Turner SR, Dupree P. Absence of branches from xylan in *Arabidopsis* gux mutants reveals potential for simplification of lignocellulosic biomass. PNAS. 2010;5:17409–14.
13. Rennie EA, Hansen SF, Baidoo EE, Hadi MZ, Keasling JD, Scheller HV. Three members of the Arabidopsis glycosyltransferase family 8 are xylan glucuronosyltransferases. Plant Physiol. 2012;159(4):1408–17.
14. Brown DM, Goubet F, Wong VW, Goodacre R, Stephens E, Dupree P, Turner SR. Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. Plant J. 2007;52(6):1154–68.
15. Busse-Wicher M, Li A, Silveira RL, Pereira CS, Tryfona T, Gomes TC, Skaf MS, Dupree P. Evolution of xylan substitution patterns in gymnosperms and angiosperms: implications for xylan interaction with cellulose. Plant Physiol. 2016;171(4):2418–31.
16. Mortimer JC, Faria-Blanc N, Yu X, Tryfona T, Sorieul M, Ng YZ, Zhang Z, Stott K, Anders N, Dupree P. An unusual xylan in *Arabidopsis* primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14. Plant J. 2015;83(3):413–26.
17. Urbanowicz BR, Peña MJ, Ratnaparkhe S, Avci U, Backe J, Steet HF, Foston M, Li H, O'Neill MA, Ragauskas AJ, Darvill AG, Wyman C, Gilbert HJ, York WS. 4-O-methylation of glucuronic acid in Arabidopsis glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. PNAS. 2012;109(35):14253–8.
18. Li X, Jackson P, Rubtsov DV, Faria-Blanc N, Mortimer JC, Turner SR, Krogh KB, Johansen KS, Dupree P. Development and application of a high throughput carbohydrate profiling technique for analyzing plant cell wall polysaccharides and carbohydrate active enzymes. Biotechnol Biofuels. 2013;6(1):94.
19. Rogowski A, Baslé A, Farinas CS, Solovyova A, Mortimer JC, Dupree P, Gilbert HJ, Bolam DN. Evidence that GH115 α-glucuronidase activity, which is required to degrade plant biomass, is dependent on conformational flexibility. J Biol Chem. 2014;289(1):53–64.
20. Zhang J, Siika-Aho M, Tenkanen M, Viikari L. The role of acetyl xylan esterase in the solubilization of xylan and enzymatic hydrolysis of wheat straw and giant reed. Biotechnol Biofuels. 2011;4(1):60.
21. Bensussan M, Lefebvre V, Ducamp A, Trouverie J, Gineau E, Fortabat MN, Guillebaux A, Baldy A, Naquin D, Herbette S, Lapierre C, Mouille G, Horlow C, Durand-Tardif M. Suppression of dwarf and irregular xylem phenotypes generates low-acetylated biomass lines in *Arabidopsis*. Plant Physiol. 2015;168(2):452–63.
22. Pawar PM, Derba-Maceluch M, Chong SL, Gómez LD, Miedes E, Banasiak A, Ratke C, Gaertner C, Mouille G, McQueen-Mason SJ, Molina A, Sellstedt A, Tenkanen M, Mellerowicz EJ. Expression of fungal acetyl xylan esterase in *Arabidopsis thaliana* improves saccharification of stem lignocellulose. Plant Biotechnol J. 2016;14(1):387–97.
23. Chong SL, Derba-Maceluch M, Koutaniemi S, Gómez LD, McQueen-Mason SJ, Tenkanen M, Mellerowicz EJ. Active fungal GH115 α-glucuronidase produced in *Arabidopsis thaliana* affects only the UX1-reactive glucuronate decorations on native glucuronoxylans. BMC Biotechnol. 2015;18(15):56.
24. Cornuault V, Buffetto F, Rydahl MG, Marcus SE, Torode TA, Xue J, Crépeau MJ, Faria-Blanc N, Willats WG, Dupree P, Ralet MC, Knox JP. Monoclonal antibodies indicate low-abundance links between heteroxylan and other glycans of plant cell walls. Planta. 2015;242(6):1321–34.
25. Lewicka AJ, Lyczakowski JJ, Blackhurst G, Pashkuleva C, Rothschild-Mancinelli K, Tautvaišas D, Thornton H, Villanueva H, Xiao W, Slikas J, Horsfall L, Elfick A, French C. Fusion of pyruvate decarboxylase and alcohol dehydrogenase increases ethanol production in *Escherichia coli*. ACS Synth Biol. 2014;3(12):976–8.
26. Wickett NJ, et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. PNAS. 2014;111:E4859–68.
27. Matasci N, et al. Data access for the 1000 plants (1KP) project. GigaScience. 2014;3:17.
28. Xie Y, et al. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics. 2014;30:1660–6.
29. Johnson MT, et al. Evaluating methods for isolating total RNA and predicting the success of sequencing phylogenetically diverse plant transcriptomes. PLoS ONE. 2012;7:e50226.
30. Nystedt B, et al. The Norway spruce genome sequence and conifer genome evolution. Nature. 2013;497(7451):579–84.
31. Dunkley TP, Hester S, Shadforth IP, Runions J, Weimar T, Hanton SL, Griffin JL, Bessant C, Brandizzi F, Hawes C, Watson RB, Dupree P, Lilley KS. Mapping the *Arabidopsis organelle* proteome. PNAS. 2006;103(17):6518–23.
32. Voiniciuc C, Schmidt MH, Berger A, Yang B, Ebert B, Scheller HV, North HM, Usadel B, Günl M. Mucilage-Related10 produces galactoglucomannan that maintains pectin and cellulose architecture in arabidopsis seed mucilage. Plant Physiol. 2015;169(1):403–20.
33. Brown DM, Zhang Z, Stephens E, Dupree P, Turner SR. Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in *Arabidopsis*. Plant J. 2009;57(4):732–46.
34. Wang Y, Fan C, Hu H, Li Y, Sun D, Wang Y, Peng L. Genetic modification of plant cell walls to enhance biomass yield and biofuel production in bioenergy crops. Biotechnol Adv. 2016;34:997–1017.
35. Marriott PE, Gómez LD, McQueen-Mason SJ. Unlocking the potential of lignocellulosic biomass through plant science. New Phytol. 2016;209:1366–81.
36. Loqué D, Scheller HV, Pauly M. Engineering of plant cell walls for enhanced biofuel production. Curr Opin Plant Biol. 2015;25:151–61.
37. Öhgren K, Bura R, Saddler J, Zacchi G. Effect of hemicellulose and lignin removal on enzymatic hydrolysis of steam pretreated corn stover. Biores Technol. 2007;98(13):2503–10.
38. Fernandes AN, Thomas LH, Altaner CM, Callow P, Forsyth VT, Apperley DC, Kennedy CJ, Jarvis MC. Nanostructure of cellulose microfibrils in spruce wood. PNAS. 2011;108(47):1195–203.
39. Cosgrove DJ, Jarvis MC. Comparative structure and biomechanics of plant primary and secondary cell walls. Front Plant Sci. 2012;22(3):204.
40. Simmons TJ, Mortimer JC, Bernardinelli OD, Pöppler AC, Brown SP, deAzevedo ER, Dupree R, Dupree P. Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. Nat Commun. 2016;21(7):13902.
41. Busse-Wicher M, Gomes TC, Tryfona T, Nikolovski N, Stott K, Grantham NJ, Bolam DN, Skaf MS, Dupree P. 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. 2014;79(3):492–506.
42. Giummarella N, Lawoko M. Structural basis for the formation and regulation of lignin-xylan bonds in birch. ACS Sustain Chem Eng. 2016;4:5319–26.
43. Watanabe T, Koshijima T. Evidence for an ester linkage between lignin and glucuronic-acid in lignin carbohydrate complexes by DDQ-oxidation. Agric Biol Chem. 1988;52:2953–5.

44. Reis D, Thellier M. Helicoidal pattern in secondary cell walls and possible role of xylans in their construction. C R Biol. 2004;327(9–10):785–90.
45. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–9.
46. Atanassov II, Atanassov II, Etchells JP, Turner SR. A simple, flexible and efficient PCR-fusion/Gateway cloning procedure for gene fusion, site-directed mutagenesis, short sequence insertion and domain deletions and swaps. Plant Methods. 2009;5:14–25.
47. Clough SJ, Bent AF. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. 1998;16(6):735–43.
48. Sainsbury F, Thuenemann EC, Lomonossoff GP. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol J. 2009;7(7):682–93.
49. Sparkes IA, Runions J, Kearns A, Hawes C. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nat Protoc. 2006;1(4):2019–25.
50. Bromley JR, Busse-Wicher M, Tryfona T, Mortimer JC, Zhang Z, Brown DM, Dupree P. GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. Plant J. 2013;74(3):423–34.
51. Goubet F, Jackson P, Deery MJ, Dupree P. Polysaccharide analysis using carbohydrate gel electrophoresis: a method to study plant cell wall polysaccharides and polysaccharide hydrolases. Anal Biochem. 2002;300(1):53–68.