STRUCTURAL DIVERSITY OF HETEROXYLANS ISOLATED FROM THE CELL WALLS OF TAXONOMICALLY DIVERSE LAND PLANTS

by

AMEYA R. KULKARNI

(Under the Direction of William S. York)

ABSTRACT

Plant secondary cell walls have received considerable attention in recent years as they account for the bulk of plant biomass and are thus a renewable source of carbon for the production of biofuels and other added-value chemicals. Heteroxylans are the predominant hemicellulose in biomass from woody plants and grasses. These xylans have a backbone composed of 1,4-linked β-D-xylopyranosyl (Xyl*p*) residues that vary in the type, location and number of substitutions to the backbone. Eudicots synthesize glucuronoxylans that have α-D-glucosyluronic acid (α-D-Glc*p*A) and/or a 4-*O*-methyl α-D-glucosyluronic acid (4-*O*-Me-Glc*p*A) linked to *O*-2 of the backbone residues. By contrast, grass xylans contain a-L-arabinofuranosyl residues linked to *O*-3 of the backbone. Dicot glucuronoxylans have a unique glycosyl sequence (4-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-GalpA-(1,4)-D-Xyl*p*) at their reducing end, which has been shown using Arabidopsis mutants to be required for normal xylan synthesis and for normal plant growth and development. In my dissertation, I describe the structural characterization of heteroxylans isolated from the walls of members of the different monocot orders and seedless land plants and discuss how these changes are correlated with the evolution of these plants. I also show that members of the Poales and Asparagales have heteroxylans that lack the glycosyl sequence 4-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-GalpA-(1,4)-D-Xyl*p*) at their reducing end. The results of my studies provide new insight into the structure of

heteroxylans in the monocots and the nature of the sequence of glycosyl residues at the reducing ends of these xylans. In addition my research provides evidence that the sequence 4-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-GalpA-(1,4)-D-Xyl*p* was not eliminated when monocots and dicots diverged but was lost during the evolution of some monocot orders. Such knowledge provides additional insight into the structural diversity of land plant cell wall polysaccharides. My studies also provide a foundation for the application of genetic and molecular approaches to modify xylan structure and its interaction with other wall components to improve the economic value of plant biomass.

INDEX WORDS: Plant cell walls, land plant evolution, monocotyledons, grasses, *physcomitrella patens*, glucuronoxylan, glucuronoarabinoxylan, reducing end sequence.

STRUCTURAL DIVERSITY OF HETEROXYLANS ISOLATED FROM THE CELL WALLS OF TAXONOMICALLY DIVERSE LAND PLANTS

by

AMEYA R. KULKARNI

B.Sc., Hislop College, R.T.M. Nagpur University, India, 2007

M.Sc., Hislop College, R.T.M. Nagpur University, India, 2007

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2013

STRUCTURAL DIVERSITY OF HETEROXYLANS ISOLATED FROM THE CELL WALLS OF TAXONOMICALLY DIVERSE LAND PLANTS

by

AMEYA R. KULKARNI

Major Professor:	William S. York
Committee:	Lance Wells
	Michael G. Hahn
	Russell Malmberg

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2013

DEDICATION

To my parents, my sister Meenal and my wife Tejashree, for their love, support, and motivation over the years.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. William S. York for his constant support for the past six and a half years. I am also very grateful to my committee members: Drs. Lance Wells, Michael G. Hahn, and Russell Malmberg.

I would like to thank my lab members: Breeanna R. Urbanowicz, Katrina Saffold, Maria J. Pena, and Sami Tuomivaara for their support. I cannot thank Malcolm A. O'Neill enough for all the help, support, and guidance.

Lastly, I am thankful to my family for their support over the years.

TABLE OF CONTENTS

Page

DEDICATION iv

ACKNOWLEDGEMENTS v

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

INTRODUCTION 1

HEMICELLULOSES AND THEIR STRUCTURE IN THE CELL WALL 5

XYLAN BIOSYNTHESIS IN DICOTS 19

XYLAN BIOSYNTHESIS IN GRASSES 24

REFERENCES 26

2 THE ABILITY OF LAND PLANTS TO SYNTHESIZE GLUCURONOXYLANS PREDATES THE EVOLUTION OF TRACHEOPHYTES 36

ABSTRACT 37

INTRODUCTION 38

RESULTS 40

DISCUSSION 59

MATERIALS AND METHODS 64

REFERENCES 75

SUPPLEMENTAL INFORMATION 83

3 COMPARISON OF ARABINOXYLAN STRUCTURES IN BIOENERGY AND MODEL GRASSES 103
ABSTRACT 104
INTRODUCTION 105
MATERIALS AND METHODS 107
RESULTS AND DISCUSSION 111
REFERENCES 126
4 STRUCTURAL DIVERSITY OF THE REDUCING END SEQUENCES OF HETEROXYLANS PRESENT IN THE CELL WALLS OF MONOCOTYLEDENOUS PLANTS 136
ABSTRACT 137
INTRODUCTION 138
RESULTS 143
DISCUSSION 161
MATERIALS AND METHODS 165
REFERENCES 169
SUPPLEMENTAL INFORMATION 187
5 CONCLUSION 192

Chapter 1

INTRODUCTION AND LITERATURE REVIEW

Introduction

The plant cell wall and its ultrastructure

Plant cells are surrounded by dynamic cell walls that play critical roles in many essential biological processes, including control of cell and tissue morphology, intercellular communication, plant–microbe interactions, and defense against pathogens. As the predominant component of terrestrial biomass, plant secondary cell walls have recently gained importance as a source of renewable fuels and as feedstocks for the production of man made materials (Pauly and Keegstra 2010, Scheller and Ulvskov 2010, Doering et al. 2012). However, a major impediment to the efficient use of plant cell walls in this context is their recalcitrance to microbial or mechanical deconstruction. Various treatments, such as heating in the presence of acid or base followed by incubation with enzymes that release soluble monosaccharides and oligosaccharides, have been developed to make recalcitrant cell walls more amenable to downstream processes such as fermentation. (Himmel et al. 2007) However, a more complete understanding of the structure and biosynthesis of the plant cell wall is required to develop better industrial protocols for biomass treatment and new biomass crop lines that produce less recalcitrant biomass.

Plant cell walls, which are complex aggregates composed predominantly of cellulose microfibrils, hemicelluloses, pectin polysaccharides and lignin, are essential for normal growth and development during the entire life cycle of a plant (Sandhu et al. 2009). Plants form two types of cell walls (primary and secondary), which differ in function and composition (**Figure 1.1**). Primary walls form a dynamic structure around growing cells and provide strength while allowing controlled growth. Secondary cell walls, which are much thicker and stronger, account for most of the carbohydrate in biomass that is deposited once the cell has ceased to grow. The secondary walls of fiber cells, tracheids, and sclereids in the xylem are further strengthened by the incorporation of lignin (Carpita and Gibeaut 1993, Carpita 1996).

Primary cell walls of flowering plants are broadly divided into type I and type II cell walls (Carpita and Gibeaut 1993, Carpita 1996, Pauly and Keegstra 2008, Vogel 2008). Type I cell walls are common features of dicots, gymnosperms and noncommelinid monocots (e.g. acorales). The predominant components of Type I cell walls are cellulose microfibrils, hemicelluloses, pectin and structural proteins. The major non-cellulosic component of type I cell walls is xyloglucan (XyG). In contrast, type II cell walls, which are present in commelinids (e.g. poales, commelinales, zingiberales), contain large amounts of cellulose, hemicelluloses along with smaller amounts of pectins and structural proteins as compared to type I cell walls. Glucuronoarabinoxylan (GAX) is the major hemicellulose in type II cell walls, (Pauly and Keegstra 2008, Vogel 2008). Significant amounts of mixed linkage glucans are found in the type II cell walls of plants in the Poaceae family.

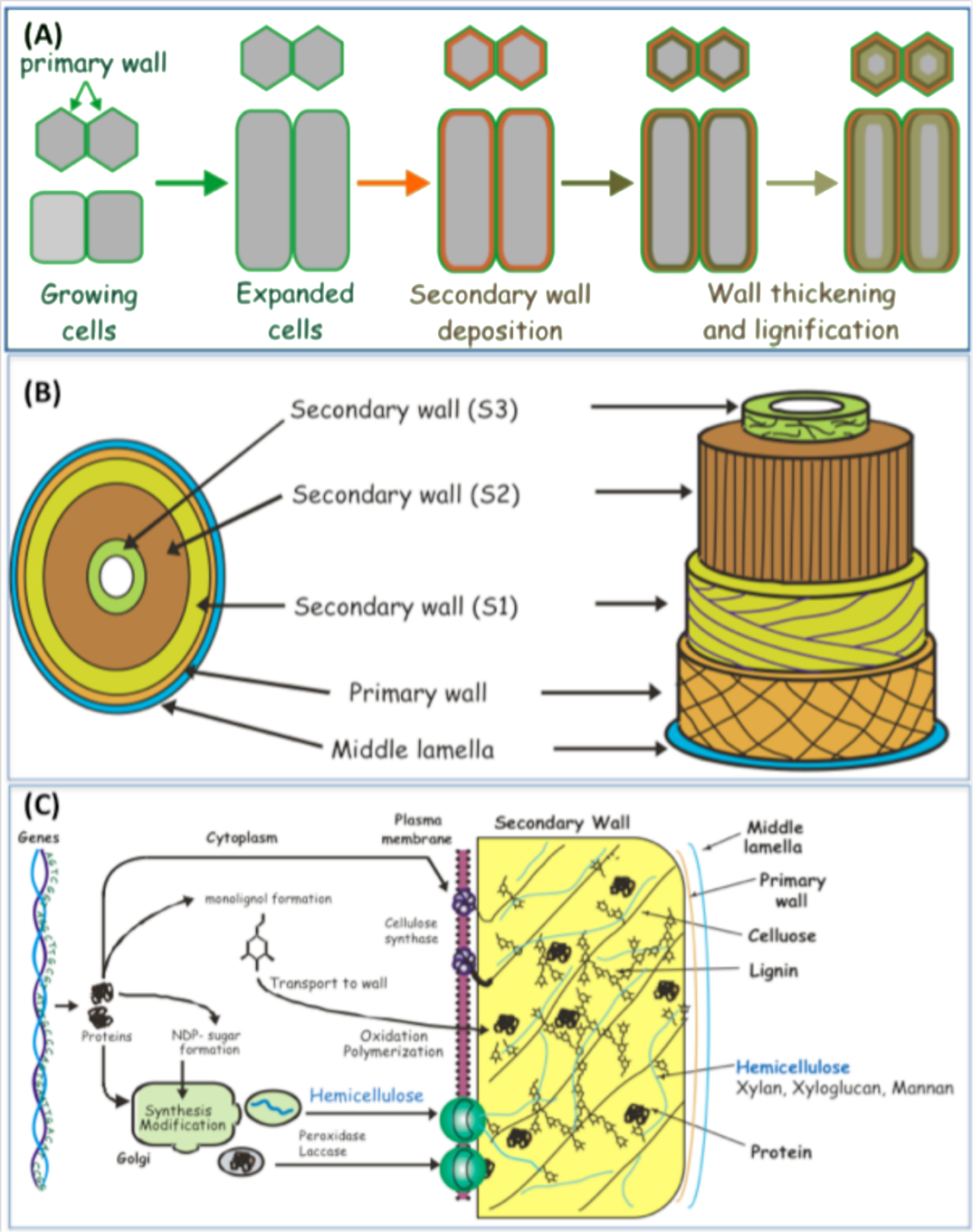


Figure 1.1: Model of primary and secondary cell walls and their synthesis.

From (Mohnen et al. 2008); originally generated by Malcolm O'Neill, CCRC.

(A) Primary cell walls surrounds dividing and expanding plant cells. Primary cell walls are capable of controlled expansion to allow the cell to grow yet remain sufficiently strong to prevent the cell from rupturing. On the other hand, secondary walls are formed after the cells have stopped growing and are often lignified to increase their strength and hydrophobicity.

(B) Once the plant cell has stopped growing, secondary thickening of its cell wall can occur. Three layers (S1, S2, and S3) of secondary walls with different cellulose microfibril orientation are laid down sequentially between the primary wall and the plasma membrane.

(C) Cellulose synthases (CesAs) catalyze the synthesis of secondary wall cellulose at the plasma membrane. Hemicelluloses are synthesized in the Golgi apparatus and then transported to the walls. Monolignols are synthesized in the cytoplasm and then transported to the cell wall, where they are polymerized to form lignin.

Hemicelluloses and their structures in the cell wall

Hemicelluloses comprise of xyloglucan (XyG), heteromannans, mixed linked glucans, and xylan. The amount of each of these hemicelluloses in a cell wall varies depending on the plant species and tissue type. For example, mixed linked glucans are commonly found in grasses but not in other species. The structures of different hemicelluloses in plant cell walls are described in **Figure 1.2**.

Xyloglucans

The xyloglucan backbone consists of 1,4 linked β-D-glucan (Glc) residues, which are substituted at O-6 with α-D-Xylose (Xyl) (**Figure 1.2**). A number of the side chains variations are observed in the XyG structures. These variations are thought to have both functional and taxonomic significance (Scheller and Ulvskov 2010). In most dicots and non-graminaceous monocots, Xyl residues substitute three out of four Glc residues of backbone, whereas, in commelinid monocots (e.g. Poales) and most lower land plants (moss), the Glc backbone is less substituted with Xyl residues (Pauly et al. 1999, Jia et al. 2003, Peña et al. 2008, Hsieh and Harris 2009). In general, the XyG side chains are divided into charged and uncharged side chains.

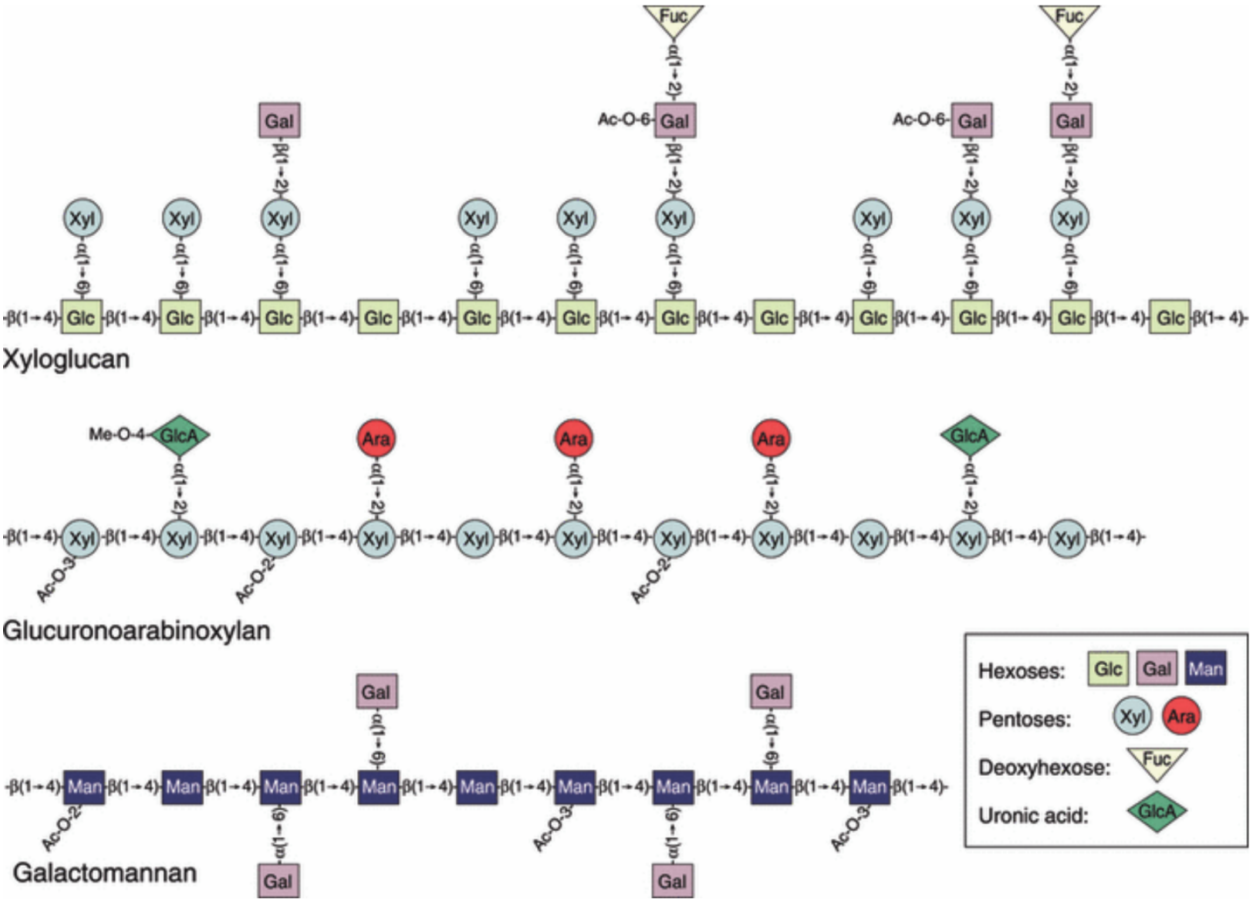


Figure 1.2: Types of hemicellulose structures present in plant cell walls.

From (Pauly and Keegstra 2008). Xyloglucan, a major component of type I cell walls, whereas glucuronoarabinoxylan is a major component of type II cell walls. Galactomannans are abundantly present in the secondary cell walls of conifers (Pauly and Keegstra 2008, Scheller and Ulvskov 2010).

Mannans

Mannans are grouped into four categories, namely, mannan, glucomannan, galactomannan, and galactoglucomannan. They are major hemicelluloses of conifer cell walls (Ebringerová et al. 2005, Scheller and Ulvskov 2010, Pauly et al. 2013). Mannans and galactomannans have backbones consisting of 1,4-linked β-D-mannose, whereas galactomannan and galactoglucomannan backbones contain both β-1,4-linked glucose and β-1,4-linked mannose residues (**Figure 1.2**).

Mixed linked glucans

Mixed linked glucans are well-established components of the primary cell walls of grasses. These hemicellulosic polysaccharides are composed of linear chain β-(1→3) and (1→4) linked glucose residues. They are proposed to play a vital role in cell expansion varying at different growth stages (Obel et al. 2002, Gibeaut et al. 2005, Scheller and Ulvskov 2010). These polysaccharides have not been observed in dicots but have been detected in *Equisetum, Charophytes,* and red algae. It has also been suggested that mixed linked glucans might also be present in liverworts (Popper and Fry 2003, Scheller and Ulvskov 2010).

Xylans

Heteroxylans are the major hemicellulosic component of secondary cell walls of land plants. These polysaccharides have a backbone composed of (1-4)-linked-β-D-xylosyl residues. True homoxylans are rare and in almost all cases the backbone is substituted, to varying degrees, with monosaccharide or disaccharide side chains (Ebringerová et al. 2005, York and O'Neill 2008,

Scheller and Ulvskov 2010). The diversity of heteroxylan structures in land plants is shown in **Figure 1.3**. In eudicots, the xylan backbone is typically substituted with α-D-glucosyluronic acid (Glc*p*A) or 4-*O*-methyl α-D-glucosyluronic acid (MeGlc*p*A) and is known as glucuronoxylan (GX) (**Figure 1.3**). The grasses (Poaceae) synthesize arabinoxylans (AX) whose backbones are substituted predominantly at *O*-3 with α-L-arabinofuranose (Ara*f*) residues. Small amounts of Glc*p*A and MeGlc*p*A may also be present in these polysaccharides, which are collectively referred to as glucuronoarabinoxylans (GAX) (Ebringerová et al. 2005, York and O'Neill 2008, Faik 2010, Scheller and Ulvskov 2010). Xylans are acetylated to varying degrees at *O*-3 of xylose residues and to a lesser extent at *O*-2, especially in dicot cell walls (Ebringerová et al. 2005). In grasses, another common feature is the *O*-5 substitution of Ara*f* residues with ferulic acid and/or *p*-coumaric acid esters.

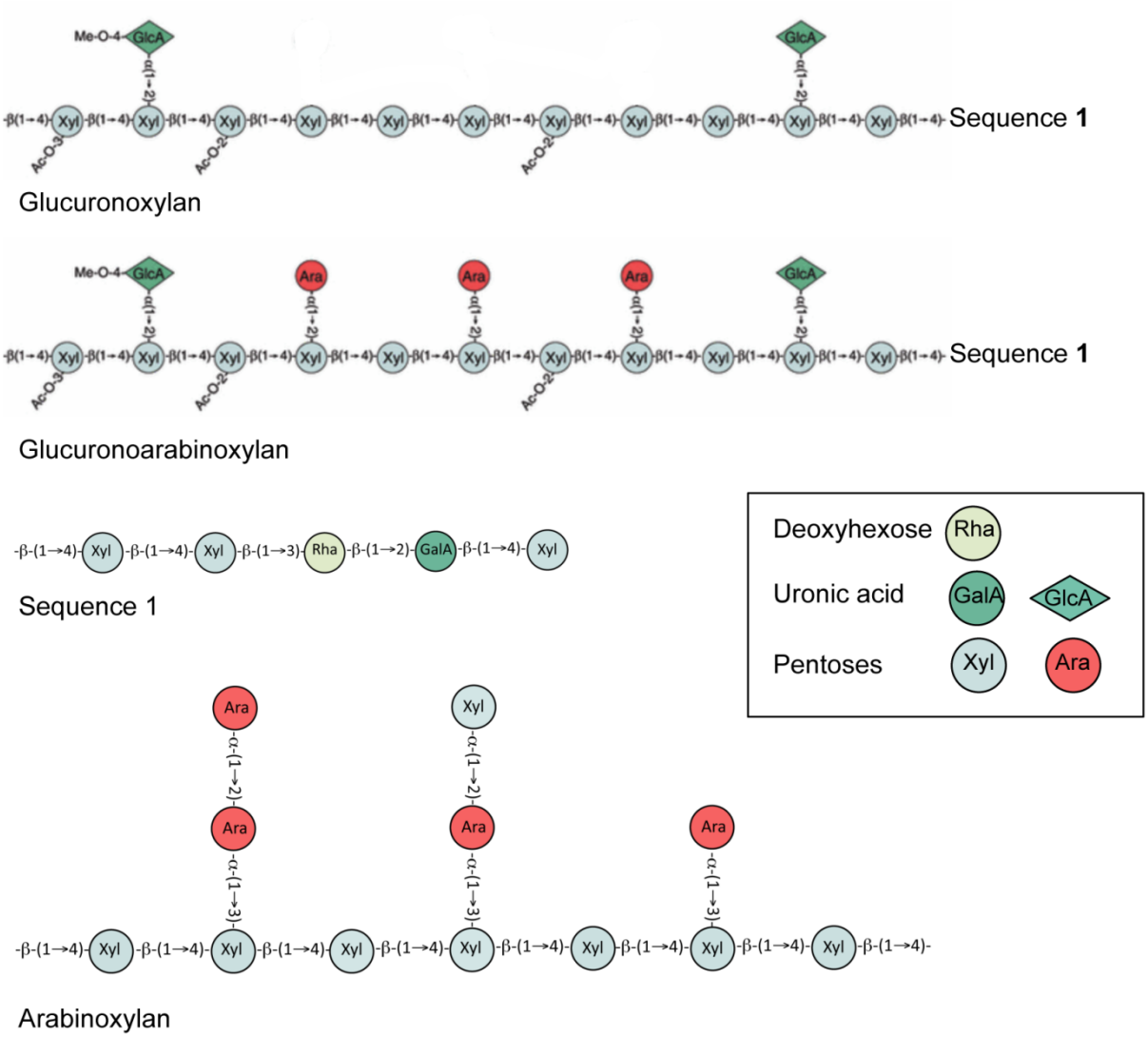


Figure 1.3: Different types of heteroxylans present in plant cell walls. Modified from (Pauly and Keegstra 2008). Glucuronoxylans produced by dicots and gymnosperms (softwood) are composed of a 1,4-linked xylan backbone that is substituted with GlcA and/or MeGlcA. Glucuronoarabinoxylan produced by gymnosperms (softwood) have Ara*f* residues along with GlcA and MeGlcA substitutions. Dicots and gymnosperms produce xylans with sequence **1** at the reducing end. Some of the xylosyl residues in the backbones of arabinoxylans in grass cell

walls bear sidechains at O-2 or O-3. These sidechains consist of an arabinosyl residue, which is sometimes substituted at O-2 with another arabinosyl residue or a xylosyl residue.

Xylan in seedless land plants

Xylans have been reported from the cell walls of several red algae (Rhodophytes), indicating that xylans are not restricted to land plants. However, the presence of xylan with sidechains containing of Glc*p*A, 4-*O*-Me-Glc*p*A, or Ara*f* residues has not been reported in these algae. For example, the cell walls of *Chaetangium fastigiatum* and *Scinaia hatei* have been reported to contain linear 1,4-linked xylans (Matulewicz and Cerezo 1987, Mandal et al. 2009), whereas, some others such as *Rhodymenia palmate* have been reported to contain xylan composed of 1,3- and 1,4-linked Xyl*p* residues (Percival and Chanda 1950). Xylans have been reported to be present in hornwort cell walls, but their structures have not been determined. However, it was unclear before I began my research whether xylan was present in the cell walls of mosses and liverworts (Carafa et al. 2005, Moller et al. 2007). An immunocytochemical study detected no xylan-specific epitopes in the cell walls of eight moss species, including *Funeria hygrometrica* (Carafa et al. 2005). However, the presence of xylan in the cell walls of *Physcomitrela patens* chloronemal filaments was indicated by immunochemical and glycosyl-residue linkage analyses (Moller et al. 2007). Before I began my research, there were no reports on complete structural characterization of xylans from seedless plants. Thus the first aim of my dissertation was to completely characterize xylans from *P. patens, S. krussiana,* and *E. hyemale* (See Chapter 2 for details). My study established that the mosses synthesize small amounts of glucuronoxylan that is structurally homologous to the glucuronoxylans present in the secondary cell walls of lycopodiophytes, monilophytes, and many seed-bearing plants

Glucuronoxylans have a unique sequence of glycosyl residues at their reducing ends

Glucuronoxylans in dicots and gymnosperm GXs contain the glycosyl sequence 4-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-GalpA-(1,4)-D-Xyl*p* (i.e., Sequence **1**) at their reducing ends (Shimizu et al. 1976, Johansson and Samuelson 1977, Peña et al. 2007). Analysis of several Arabidopsis mutants suggest that Sequence **1** is involved in controlling the xylan chain length, acting either as a primer to initiate the chain synthesis or as a substrate for enzyme catalyzed reactions that terminate xylan (Peña et al. 2007, York and O'Neill 2008). Two models for the biosynthesis of GX have been proposed (York and O'Neill 2008). In the first model (**Figure 1.4A**), the backbone is elongated by the addition of the xylosyl residues to the reducing end and the nascent polymer is transfered to Sequence **1**, which acts as a chain terminator by preventing further addition of residues to the reducing end. In the second model, Sequence **1** acts as a primer for xylan biosynthesis and the xylan chain is elongated by iterative addition of xylosyl residues to the nonreducing end of the nascent polymer (**Figure 1.4B**) (Persson et al. 2005, Brown et al. 2007, Lee et al. 2007a, York and O'Neill 2008). Chain elongation ceases when the connection of Sequence **1** to a putative carrier is broken so the xylan can escape the xylan enzyme complex that catalyzes xylan transfer.

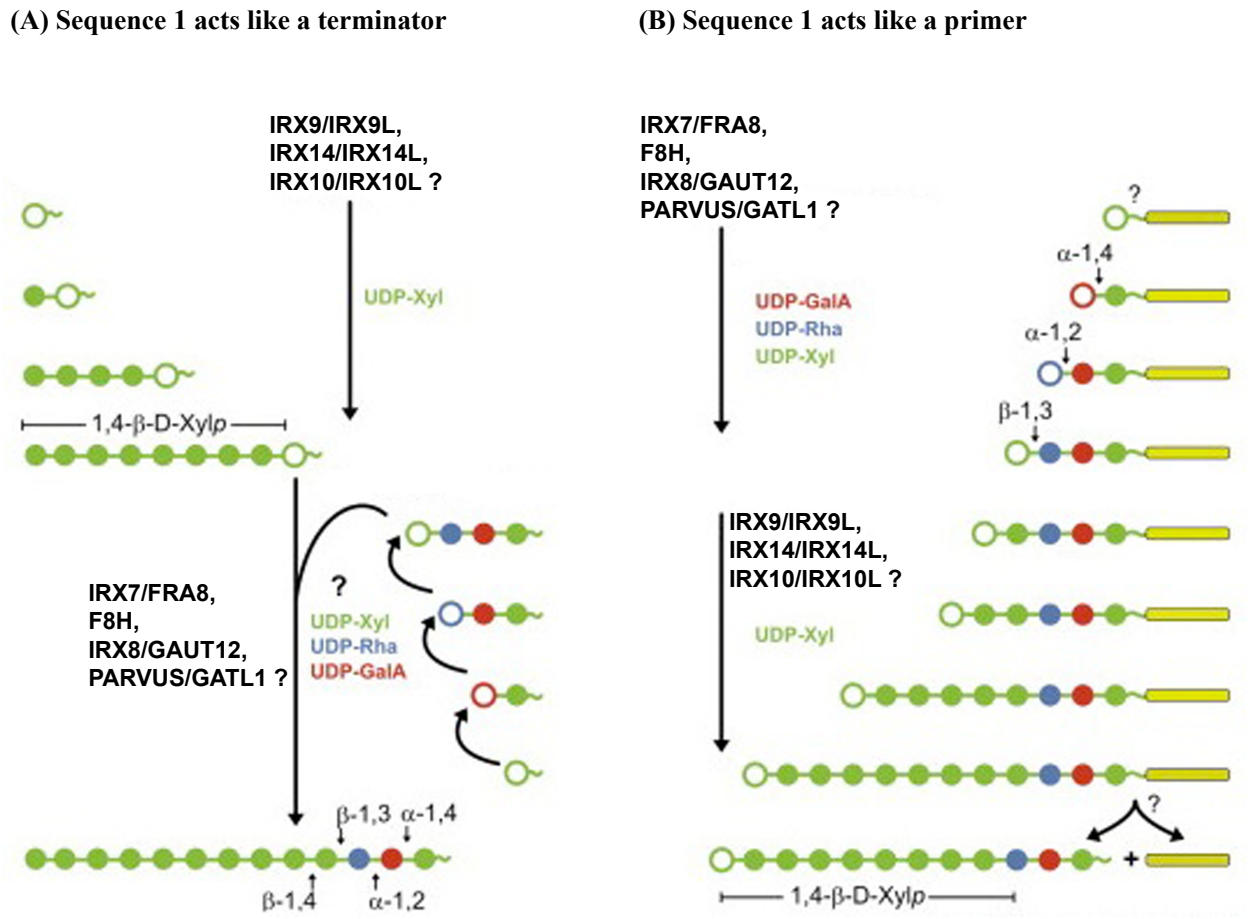


Figure 1.4: Two hypothetical models for GX biosynthesis. The most recently added glycosyl residues are represented by open circles. (A) GX is synthesized by transfer of xylosyl residues to the reducing end of the chain. The elongation process is terminated by transfer of the nascent chain to Sequence **1**. (B) Sequence **1** acts as a primer and xylosyl residues are sequentially added to the nonreducing end (Figure from York and O'Neill. Curr. Opin. Plant Biol. **11**: 258-265.).

Other models have also been suggested that involves synthesis of xylan in blocks. These models are based on the mechanism for the synthesis of bacterial lipopolysaccharide *O*-anitgen and capsular polysaccharides (York and O'Neill 2008, Willis and Whitfield 2013). In these models also, Sequence **1** act as a terminator or a primer to control polymer chain length (**Figure 1.5**). In one model, initial monosaccharide or oligosaccharides residue is attached to a substrate which can be protein, lipid, or another carbohydrate and the elongation occurs by addition of xylose residues to the non-reducing end (**Figure 1.5A**). Once the blocks have reached a certain length, they are attached together to make a polymer. Finally, sequence 1 is transferred onto the polymer thus terminating the synthesis. In another model, sequence 1 is attached to a substrate that initiates the addition of xylose residues to the terminal end (**Figure 5.1B**). Once the block of a particular length is synthesized, an enzyme then releases the block that further gets attached at the terminal end of another block with sequence 1 attached to a substrate. Once a certain length is achieved, the polysaccharide is released with sequence 1 at its reducing end.

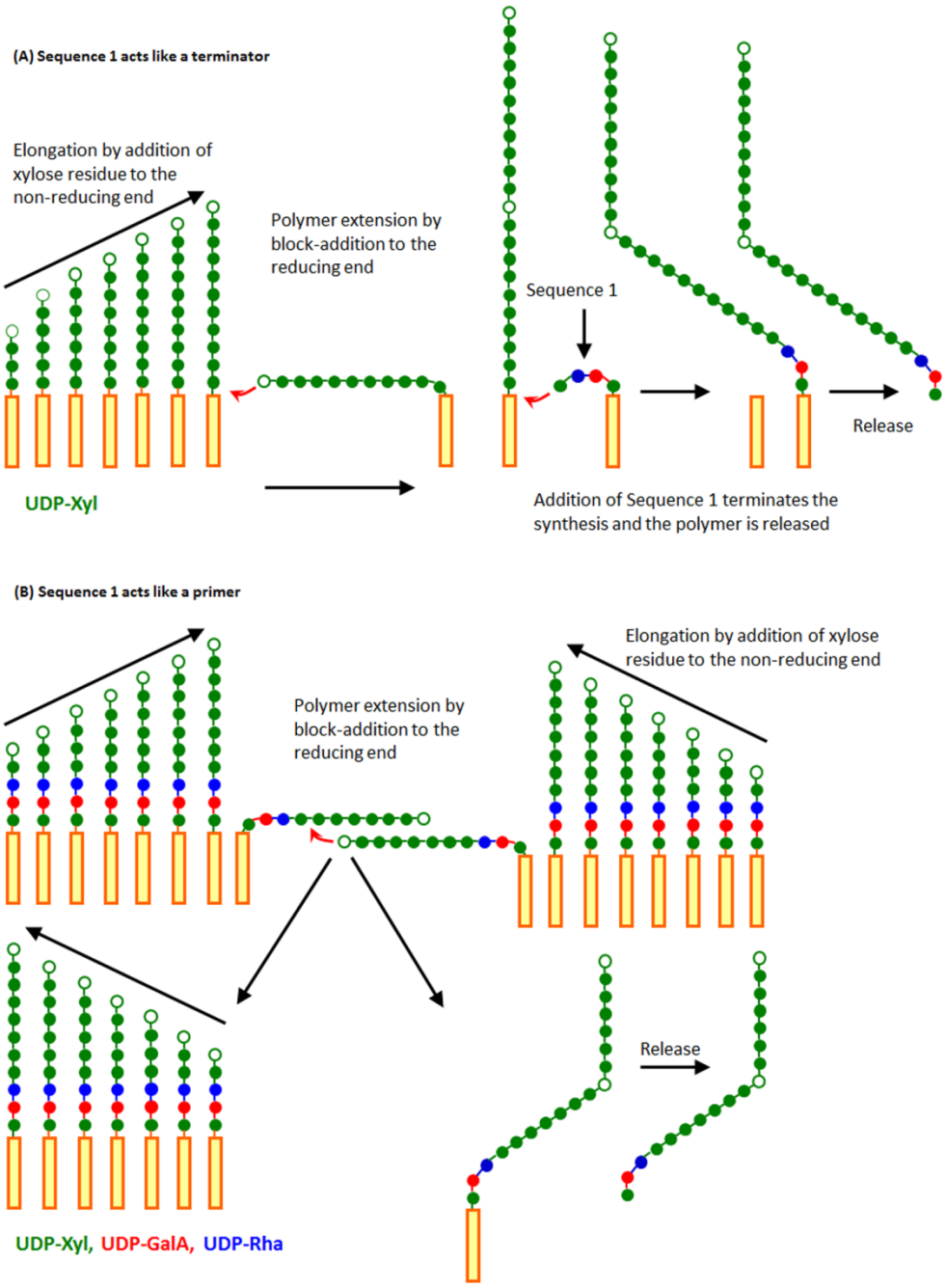


Figure 1.5: Hypothetical models for GX biosynthesis. Adapted from (York W. S. et al 2008 Plant Polysaccharide Workshop, Sigtuna Sweden). Sequence 1 acts like a (A) terminator or like a (B) primer sequence by addition of xylose residues to the non-reducing ends. The most recently added glycosyl residues are represented by open circles.

Sequence **1** has been detected only in seed-bearing land plants (Spermatophyta). I was not able to find these structures in seedless plants such as *P. patens, S. Kraussiana,* and *E. hyemale* (Kulkarni et al. 2012) (See Chapter 2 for details). I also compared xylan structures in five bioenergy grasses (*P. virgatum, Brachypodium, Oryza, Miscanthus* and *foxtail millet*) and provided evidence for the absence of this reducing end sequence in grasses (See Chapter 3 for details). However, the possibility cannot be discounted that grass arabinoxylans contain other unique glycosyl sequences at their reducing end. In my dissertation research, I have compared the structure of heteroxylans in the cells walls of grasses to determine if grasses contain any discernable amounts of sequence **1** at their reducing ends. My study established that grasses lacked discernable amounts of sequence **1** at their reducing ends (See Chapter 3 for details). Therefore I structurally characterized the reducing end sequence(s) of grass xylans (See Chapter 4 for details).

Little is known about xylan structure in monocots other than grasses

Monocots (Liliopsida) are the most diverse, morphologically varied, ecologically successful and economically important taxonomic class of angiosperms (**Figure 1.6**) (Bremer 2000, Givnish and Sytsma 2006, Givnish et al. 2010). Monocots comprise nearly one fourth of all species and families of flowering plants and dominate many terrestrial and aquatic ecosystems. There are approximately 65,000 species of monocots in 82 families. Monocots are typically divided into two categories, namely, commelinids and non-commelinids (**Figure 1.6**). Commelinids comprise the orders Poales, Zingiberales, Commelinales, and Arecales, with Arecales being closely related to non-commelinids. Poales are considered to be the most recently diverged monocots and are thus most distantly related to dicotyledons (Chase et al. 2006,

Givnish and Sytsma 2006, Givnish et al. 2010). The non-commelinid monocots include Asparagales, Liliales, Pandanales, Dioscorerales, Alismatales and Acorales. Acorales are the most closely related to dicotyledons and so are considered the most basal monocots (Chase et al. 2006, Givnish and Sytsma 2006, Givnish et al. 2010). Examples of monocots include the grasses, bromeliads, ginger, cardamom, coconut, banana, onions, asparagus, lilies, palms, and seagrasses. The grasses (for example, barley, wheat and rice) are major components of the human diet and are thus intensively farmed. Several grasses, including switchgrass and miscanthus, are being developed as potential energy crops for the production of biofuels (Givnish et al. 2010).

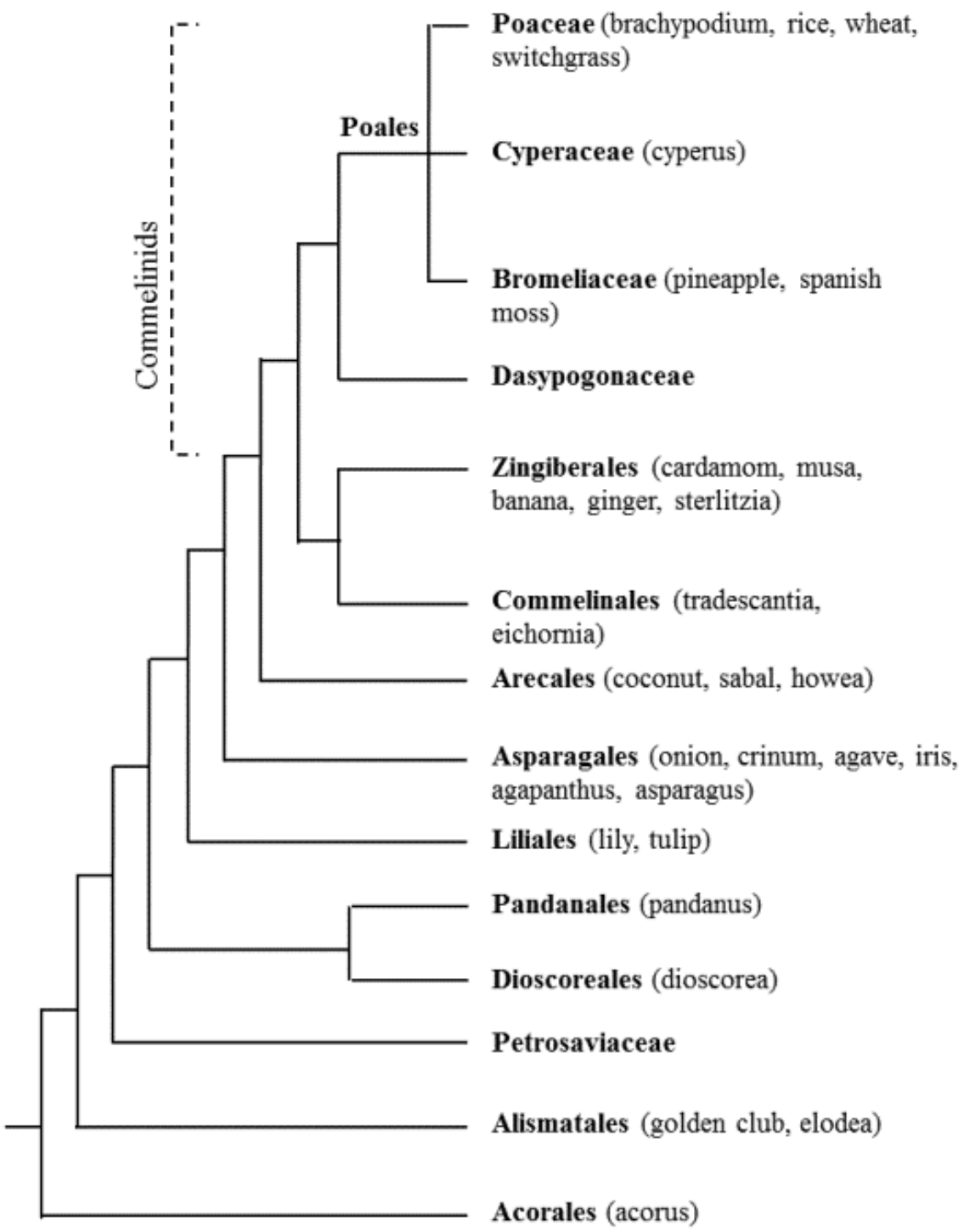


Figure 1.6: Phylogenetic tree showing different orders of monocots with examples of plant species belonging to individual orders. Adapted from (Chase et al. 2006).

Most studies of monocot cell walls have emphasized the grasses (Poaceae), and very little information is available about xylan in the cell walls of other monocots. However, xylans from a few non-grass monocot species have been studied (Buchala and Meier 1972, Jacobs et al. 2003, Bendahou et al. 2007, Scheller and Ulvskov 2010, Simas-Tosin et al. 2013). Structural characterization of heteroxylans from the cell walls of cyperus, pineapple, flax and palms showed that their cell walls contain glucuronoarabinoxylan (Buchala and Meier 1972, Jacobs et al. 2003, Bendahou et al. 2007, Simas-Tosin et al. 2013). Comparison of xyloglucan and pectins from the cell walls of monocotyledons provided evidence for structural diversity in the cell walls within this taxonomic class (Jarvis et al. 1988, Hsieh and Harris 2009). However, few complete characterizations of monocot xylans were previously reported, making an in-depth comparison of xylan structure in different monocots impossible. In my dissertation, I have structurally characterized xylans in the walls of members of the different monocot families and determined if there are any unique glycosyl sequences at their reducing ends. This made it possible to establish a general model for the diversity of xylan structures, including the glycosyl sequences at their reducing ends, across various orders of monocots

Xylan structures vary from plant to plant, tissue to tissue, and between different cell types (Pauly and Keegstra 2008). There is increasing awareness that xylans may impact the value of plant biomass as a source of renewable energy and as a feedstock for the production of value-added chemicals as these polysaccharides are very abundant components of the primary cell walls of grasses and of the secondary cell walls of all angiosperms (Doering et al. 2012). Altering xylan structure or abundance in the cell wall is likely to affect its recalcitrance to deconstruction. A major aim of my research has been to better understand the biological

mechanisms for synthesis and modification of xylan. Such understanding could be used as a rational basis for developing crops that produce biomass with economically advantageous properties.

Xylan biosynthesis in dicots

Several studies have provided evidence that members of glycosyltransferase (GT) families 2, 8, 43, and 47 have a role in the biosynthesis of xylan. A summary of proven and putative proteins involved in xylan biosynthesis is outlined in **Table 1.1**.

Characterization of several *irregular xylem* (*irx*) mutants including *irx9, irx10,* and *irx14* has indicated that these genes encode proteins involved in xylan backbone synthesis (Peña et al. 2007, Brown et al. 2009, Wu et al. 2009, Lee et al. 2010, Wu et al. 2010). All these mutants exhibit a dwarf phenotype along with irregular xylem morphology. Close homologs of these genes (*IRX9L*, *IRX10L*, and *IRX14L*) have also been found to be involved in xylan synthesis (Wu et al. 2009, Keppler and Showalter 2010, Wu et al. 2010). Co-expression of Arabidopsis *IRX9* and *IRX14* in tobacco cells resulted in increased xylosyl transferase activity in microsomal fractions, and led the authors to suggest that these genes may be responsible for the synthesis of the xylan backbone (Lee et al. 2012a). Arabidopsis *irx9 irx9-L*, *irx10 irx10-L*, and *irx14 irx14-L* double mutants showed a more severe phenotype than single mutants and produced less xylan, which had a reduced degree of polymerization, thus suggesting their possible role in xylan backbone synthesis. However, these mutants were seen to retain the reducing end sequence **1** (Wu et al. 2009, Keppler and Showalter 2010, Wu et al. 2010). The drastic decrease in the abundance of xylan in the *irx10 irx10-L*, and *irx14 irx14-L* made it difficult to isolate this

polysaccharide from these double mutants, suggesting these genes have a vital role in xylan synthesis (Wu et al. 2010). Co-expression analysis of poplar *PoGT43B/PtrGT43B*, and *PtrGT43C/D* in tobacco BY2 cells exhibited a high xylosyl transferase activity indicating that they are functional orthologs of *IRX9* and *IRX14* respectively (Zhou et al. 2006, Zhou et al. 2007, Lee et al. 2012b).

IRX15 and IRX15-LIKE proteins have recently emerged as new players in xylan synthesis. The *irx15* and *irx15-like* mutants have irregular xylem vessels and synthesize xylan with a lower degree of polymerization than wild type xylan (Brown et al. 2011, Jensen et al. 2011). Mortimer et al., (2010) showed that GUX 1 and 2 (GlucUronic acid substitution of Xylan) genes are involved in the addition of GlcA and MeGlcA side chains to the xylan backbone in Arabidopsis (Mortimer et al. 2010, Oikawa et al. 2010). The *gux1 gux2* double mutant has virtually undetectable amounts of GlcA side chains and low GlcA GT activity. Recently, a detailed analysis of the *gux* mutants showed differences in the [Me]GlcA substitution patterns (Bromley et al. 2013). It was observed that GUX1 substituted the xylan backbone with [Me]GlcA at intervals of 8 to 10 residues, whereas, GUX2 substituted the xylan backbone more tightly at intervals of 5, 6, and 7 residues (Bromley et al. 2013). As there is no change in the amount of backbone formed by these mutants, the authors concluded that backbone synthesis and the addition of GlcA residues are separate processes. GUX3 and GUX4 have been observed to be homologs of GUX1 and GUX2 and hence have been reported to be xylan glycosyltransferase (Rennie et al. 2012).

Other proteins, including FRA8, F8H, IRX8, and PARVUS/GATL1, have been implicated in the synthesis of the glycosyl reducing end Sequence **1** (Lee et al. 2007b, Peña et al.

2007, York and O'Neill 2008, Lee et al. 2009a, Wu et al. 2010). Arabidopsis plants with mutations in the genes encoding the above proteins have been shown to retain in vitro xylan synthase activity while being depleted in Sequence **1**. *FRA8* and *F8H* have been proposed to be important for chain elongation, and may act in a complex with other xylan-related enzymes (Wu et al. 2010). IRX8 (also known as GAUT12) has been proposed to mediate the formation of the galacturonic acid - xylose linkage in Sequence **1**, as it is a homolog of GAUT1, a galacturonic acid transferase (Sterling et al. 2006). PARVUS was reported to be present in ER, suggesting that it may be a α-xylosyltransferase that transfers the reducing end xylose to a primer (Scheller and Ulvskov 2010) or initiate the formation of reducing end sequence by transferring xylose residue to an acceptor (Lee et al. 2009a). *IRX7* which is a homolog of *F8H* (Lee et al. 2009a), may function *in vitro* as a xylosyltransferase to form the Xyl-Rha linkage in Sequence **1**. Thus, it has been proposed to be a Rha-specific xylosyl transferase (Scheller and Ulvskov 2010), although this activity has not been demonstrated biochemically. It was possible to rescue *fra8* mutant by expression of poplar *PoGT47C*, which was thus proposed to be the *FRA8* ortholog in poplar (Zhou et al. 2006). *PdGATL1.1, PdGATL1.2, PoGT8E*, and *PoGT8F* have been identified as orthologs of *PARVUS* by similar complementation experiments (Kong et al. 2009, Lee et al. 2009b).

An Arabidopsis protein previously classified as domain of unknown function (DUF) 579 was recently shown to be a glucuronoxylan methyltransferase (AtGXMT) that catalyzes the 4-*O*-methylation of glucuronic acid linked to the xylan backbone (Urbanowicz et al. 2012). Xylans and several other polysaccharides are typically *O*-acetylated, which may impact the polymer's susceptibility to enzymatic fragmentation. A four-member gene family (*REDUCED WALL*

ACETYLATION) has been identified in Arabidopsis on the basis of their homology to putative bacterial *O*-acetyl transferases. Plant carrying mutations in the *RWA* genes synthesize wall polysaccharides, including xylan, with reduced levels of *O*-acetylation. However, it remains to be demonstrated that RWA proteins are directly involved in the *O*-acetylation of xylan (Lee et al. 2011, Manabe et al. 2011). Recently, histochemical analysis of *rwa* quadruple mutant showed reduction in *O*-acetylation by 63%, which was correlated to defects in the differentiation of cell with secondary cell walls (Manabe et al. 2013). ESK1/TBL29, whose sequence is not homologous to that of the RWA proteins, was recently identified as a (putative) xylan *O*-acetyltransferase. The *tbl29* mutant exhibits a ~60% reduction in the degree of xylan acetylation (Xiong et al. 2013).

Table 1.1: Summary of proteins known to play a role in xylan biosynthesis.

CAZy Family	Protein name	Protein function and their orthologs	References
GT 43	IRX9, IRX9-L	Involved in **xylan backbone** chain elongation by adding β-(1,4)-D-xylose residues. Complementation studies have indicated that poplar *Pt*GT43B and *Pt*GT43C are functional orthologs of IRX9 and IRX14 respectively. Rice OsIRX9, OsIRX9L and OsIRX14 identified as putative orthologs of respective counterparts by complementation studies.	Brown et al. 2007, Lee et al. 2007, Peña et al. 2007, Lee et al. 2010, Wu et al. 2010, Lee et al. 2012, Chiniquy et al. 2013
	IRX14, IRX14-L		
GT 47	IRX10, IRX10-L		Brown et al. 2009, Wu et al. 2009
-	IRX15, IRX15-L		Brown et al. 2011, Jensen et al. 2011
GT47	IRX7/FRA8	Implicated in the synthesis of the **reducing end sequence**. No biochemical activity has been demonstrated. Poplar *Pd*GATL1.1, *Pd*GATL1.2, *Po*GT8E and *Po*GT8F are identified as functional orthologs of PARVUS.	Peña et al. 2007
	F8H		Lee et al. 2009
GT8	IRX8/GAUT12		Peña et al. 2007
	PARVUS/ GATL1/GLZ1		Brown et al. 2007, Lee et al. 2007, Kong et al. 2009, Lee et al. 2009
GT8	GUX1-5	Involved in the **addition of GlcA and MeGlcA side chains** to the xylan backbone. Over-expression of GUX1-4 in tobacco BY2 cells yielded in xylan:GlcAT activity.	Mortimer et al. 2010, Rennie et al. 2012
GT61	XAT1, XAT2	Knockdown expression in wheat **endosperms decreased α-1,3-Ara*f*** sidechains.	Anders et al. 2102
GT61	XAX1	Involved in **addition of xylose residue to form β-Xyl*p*-(1,2)-α-Ara*f*** sidechain to xylan backbone in rice. Proposed to be involved in feruloylation of grasses.	Chiniquy et al. 2012
-	GXMT	Catalyzes the **4-*O*-methylation of glucuronic acid** linked to the xylan backbone.	Urbanowicz et al. 2012
-	TBL29	Reported to **decrease *O*-acetylation** by ~60%. Mutant observed to have severe collapsed xylem phenotype.	Xiong et al. 2013
-	RWA1-4	Mutants observed to have **reduced levels of *O*-acetylation** in xylan. May not be involved directly in acetylation. No biochemical activity demonstrated.	Manabe et al. 2011, Lee et al. 2011, Manabe et al. 2013

Xylan biosynthesis in grasses

Xylans in grass cell walls have diverse structures that are likely to be modified in response to the changing functional requirements of the wall during growth and development. It is highly likely that the xylan biosynthesis in grasses is catalyzed by a complex of synthases and/or other GT enzymes.

Members of GT families GT43, GT47, GT61, and GT75 have been implicated in xylan biosynthesis (Zeng et al. 2010, Hirano et al. 2013) in grasses. Some GT family 75 enzymes have been proposed to have GAX:GlcAT activity (Dhugga et al. 1997). Recently, over-expression of *OsIRX9*, *OsIRX9L*, and *OsIRX14* was found to partially or fully complement *irx9* and *irx14* mutants in Arabidopsis, suggesting that these genes might function in xylan backbone synthesis in rice (Chiniquy et al. 2013). Alterations in the expression of members of family GT61 showed a reduction in Ara*f* substitution of xylan. Heterologous expression of wheat and rice xylan arabinosyltransferase genes in *Arabidopsis* has also provided gain of function evidence for GT61 family being xylan α-(1,3)-arabinosyltransferases (Anders et al. 2012). Two of the genes (*TaXAT1* and *TaXAT2*), are abundantly expressed in wheat grain (Toole et al. 2010) which was consistent with their involvement in AX biosynthesis. Gain of function was observed by heterologous expression of wheat and rice XAT in Arabidopsis, thus providing evidence for α-(1,3)-arabinosyltransferases activity (Anders et al. 2012). Recently, the rice XAX1 protein (family GT61) was proposed to be involved in adding xylose residues to form a β-Xyl*p*-(1,2)-α-Ara*f* sidechain. (Chiniquy et al. 2012). The rice *xax1* mutant is deficient in ferulic and coumaric acids, suggesting a role for XAX1 in the feruloylation of grasses.

In my dissertation research, I characterized the structures of xylans from taxonomically diverse land plants to identify key structural transitions that occurred during the evolution of these plants (See Chapter 2, 3, and 4). For example, my research has provided insight into the transition from glucuronoxylan to arabinoxylan during monocot evolution. I also established that grasses lacked discernable amounts of Sequence **1** (See Chapter 3 and 4). My research also enabled me to make a predictive model for variation of xylan structure in monocots, providing a basis for understanding the mechanism of xylan biosynthesis in grasses. Xylans have key roles in the assembly of plant cell walls and their structure and abundance almost certainly affect the recalcitrance of these walls to chemical and/or enzymatic deconstruction. Development of advanced technologies for biomass processing will thus depend on understanding xylan structure and biosynthesis in energy crops. My results are thus likely to have important implications for commercial biofuel production from grasses and other crops.

References

Anders N, Wilkinson MD, Lovegrove A, Freeman J, Tryfona T, Pellny TK, Weimar T, Mortimer JC, Stott K, Baker JM*, et al.* 2012. Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proc. Natl. Acad. Sci. U S A* 109:989-993.

Bendahou A, Dufresne A, Kaddami H, Habibi Y. 2007. Isolation and structural characterization of hemicelluloses from palm of *Phoenix dactylifera L. Carbohydr. Polym.* 68:601-608.

Bremer K. 2000. Early Cretaceous lineages of monocot flowering plants. *Proc. Natl. Acad. Sci. U S A* 97:4707-4711.

Bromley JR, Busse-Wicher M, Tryfona T, Mortimer JC, Zhang Z, Brown DM, Dupree P. 2013. GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *Plant J* 74:423-434.

Brown D, Wightman R, Zhang Z, Gomez LD, Atanassov I, Bukowski JP, Tryfona T, McQueen-Mason SJ, Dupree P, 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 J.* 66:401-413.

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

Brown DM, Zhang Z, Stephens E, Dupree P, Turner SR. 2009. Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *Plant J.* 57:732-746.

Buchala A, Meier H. 1972. Hemicelluloses from the stalk of *Cyperus papyrus*. *Phytochemistry* 11:3275-3278.

Carafa A, Duckett JG, Knox JP, Ligrone R. 2005. Distribution of cell-wall xylans in bryophytes and tracheophytes: new insights into basal interrelationships of land plants. *New Phytol.* 168:231-240.

Carpita NC. 1996. Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47:445-476.

Carpita NC, Gibeaut DM. 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3:1-30.

Chase MW, Fay MF, Devey DS, Maurin O, R√∏nsted N, Davies TJ, Pillon Y, Petersen G, Seberg O, Tamura MN. 2006. Multigene analyses of monocot relationships: a summary. *Aliso* 22:63-75.

Chiniquy D, Sharma V, Schultink A, Baidoo EE, Rautengarten C, Cheng K, Carroll A, Ulvskov P, Harholt J, Keasling JD*, et al.* 2012. XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan. *Proc. Natl. Acad. Sci. U S A* 109:17117-17122.

Chiniquy D, Varanasi P, Oh T, Harholt J, Katnelson J, Singh S, Auer M, Simmons B, Adams PD, Scheller HV*, et al.* 2013. Three Novel Rice Genes Closely Related to the Arabidopsis IRX9, IRX9L, and IRX14 Genes and Their Roles in Xylan Biosynthesis. *Front Plant Sci.* 4:83.

Dhugga KS, Tiwari SC, Ray PM. 1997. A reversibly glycosylated polypeptide (RGP1) possibly involved in plant cell wall synthesis: purification, gene cloning, and trans-Golgi localization. *Proc Natl Acad Sci U S A* 94:7679-7684.

Doering A, Lathe R, Persson S. 2012. An update on xylan synthesis. *Mol Plant* 5:769-771.

Ebringerová A, Hromadkova Z, Heinze T. 2005. Hemicellulose. *Adv. Polym. Sci.* 186:1-67.

Faik A. 2010. Xylan biosynthesis: News from the grass. *Plant Physiol.* 153:396-402.

Gibeaut DM, Pauly M, Bacic A, Fincher GB. 2005. Changes in cell wall polysaccharides in developing barley (*Hordeum vulgare*) coleoptiles. *Planta* 221:729-738.

Givnish T, Sytsma K. 2006. Phylogeny of the monocotyledons based on the highly informative plastid gene ndhF: Evidence for widespread concerted convergence. *Monocots: Comparative Biology and Evolution (Excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California*:28-51.

Givnish TJ, Ames M, McNeal JR, McKain MR, Steele PR, dePamphilis CW, Graham SW, Pires JC, Stevenson DW, Zomlefer WB. 2010. Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of Poales 1. *Annals of the Missouri Botanical Garden* 97:584-616.

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315:804-807.

Hirano K, Aya K, Morinaka Y, Nagamatsu S, Sato Y, Antonio BA, Namiki N, Nagamura Y, Matsuoka M. 2013. Survey of genes involved in rice secondary cell wall formation through a co-expression network. *Plant Cell Physiol.* 54:1803-1821.

Hsieh YS, Harris PJ. 2009. Xyloglucans of monocotyledons have diverse structures. *Molecular Plant* 2:943-965.

Jacobs A, Palm M, Zacchi G, Dahlman O. 2003. Isolation and characterization of water-soluble hemicelluloses from flax shive. *Carbohydr. Res.* 338:1869-1876.

Jarvis MC, Forsyth W, Duncan HJ. 1988. A survey of the pectic content of nonlignified monocot cell walls. *Plant Physiol.* 88:309-314.

Jensen JK, Kim H, Cocuron JC, Orler R, Ralph J, Wilkerson CG. 2011. The DUF579 domain containing proteins IRX15 and IRX15-L affect xylan synthesis in Arabidopsis. *Plant J.* 66:387-400.

Jia Z, Qin Q, Darvill AG, York WS. 2003. Structure of the xyloglucan produced by suspension-cultured tomato cells. *Carbohydr. Res.* 338:1197-1208.

Johansson M, Samuelson O. 1977. Reducing end groups in brich xylan and their alkaline degradation. *Wood Sci. Technol.* 11:251-263.

Keppler BD, Showalter AM. 2010. IRX14 and IRX14-LIKE, two glycosyl transferases involved in glucuronoxylan biosynthesis and drought tolerance in Arabidopsis. *Mol. Plant* 3:834-841.

Kong Y, Zhou G, Avci U, Gu X, Jones C, Yin Y, Xu Y, Hahn MG. 2009. Two poplar glycosyltransferase genes, PdGATL1.1 and PdGATL1.2, are functional orthologs to PARVUS/AtGATL1 in Arabidopsis. *Mol. Plant* 2:1040-1050.

Kulkarni AR, Pena MJ, Avci U, Mazumder K, Urbanowicz BR, Pattathil S, Yin Y, O'Neill MA, Roberts AW, Hahn MG*, et al.* 2012. The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes. *Glycobiology* 22:439-451.

Lee C, O'Neill MA, Tsumuraya Y, Darvill AG, Ye ZH. 2007a. The irregular xylem9 mutant is deficient in xylan xylosyltransferase activity. *Plant Cell Physiol.* 48:1624-1634.

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

Lee C, Teng Q, Huang W, Zhong R, Ye ZH. 2009a. The F8H glycosyltransferase is a functional paralog of FRA8 involved in glucuronoxylan biosynthesis in Arabidopsis. *Plant Cell Physiol.* 50:812-827.

Lee C, Teng Q, Huang W, Zhong R, Ye ZH. 2009b. The poplar GT8E and GT8F glycosyltransferases are functional orthologs of Arabidopsis PARVUS involved in glucuronoxylan biosynthesis. *Plant Cell Physiol.* 50:1982-1987.

Lee C, Teng Q, Zhong R, Ye ZH. 2011. The four Arabidopsis reduced wall acetylation genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. *Plant. Cell. Physiol.* 52:1289-1301.

Lee C, Zhong R, Richardson EA, Himmelsbach DS, McPhail BT, Ye ZH. 2007b. The PARVUS gene is expressed in cells undergoing secondary wall thickening and is essential for glucuronoxylan biosynthesis. *Plant Cell Physiol.* 48:1659-1672.

Lee C, Zhong R, Ye ZH. 2012a. Arabidopsis family GT43 members are xylan xylosyltransferases required for the elongation of the xylan backbone. *Plant Cell Physiol.* 53:135-143.

Lee C, Zhong R, Ye ZH. 2012b. Biochemical characterization of xylan xylosyltransferases involved in wood formation in poplar. *Plant Signal Behav* 7:332-337.

Manabe Y, Nafisi M, Verhertbruggen Y, Orfila C, Gille S, Rautengarten C, Cherk C, Marcus SE, Somerville S, Pauly M*, et al.* 2011. Loss-of-function mutation of REDUCED WALL ACETYLATION2 in Arabidopsis leads to reduced cell wall acetylation and increased resistance to Botrytis cinerea. *Plant Physiol.* 155:1068-1078.

Manabe Y, Verhertbruggen Y, Gille S, Harholt J, Chong S-L, Pawar PM, Mellerowicz E, Tenkanen M, Cheng K, Pauly M. 2013. RWA Proteins Play Vital and Distinct Roles in Cell Wall O-Acetylation in Arabidopsis thaliana. *Plant Physiol.*:pp. 113.225193.

Mandal P, Pujol C, Damonte E, Ghosh T, Ray B. 2009. Xylans from Scinaia hatei: Structural features, sulfation and anti-HSV activity. *Int. J. Biol. Macromol.* 46:173-178.

Matulewicz M, Cerezo A. 1987. Alkali-soluble polysaccharides from *Chaetangium fastigiatum*: Structure of a xylan. *Phytochemistry* 26:1033-1035.

Mohnen D, Bar-Peled M, Somerville C. 2008. Cell wall polysaccharide synthesis. In: Himmel ME editor. Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. : Blackwell Publishing, Singapore. p. 94-159.

Moller I, Sorensen I, Bernal AJ, Blaukopf C, Lee K, Obro J, Pettolino F, Roberts A, Mikkelsen JD, Knox JP*, et al.* 2007. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *Plant J.* 50:1118-1128.

Mortimer JC, Miles GP, Brown DM, Zhang Z, Segura MP, Weimar T, Yu X, Seffen KA, Stephens E, Turner SR*, et al.* 2010. Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. *Proc. Natl. Acad. Sci. U S A* 107:17409-17414.

Obel N, Porchia AC, Scheller HV. 2002. Dynamic changes in cell wall polysaccharides during wheat seedling development. *Phytochemistry* 60:603-610.

Oikawa A, Joshi HJ, Rennie EA, Ebert B, Manisseri C, Heazlewood JL, Scheller HV. 2010. An integrative approach to the identification of Arabidopsis and rice genes involved in xylan and secondary wall development. *PLoS One* 5:e15481.

Pauly M, Albersheim P, Darvill A, York WS. 1999. Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. *Plant J* 20:629-639.

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

Pauly M, Keegstra K. 2008. Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J.* 54:559-568.

Pauly M, Keegstra K. 2010. Plant cell wall polymers as precursors for biofuels. *Curr. Opin. Plant Biol.* 13:304-311.

Peña MJ, Darvill AG, Eberhard S, York WS, O'Neill MA. 2008. Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology* 18:891-904.

Peña MJ, Zhong R, Zhou GK, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye ZH. 2007. Arabidopsis irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19:549-563.

Percival EGV, Chanda SK. 1950. The Xylan of Rhodymenia palmata. *Nature* 166:787-788.

Persson S, Wei H, Milne J, Page GP, Somerville CR. 2005. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Natl Acad Sci U S A* 102:8633-8638.

Popper ZA, Fry SC. 2003. Primary cell wall composition of bryophytes and charophytes. *Ann. Bot.* 91:1-12.

Rennie EA, Hansen SF, Baidoo EE, Hadi MZ, Keasling JD, Scheller HV. 2012. Three members of the Arabidopsis glycosyltransferase family 8 are xylan glucuronosyltransferases. *Plant Physiol.* 159:1408-1417.

Sandhu AP, Randhawa GS, Dhugga KS. 2009. Plant cell wall matrix polysaccharide biosynthesis. *Mol. Plant* 2:840-850.

Scheller HV, Ulvskov P. 2010. Hemicelluloses. *Ann. Rev. Plant Biol.* 61:263-289.

Shimizu K, Ishihara M, Ishihara T. 1976. Hemicellulases of brown rotting fungus Tyromyces palustris. II. The oligosaccharides from the hydrolysate of a hardwood xylan by the intracellular xylanase. *Mokuzai Gakkaishi*:618-625.

Simas-Tosin FF, de Souza LM, Wagner R, Pereira GC, Barraza RR, Wendel CF, Sassaki GL, Iacomini M, Gorin PA. 2013. Structural characterization of a glucuronoarabinoxylan from pineapple (*Ananas comosus* (L.) Merrill) gum exudate. *Carbohydr. Polym.* 94:704-711.

Sterling JD, Atmodjo MA, Inwood SE, Kumar Kolli VS, Quigley HF, Hahn MG, Mohnen D. 2006. Functional identification of an Arabidopsis pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proc Natl Acad Sci U S A* 103:5236-5241.

Toole GA, Le Gall G, Colquhoun IJ, Nemeth C, Saulnier L, Lovegrove A, Pellny T, Wilkinson MD, Freeman J, Mitchell RA*, et al.* 2010. Temporal and spatial changes in cell wall composition in developing grains of wheat cv. Hereward. *Planta* 232:677-689.

Urbanowicz BR, Pe√±a MJ, Ratnaparkhe S, Avci U, Backe J, Steet HF, Foston M, Li H, O,ÄôNeill MA, Ragauskas AJ. 2012. 4-O-methylation of glucuronic acid in Arabidopsis glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. *Proc. Natl. Acad. Sci. U S A* 109:14253-14258.

Vogel J. 2008. Unique aspects of the grass cell wall. *Curr. Opin. Plant Biol.* 11:301-307.

Willis LM, Whitfield C. 2013. Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways. *Carbohydr. Res.* 378:35-44.

Wu AM, Hornblad E, Voxeur A, Gerber L, Rihouey C, Lerouge P, 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 Physiol.* 153:542-554.

Wu AM, Rihouey C, Seveno M, Hornblad E, Singh SK, Matsunaga T, Ishii T, Lerouge P, Marchant A. 2009. The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J.* 57:718-731.

Xiong G, Cheng K, Pauly M. 2013. Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the Arabidopsis mutant tbl29. *Mol Plant* 6:1373-1375.

York WS, O'Neill MA. 2008. Biochemical control of xylan biosynthesis - which end is up? *Curr. Opin. Plant Biol.* 11:258-265.

Zeng W, Jiang N, Nadella R, Killen TL, Nadella V, Faik A. 2010. A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. *Plant Physiol.* 154:78-97.

Zhou GK, Zhong R, Himmelsbach DS, McPhail BT, Ye ZH. 2007. Molecular characterization of PoGT8D and PoGT43B, two secondary wall-associated glycosyltransferases in poplar. *Plant Cell Physiol.* 48:689-699.

Zhou GK, Zhong R, Richardson EA, Morrison WH, 3rd, Nairn CJ, Wood-Jones A, Ye ZH. 2006. The poplar glycosyltransferase GT47C is functionally conserved with Arabidopsis Fragile fiber8. *Plant Cell Physiol.* 47:1229-1240.

Chapter 2

THE ABILITY OF LAND PLANTS TO SYNTHESIZE GLUCURONOXYLANS PREDATES THE EVOLUTION OF TRACHEOPHYTES.[1]

[1]Ameya R. Kulkarni, Maria J. Peña, Utku Avci, Koushik Mazumder, Breeanna R. Urbanowicz, Sivakumar Pattathil, Yanbin Yin, Malcolm A. O'Neill, Alison W. Roberts, Michael G. Hahn, Ying Xu, Alan G. Darvill, and William S. York .(2012). The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes. Glycobiology **22**: 439-451. Reprinted here with permission of the publisher.

Abstract

Glucuronoxylans with a backbone of 1,4-linked β-D-xylosyl residues are ubiquitous in the secondary walls of gymnosperms and angiosperms. Xylans have been reported to be present in hornwort cell walls, but their structures have not been determined. By contrast, the presence of xylans in the cell walls of mosses and liverworts remains a subject of debate. Here we present data that unequivocally establishes that the cell walls of leafy tissue and axillary hair cells of the moss *Physcomitrella patens* contain a glucuronoxylan that is structurally homologous to glucuronoxylans in the secondary cell walls of vascular plants. Some of the 1,4-linked β-D-xylopyranosyl residues in the backbone of this glucuronoxylan bear an α-D-glucosyluronic acid (Glc*p*A) sidechain at O-2. By contrast, the lycopodiphyte *Selaginella kraussiana* synthesizes a glucuronoxylan substituted with 4-*O*-Me-α-D-Glc*p*A sidechains, as do many hardwood species. The monilophyte *Equisetum hyemale* produces a glucuronoxylan with both 4-*O*-Me-α-D-Glc*p*A and α-D-Glc*p*A sidechains, as does Arabidopsis. The seedless plant glucuronoxylans contain no discernible amounts of the reducing-end sequence that is characteristic of gymnosperm and eudicot xylans. Phylogenetic studies showed that the *P. patens* genome contains genes with high sequence similarity to Arabidopsis CAZy family GT8, GT43 and GT47 glycosyltransferases that are likely involved in xylan synthesis. We conclude that mosses synthesize glucuronoxylan that is structurally homologous to the glucuronoxylans present in the secondary cell walls of lycopodiophytes, monilophytes, and many seed-bearing plants, and that several of the glycosyltransferases required for glucuronoxylan synthesis evolved before the evolution of tracheophytes.

Introduction

The secondary walls of vascular plants have important roles in specialized cells that provide mechanical support to tissues and in specialized tissues (xylem) that are involved in the movement of water throughout the plant body (Evert 2006). Secondary wall deposition typically begins when a plant cell has ceased to expand and is accompanied by changes in enzyme activities (Dalessandro and Northcote 1977) and gene expression (Aspeborg et al. 2005, Zhong and Ye 2007) that lead to the formation of a wall that is composed predominantly of cellulose, hemicellulose (heteroxylan and/or glucomannan), and lignin (Mellerowicz and Sundberg 2008). The ubiquitous presence of branched 1,4-linked β-D-xylans with glucuronosyl sidechains in the secondary cell walls of vascular plants has led to the suggestion that the ability to synthesize these polysaccharides was a necessary event for the evolution of vascular and mechanical tissues that enabled tracheophytes to fully exploit the terrestrial environment (Carafa et al. 2005). However, the identity of the first land plants that were capable of synthesizing polysaccharides homologous to the glucuronoxylans in the secondary cell walls of vascular plants remains a subject of debate (Carafa et al. 2005, Popper and Tuohy 2010, Sorensen et al. 2010, Popper 2011).

No xylan has been isolated from a bryophyte (liverworts, mosses, and hornworts) and structurally characterized. However, a monoclonal antibody (LM11) that binds to xylan has been reported to label the cell walls of hornwort spores and sporophyte pseudoelators (Carafa et al. 2005). No labeling of liverwort and moss cell walls was observed with LM11 or with LM10, another monoclonal antibody that binds to xylan (Carafa et al. 2005). Based on these results, Carafa et al. (2005) suggested that the ability to synthesize xylan predates the appearance of

vascular plants and that the presence of xylan separates hornworts from the other bryophytes. However, LM10 has been reported to bind, albeit rather weakly, to aqueous buffer and alkali extracts of cell walls from the moss *Physcomitrella patens* (Moller et al. 2007). Small amounts of 4-linked xylose have been detected in the cell walls of *P patens* (Moller et al. 2007) and the moss *Sphagnum novo-zelandicum* (Kremer et al. 2004). Nevertheless, such results by themselves do not establish whether the cell walls of bryophytes contain glucuronoxylans similar to those synthesized by vascular plants.

Xylans from seed-bearing vascular plants (Gymnosperms and angiosperms) have a backbone composed of 1,4-linked β-D-xylopyranosyl (Xyl*p*) residues but differ in the type, location, and number of glycosyl residues attached to this backbone. For example, many eudicots synthesize glucronoxylans that have α-D-glucosyluronic acid (α-D-Glc*p*A) and/or a 4-*O*-methyl α-D-glucosyluronic acid (4-*O*-Me-Glc*p*A) sidechains at O-2 of the backbone residues. Gymnosperms synthesize glucuronoarabinoxylans in which backbone residues are substituted at *O*-2 with 4-*O*-Me-Glc*p*A and at *O*-3 with α-L-arabinofuranosyl (Ara*f*) residues (Ebringerová et al. 2005). The glucuronoxylans of two gymnosperms (spruce [*Picea abies*] and birch [*Betula verrucosa]*) and the eudicot Arabidopsis have been shown to contain the glycosyl sequence 4-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-GalpA-(1,4)-D-Xyl*p* at their reducing ends (Shimizu et al. 1976, Johansson and Samuelson 1977, Peña et al. 2007). The Poaceae (grasses) typically produce arabinoxylans and glucuronoarabinoxylans substituted predominantly with α-L-Ara*f* residues at *O*-2 and/or *O*-3 and less frequently with Glc*p*A and/or 4-*O*-Me-Glc*p*A at *O*-2 (Izydorczyk and Biliaderis 1995, Smith and Harris 1999). The limited data available suggest that

monilophytes (a group of seedless vascular plants) synthesize glucuronoxylans that are substituted at *O*-2 with 4-*O*-Me-Glc*p*A (Bremner and Wilkie 1966).

We now report the results of chemical, biochemical, immunocytochemical, and phylogenetic analyses that together provide compelling evidence that *P. patens* produces a glucuronoxylan that is structurally homologous to glucuronoxylans located in the secondary cell walls of many vascular plants. Thus, the basic machinery required to synthesize this polysaccharide predates the appearance of vascularization in land plants.

Results

Monoclonal antibodies that recognize xylan epitopes label walls of specific Physcomitrella leafy gametophore cells

Moller et al. (2007) have presented evidence indicating that *P. patens* chloronemal filament cell walls contain xylan. Nevertheless, these authors do not explicitly state that the walls of this moss contain branched xylans, and their data do not provide strong evidence for the presence of glucuronoxylans in the cell walls of this plant. To extend these studies and identify *P. patens* tissue that may contain xylan, a series of leafy gametophore cross sections were prepared. Overall cellular organization was visualized using Toluidine blue staining (**Figures 2.1a, 2.1c and S2.1a**) and specific polysaccharide epitopes were localized by immunolabeling (**Figures 2.1b, 1d – h and S1c - h**).

The sections were immunolabeled with several monoclonal antibodies that recognize diverse and distinct xylan epitopes. LM11, a monoclonal antibody that binds to linear and

substituted xylan (McCartney et al. 2005), labeled the walls of pairs of cells (**Figures 2.1b and 2.1d, arrows**) identified as axillary hair cells (Ligrone 1986, Hiwatashi et al. 2001), which are located between leaves and the leafy shoot. High-resolution transmission electron microscopy (**Figure S2.2**) showed that the outermost layer of the axillary hair cells was frequently lifted or separated from the rest of the cell, as described for axillary hair cells in other moss species (Ligrone 1986). A similar pattern of labeling was observed with three additional monoclonal antibodies (CCRC-M147, CCRC-M154, and CCRC-M160) (**Figure S2.1c – e**) that recognize xylan epitopes structurally distinct from that recognized by LM11 (Pattathil et al. 2010). Axillary hair cell walls were also strongly labeled by CCRC-M137 (**Figures 2.1f and 2.1h**), which binds to a xylan epitope distinct from the epitopes recognized by the other four xylan-directed antibodies above (Pattathil et al. 2010). CCRC-M137 labeled leaf cell walls more strongly than did any of the other four anti-xylan antibodies (**Figures 2.1f and 2.1h**). LM10, which binds to unsubstituted xylan (McCartney et al. 2005), did not label the axillary hair cell walls, and labeling of the leaf cell walls was very weak (**Figure S2.1f**). Leafy shoot cells were not labeled by any of the xylan-directed monoclonal antibodies used in this study (**Figures 2.1b, 2.1d, 2.1f, 2.1h, and S2.1c – f**).

The *P. patens* sections were also labeled with monoclonal antibodies that recognize other polysaccharides (non-fucosylated xyloglucan, de-esterified pectin, and rhamnogalacturonan I) known to be present in Physcomitrella cell walls (Moller et al. 2007; Peña et al. 2008) in part as a control to ensure that all walls in the sections were accessible to antibodies. CCRC-M88, which binds to a non-fucosylated xyloglucan epitope (Pattathil et al. 2010), strongly labeled leafy shoot cell walls (**Figures 2.1e and 2.1g**) and leaf cells to a lesser extent. CCRC-M38,

which recognizes de-esterified pectin (unpublished results of the authors) strongly labeled leaf cell walls and weakly labeled leafy shoot walls (**Figure S2.1g**), while CCRC-M35, which recognizes the rhamnogalacturonan I backbone (Young et al. 2008) weakly labeled the cell walls of both leafy shoots and leaves (**Figures S2.1h**).

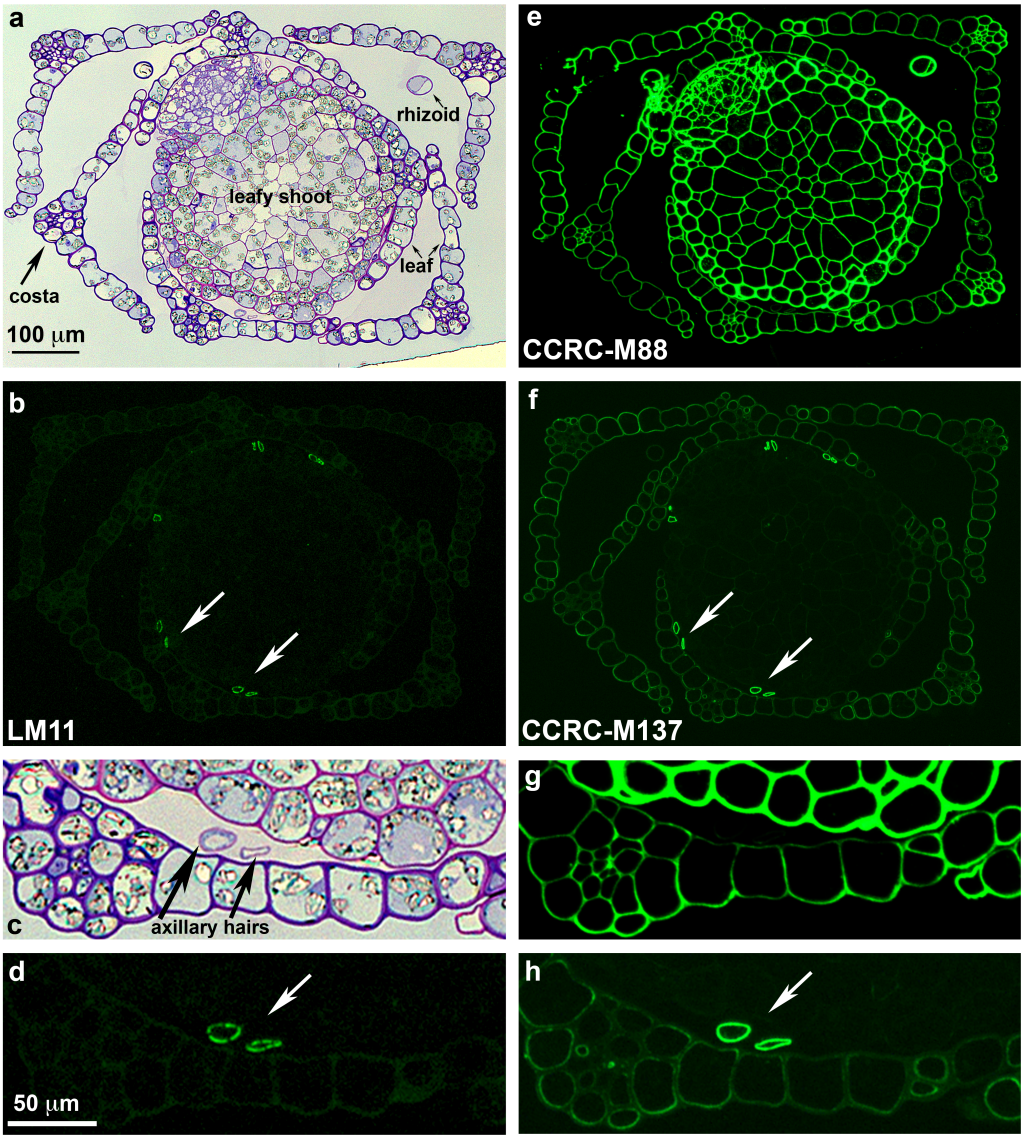


Figure 2.1. Immunofluorescence light microscope images of cross sections of *P.patens* leafy gametophores. Epitopes in serial cross-sections were visualized by their interactions with specific monoclonal antibodies. The scale for panels a, b, e and f is 100 μm as indicated by the bar in a. The scale for panels c, d, g and h is 50 μm as indicated by the bar in d.

(a) Toluidine blue stained cross section.

(b) LM11 that recognizes linear and substituted xylans specifically labeled axillary hair cell walls (arrows pointing to four axillary hairs).

(c) Magnified portion of panel a showing two typical axillary hair cells (arrows).

(d) Magnified portion of panel b showing LM11 labeling of the two axillary hair cells (arrow).

(e) CCRC-M88, specific for non-fucosylated xyloglucan, labeled all cell walls but labeling was weaker in axillary hair cell walls.

(f) CCRC-M137 that recognizes xylan labeled axillary hair cell walls (arrows) more strongly than leaf cell walls.

(g) Magnified portion of panel e showing the weak labeling of CCRC-M88 in the two axillary hair cells.

(h) Magnified portion of panel f showing CCRC-M137 labeled the two axillary hair cell walls (arrow). Labeling was weaker in leaf cell walls.

Structural characterization of the glucuronoxylan in cell walls of P. patens leafy gametophores

Previous studies have shown that glucuronoxylan is solubilized by treating vascular plant cell walls with alkali (Ebringerová et al. 2005, Zhong et al. 2005). Thus, the de-starched alcohol insoluble residue (AIR) generated from *P. patens* leafy gametophores was sequentially extracted with ammonium oxalate, 1M KOH, 4M KOH, chlorite, post chlorite 4M KOH and 5M KOH containing 4% (w/v) boric acid. Immunological glycome profiling (**Figure S2.3**) suggested that epitopes recognized by monoclonal antibodies that bind to xylan are more abundant in the 4M KOH extract than in the 1M KOH extract. However, this fraction is also rich in epitopes recognized by monoclonal antibodies that bind to xyloglucan and pectic polysaccharides. Glycosyl-linkage composition analysis also revealed that derivatives of 1,4-linked and 1,2,4-linked Xyl*p* residues are abundant in the 4M KOH extract (**Figure S2.4**). These data are consistent with the results of Moller et al. (2007) and suggest that Physcomitrella cell walls contain a branched xylan that is more difficult to extract than the branched xylan in vascular plant cell walls, which is efficiently solubilized by treatment with 1M KOH.

The results described above led us to perform detailed analyses of the 4M KOH soluble materials, which provided chemical and spectroscopic evidence for the presence of glucuronoxylan in *P. patens*. The 4M KOH extract was treated with an endo-xylanase to generate oligosaccharides. The high-molecular weight, xylanase-resistant material was precipitated by the addition of ethanol (to 60% v/v) and the ethanol-soluble products were then separated by size-exclusion chromatography (SEC). The oligosaccharide-containing fractions were collected and analyzed by matrix-assisted laser-desorption ionization time-of-flight mass

spectrometry (MALDI-TOF-MS), one and two dimensional ^{1}H NMR spectroscopy, and electrospray ionization multiple mass spectrometry (ESI-MS^{n}).

Virtually all of the detected oligosaccharides generated by xylanase treatment of the 4M KOH soluble extract were shown by MALDI-TOF-MS (**Figure S2.5**) to have molecular weights greater than 1000 daltons and to contain from seven to nine pentosyl residues together with one or two hexuronosyl residues. Lower mass ions are much less abundant in this spectrum or, in the case of di- and tri-saccharides, obscurred by matrix ions.

The purified, xylanase-generated *P. patens* oligosaccharides were further characterized by ^{1}H NMR spectroscopy. The 2D gCOSY spectrum (**Figure 2.2a**) provided chemical shift and scalar coupling information that, in combination with previously published data (Verbruggen et al. 1998, Peña et al. 2007), allowed the anomeric and ring proton resonances to be assigned for the terminal non-reducing β-D-xylosyl, internal 4-linked β-D-xylosyl, internal 2,4-linked β-D-xylosyl and reducing xylosyl residues as well as α-D-Glc*p*A residues linked to O2 of the xylosyl backbone (residues A-G, **Table 2.1, Figure 2.2A**). The downfield shift of H-2 of β-Xyl*p* residue F (δ 3.482) relative to H-2 of unbranched β-Xyl*p* residues B-F (δ 3.25-3.29) confirmed that the xylan backbone is substituted at O-2 by α-D-Glc*p*A (Peña et al. 2007). No resonance that could be assigned to 4-*O*-Me Glc*p*A, Ara*f* sidechains attached to Xyl*p* residues or Xyl*p* residues bearing Ara*f* sidechains were detected in the 2D gCOSY NMR. Resonances with chemical shifts corresponding to branched (pectic) arabinans (Cartmell et al. 2011) were detected in the ^{1}H-NMR spectra of the crude endoxylanase-treated 4 M KOH extract, but these were not observed in the ^{1}H-NMR spectra of the purified glucuronoxylan oligomers. (**See Figure S2.6**). Thus, the Ara*f* residues detected by glycosyl-linkage analysis of the 4 M KOH extract from which the

oligosaccharides were prepared are most likely components of pectic polysaccharides present in this extract (**Figure S2.3**). These data suggest that Ara*f* sidechains, if present at all, are a minor component of the *P. patens* glucuronoxylan.

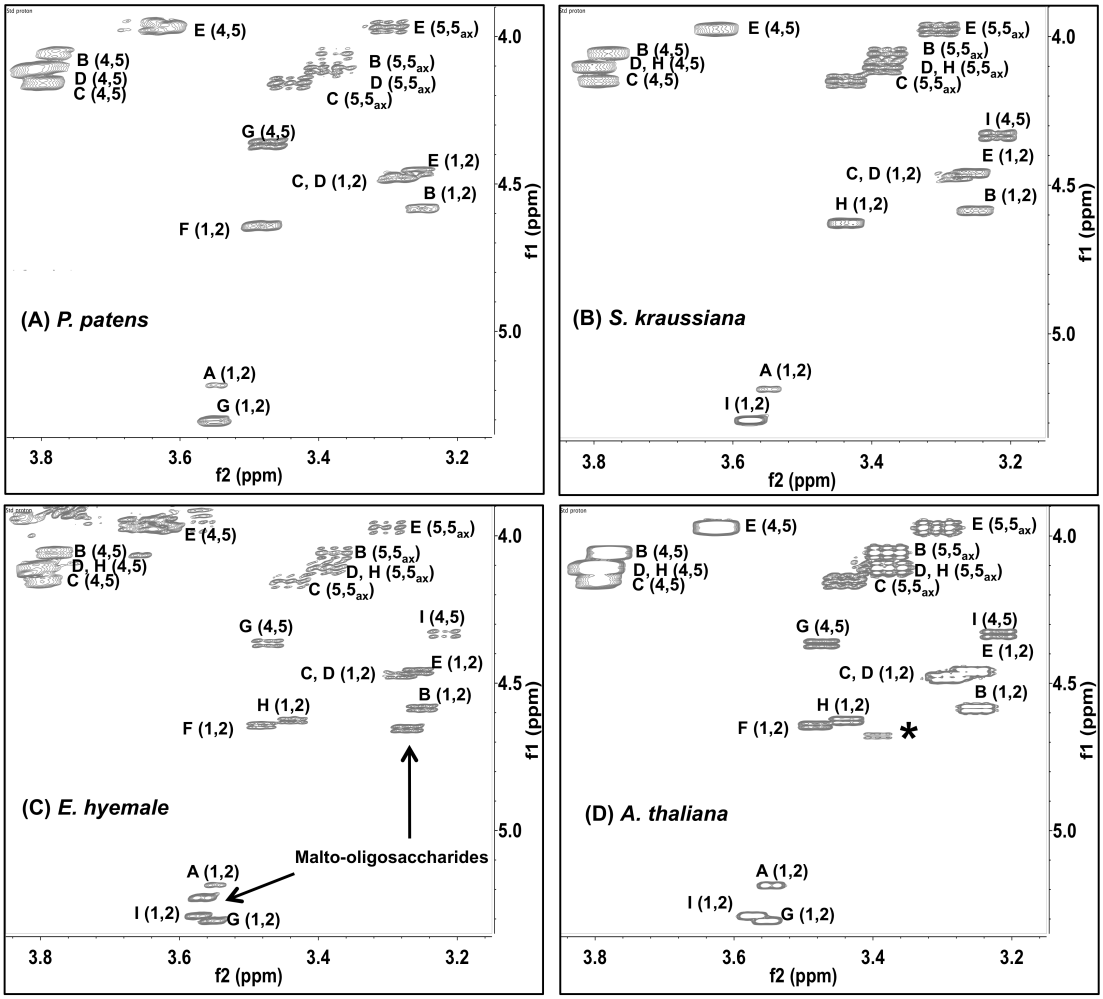


Figure 2.2: Partial 600-MHz gCOSY NMR spectrum of purified xylo-oligosaccharides generated by β-endoxylanase digestion of the 4M KOH extracts of AIR from (A) *P. patens* gametophores, (B) *S. kraussiana* sporophytes, (C) *E. hyemale* sporophytes and (D) *A. thaliana* stems. Crosspeak assignments are indicated using an uppercase letter to indicate the glycosyl residue that contains the protons (Tables 1-3) and numbers indicating the position of the protons in the residue. Resonances due to contaminating malto-oligosaccharides are also labeled in the *E. hyemale* spectrum and a crosspeak (marked with an asterisk) due to the presence of the reducing end sequence of the glucuronoxylan from *A. thaliana* are also indicated.

Table 2.1. ^{1}H NMR assignments of the xylo-oligosaccharides generated by endoxylanase treatment of the 4 M KOH extract of AIR from *P. patens* gametophores

Key	Residue	H1	H2	H3	H4	$H5_{ax}$	$H5_{eq}$
A	α-Xyl*p* (reducing)	5.186	3.546	3.787	---	---	---
B	β-Xyl*p* (reducing)	4.584	3.251	3.546	3.7829	3.377	4.058
C	β-1,4-Xyl*p* (internal) (major)	4.474	3.283	3.578	3.795	3.439	4.155
D	β-1,4-Xyl*p* (internal) (minor)	4.47	3.29	3.56	3.793	3.377	4.107
E	β-Xyl*p* (terminal)	4.457	3.254	3.426	3.624	3.296	3.964
F	β-1,2,4-Xyl*p* (α-Glc*p*A)	4.642	3.482	3.635	3.812	3.391	4.11
G	α-Glc*p*A	5.305	3.552	3.733	3.473	4.361	---

Chemical shifts are reported in ppm relative to internal acetone, δ 2.225. β-Xyl (α-Glc*p*A) is a β-linked xylosyl residue that bears a Glc*p*A sidechain at O-2. H-4 and H-5 of the reducing α-xylose were not assigned. Chemical shifts of protons with overlapping resonances are given to two decimal places. Residues are indicated by an uppercase letter as a key for cross referencing with Figure 2A.

To gain insight into the distribution of the sidechain residues along the xylan backbone, the xylo-oligosaccharides were per-*O*-methylated and then analyzed by ESI-MS^n. Examination of the resulting spectra indicated that each quasimolecular ion that was selected for fragmentation corresponded to the presence of several oligosaccharide isomers. A detailed discussion of the fragmentation pathways leading to this conclusion is given in Supplemental Information. One diagnostic fragmentation pathway (*m/z* 1465 – 1291 – 913 – 753 – 375) provides strong evidence that the most abundant structure corresponding to the quasimolecular ion at *m/z* 1465 (P_6G_2) is an oligosaccharide with two hexuronosyl sidechains separated by a single xylosyl residue (**Figure 2.3**). MS^n of the low-abundance quasimolecular ion at *m/z* 1407 (P_7G, Figure S7) revealed a fragmentation pathway (*m/z* 1407 – 1233 – 1059) that occurs by two sequential losses of 174 Da. This is consistent with the presence of oligosaccharides with a pentosyl sidechain (see Supplemental Information), which may make these quantitatively minor oligosaccharides resistant to further fragmentation by the endoxylanase even though they only have a single glucuronic acid sidechain. Insufficient material was available to fully characterize the pentosyl sidechain. Nevertheless, ^{1}H-NMR analysis indicates that the purified *P. patens* endoxylanase-generated oligosaccharides contain few, if any, arabinofuranosyl sidechains.

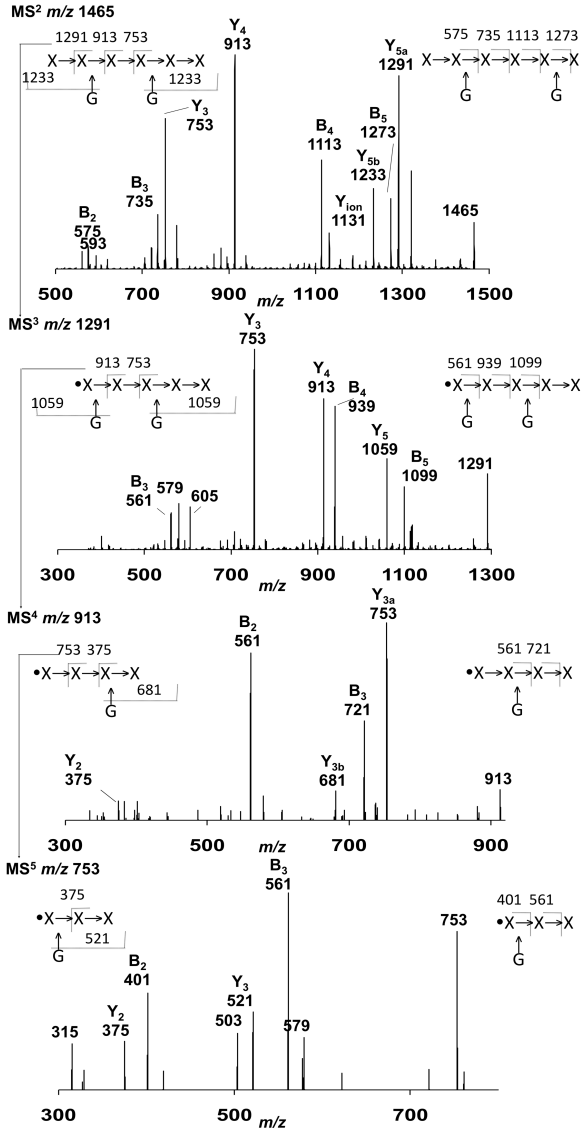


Figure 2.3: ESI-MS^n of the *m/z* 1465 precursor ion of oligosaccharides prepared by xylanase-treatment of the 4M KOH extract of *P. patens* AIR and subsequently per-*O*-methylated. The fragmentation pathway (*m/z* 1465 – 1291 – 913 – 753 – 375) illustrated provides strong evidence for the glycosyl sequence shown in the top panel, although the spectra also reveal the presence of other sequences. Fragmentation events leading to Y-ions are shown on the left side of each spectrum and events leading to B-ions are shown on the right side of each spectrum.

Xylans from seedless vascular plants and angiosperms are structurally similar

To provide an evolutionary context for our structural analysis of *P. patens* glucuronoxylan, we characterized the glucuronoxylans solubilized by 1 M KOH treatment of the cell walls of two seedless vascular plants - *S. kraussiana* (a lycopodiophyte) and *E. hyemale* (a monilophyte). The material solubilized from the AIR of these plants with 1M KOH was treated with an endo-xylanase to generate oligosaccharides, which were partially purified by SEC and analyzed by ^{1}H-NMR spectroscopy. The 2D gCOSY NMR spectrum of the *S. kraussiana* xylo-oligosaccharides (**Figure 2.2B, Table 2.2**) revealed resonances with chemical shifts and scalar coupling patterns (Peña et al. 2007) diagnostic for the presence of 1,4-linked β-D-Xyl*p* residues, some of which are substituted at O-2 with 4-*O*-Me-α-D-Glc*p*A. No resonances indicating the presence of (unmethylated) α-D-Glc*p*A sidechains were observed in this spectrum. The 2D gCOSY spectrum of *E. hyemale* xylo-oligosaccharides (**Figure 2.2C, Table 2.3**) contained resonances diagnostic for the presence of both 4-*O*-Me-α-D-Glc*p*A and α-D-Glc*p*A sidechains at O-2 of 1,2,4-linked β-D-Xyl*p* residues (Peña et al. 2007). This spectrum is similar to the gCOSY spectrum of the glucuronoxylan oligosaccharides prepared from wild-type Arabidopsis stems (Figure 2D), which have 4-*O*-Me-α-D-Glc*p*A and α-D-Glc*p*A sidechains (Peña et al. 2007)

Table 2.2. ^{1}H NMR assignments of the xylo-oligosaccharides generated by endoxylanase treatment of the 1 M KOH extract of AIR from Selaginella.

Key	Residue	H1	H2	H3	H4	$H5_{ax}$	$H5_{eq}$
A	α-Xyl*p* (reducing)	5.184	3.545	3.783			
B	β-Xyl*p* (reducing)	4.584	3.251	3.546	3.778	3.374	4.056
C	β-1,4-Xyl*p* (internal) (major)	4.469	3.276	3.570	3.791	3.435	4.146
D	β-1,4-Xyl*p* (internal) (minor)	4.477	3.288	3.55	3.79	3.37	4.103
E	β-Xyl*p* (terminal)	4.457	3.254	3.427	3.623	3.300	3.975
H	β-1,2,4-Xyl*p* (α-4Me-Glc*p*A)	4.625	3.435	3.620	3.803	3.382	4.102
I	α-4Me-Glc*p*A	5.291	3.574	3.758	3.214	4.330	---

See footnotes of Table I, except data here refers to the spectrum illustrated in Figure 2B.

Table 2.3. ^{1}H NMR assignments of the xylo-oligosaccharides generated by endoxylanase treatment of the 4 M KOH extract of AIR from Equisetum.

Key	Residue	H1	H2	H3	H4	H5	$H5_{ax}$
A	α-Xyl*p* (reducing)	5.183	3.543	3.77			
B	β-Xyl*p* (reducing)	4.583	3.251	3.546	3.778	3.376	4.055
C	β-1,4-Xyl*p* (internal) (major)	4.473	3.279	3.575	3.797	3.439	4.149
D	β-1,4-Xyl*p* (internal) (minor)	4.470	3.282	3.56	3.79	3.37	4.10
E	β-Xyl*p* (terminal)	4.457	3.257	3.427	3.622	3.301	3.970
F	β-1,2,4-Xyl*p* (α-Glc*p*A)	4.643	3.480	3.635	3.810	3.390	4.109
G	α-Glc*p*A	5.305	3.552	3.738	3.470	4.363	---
H	β-1,2,4-Xyl*p*(α-4Me-Glc*p*A)	4.627	3.436	3.623	3.803	3.383	4.102
I	α-4Me-Glc*p*A	5.290	3.571	3.763	3.216	4.332	---

See footnotes of Table I, except data here refers to the spectrum illustrated in Figure 2C.

Two gymnosperms (spruce and birch) and the eudicot Arabidopsis have been shown to synthesize xylans with the glycosyl sequence 4-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-GalpA-(1,4)-D-Xyl*p* at their reducing end (Shimizu et al. 1976, Johansson and Samuelson 1977, Peña et al. 2007). Resonances (marked with an asterisk in Figure 2D) that are diagnostic for this glycosyl sequence are clearly visible in the 2D NMR spectra of the *A. thaliana* xylo-oligosaccharides, but these resonance are not discernible in the spectra of the *S. kraussiana, E. hyemale* or *P. patens* xylo-oligosaccharides. Thus, this glycosyl sequence is absent, or present in amounts below our detection limits, in the xylans of these seedless plants.

The Physcomitrella genome contains putative orthologs of glycosyltransferase genes implicated in xylan biosynthesis

A combination of molecular and biochemical studies have identified numerous Arabidopsis and Poplar genes encoding glycosyltransferases that are likely to participate in xylan biosynthesis in secondary walls. These include members of CAZy families GT8, GT43 and GT47 (Zhong and Ye 2003, Zhong et al. 2005, Zhou et al. 2006, Brown et al. 2009). *IRX8* (also known as *GAUT12*, At5g54690) and *PARVUS* (also known as *GATL1*, At1g19300) encode family GT8 proteins that have been implicated in the synthesis of the glucuronoxylan reducing end sequence (Peña et al. 2007, Kong et al. 2009, Lee et al. 2009a). Two genes, referred to as *GUX1* (At3g18660) and *GUX2* (At4g33330), encode family GT8 enzymes that have been implicated in the attachment of GlcA and 4-*O*-Me-GlcA to the xylan backbone (Mortimer et al. 2010). Four family GT43 members, *IRX9* (At2g37090) *IRX9L* (At1g27600), *IRX14* (At4g36890) and *IRX14L* (At5g67230) are likely to have roles in xylan backbone synthesis

(Peña et al. 2007, Lee et al. 2010b, Wu et al. 2010). Three genes, *IRX7* (*FRA8*, At2g28110), *IRX10* (*GUT2*, At1g27440) and *IRX10L* (*GUT1*, At5g61840), encode GT47 enzymes that have also been implicated in xylan synthesis. Other genes (*IRX10* and *IRX10L*) may function in backbone synthesis (Wu et al. 2009), whereas *IRX7* may be involved in the synthesis of the glucuronoxylan reducing end sequence (Lee et al. 2010b). Family GT8 has recently been the subject of a detailed phylogenetic analysis (Yin et al. 2010), which reveals three potential orthologs of *IRX8* and five potential orthologs of *PARVUS* in the *P. patens* genome (**Table S2.2**).

We generated GT43 and GT47 phylogenies based on the amino acid sequences deduced from the genomes of 10 land plants and six green algae (**see Table S2.1**). The tree generated for GT family 43 is rooted with a green algal sequence (C) and consists of two clades (A and B), which contain only land plant genes (**Figure 2.4**). All of the land plants examined, including the moss *P. patens* and the lycophyte *S. moellendorffii*, have putative orthologs of IRX14 in GT43 clade A (**Figure 2.4, Table S2.2**). However, it seems unlikely that the GT43 clade B3 genes of *S. moellendorfii* and *P.patens* are orthologous to *IRX9*, which resides in a different clade (B1). Four major clades (A-D) were identified for the family GT47 glycosyltranseferases (**Figure S2.8**). The *IRX7* and *IRX10* genes are located in GT47 clade D1, as are the potential *P. patens* and *S. moellendorffii* orthologs of these genes (**Figure 2.5, Table S2.2**).

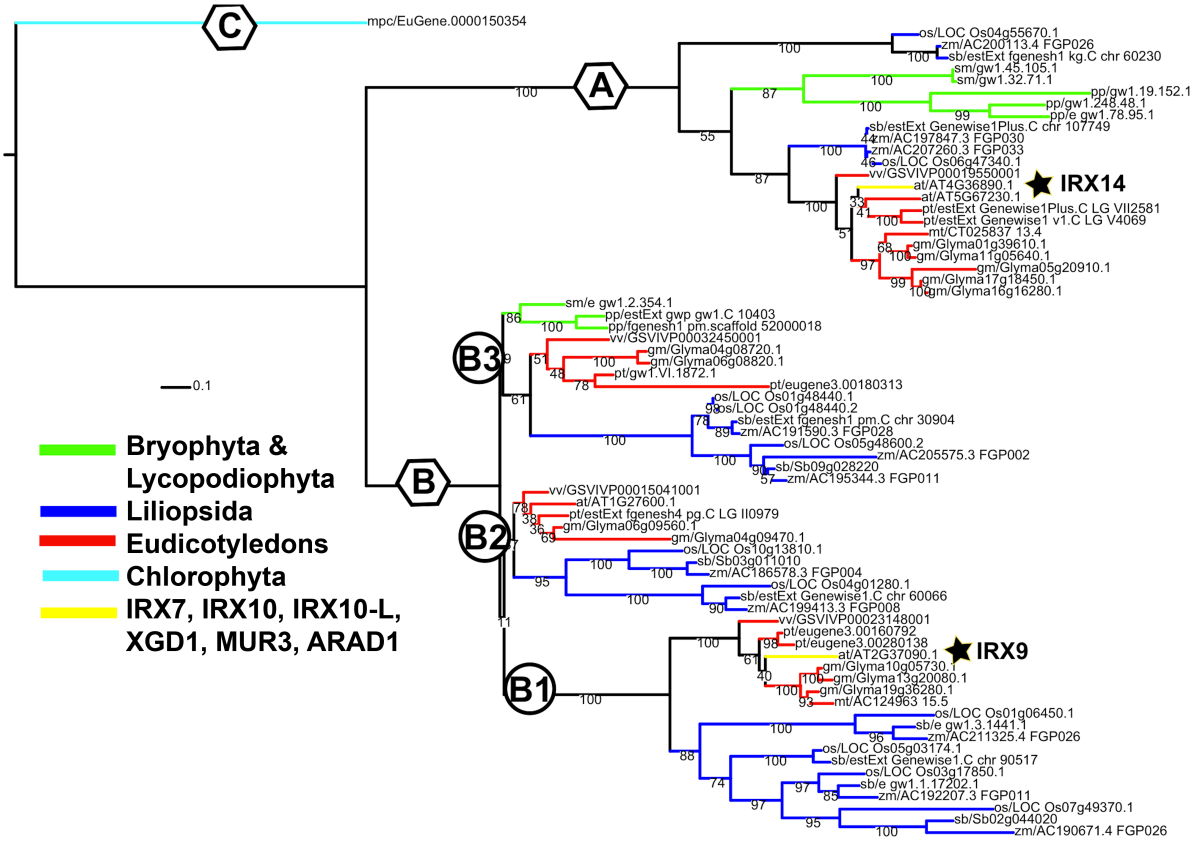


Figure 2.4: The maximum likelihood phylogeny of GT43 family proteins from 16 plant genomes. Multiple sequence alignment (MSA) of the conserved Pfam GT43 domain was performed using MAFFT v6.603 (Katoh et al. 2005b) using L-INS-I. The phylogeny was reconstructed using the PhyML v2.4.4. Clade A includes potential orthologs of *A. thaliana* IRX14 (At4g36890) and Clade B includes homologs of *A. thaliana* IRX9 (At2g37090).

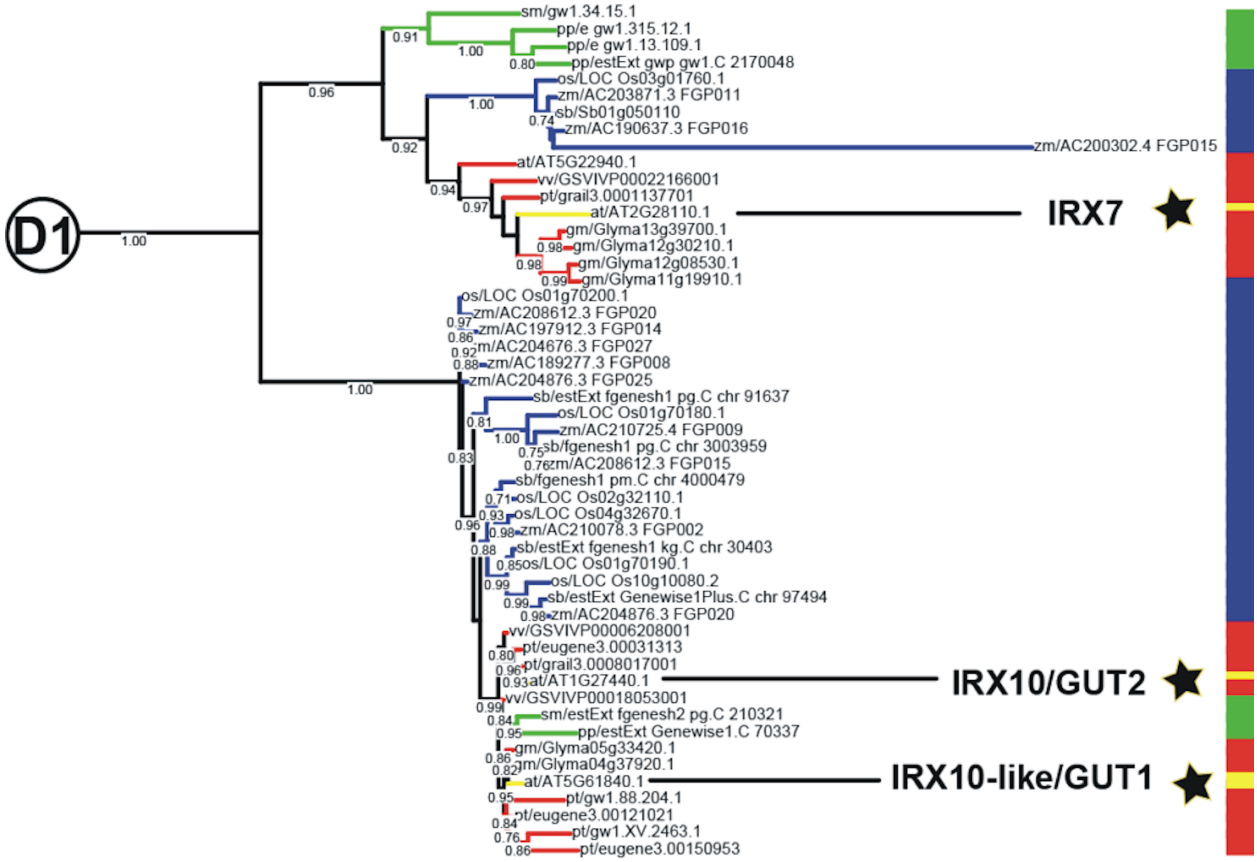


Figure 2.5: The maximum likelihood phylogeny of clade D1 of GT47 family proteins from 16 plant genomes. (See Figure S8 for more details.). Clade D1 includes potential orthologs of three *A. thaliana* proteins, IRX7 (FRA8, At2g28110), IRX10 (GUT2, At1g27440) and IRX10-like (GUT1, At5g61840).

Discussion

Bryophytes are a diverse group of avascular land plants that includes mosses (Bryophyta), liverworts (Marchantiophyta) and hornworts (Anthocerotophyta). Extant members of this paraphyletic group are believed to be the closest living relatives of the first plants to adapt to life on the land about 450 million years ago (Qiu et al. 2006, Mishler and Kelch 2009). Subsequent evolutionary innovations led to the appearance of vascular plants (tracheophytes) which diverged from the bryophytes around 420 million years ago (Niklas 1997, Graham et al. 2000, Taylor et al. 2009). Some of these evolutionary innovations are believed to have involved changes in the structure and composition of the plant cell wall (Popper and Fry 2003, Matsunaga et al. 2004, Carafa et al. 2005, Peña et al. 2008, Popper 2008, Popper 2011). For example, Carafa et al. (2005) have hypothesized that the ability to form secondary walls that contain xylans was one of the factors that facilitated the evolution of vascular and mechanical tissues. Nevertheless, evolution of the structural features of xylan that enable them to perform their biological functions in vascular plants is poorly understood and the identity of the first plants that were capable of synthesizing xylans with these specific features remains a subject of debate (Carafa et al. 2005).

Xylans composed of β-linked xylosyl residue are not restricted to land plants as they are also present in the cell walls of several red algae (Rhodophytes), although none of these algal xylans has been shown to be substituted with side chains composed of Glc*p*A, 4-*O*-Me-Glc*p*A, or Ara*f* residues. For example, the cell walls of *Chaetangium fastigiatum* and *Scinaia hatei* have been reported to contain linear 1,4-linked xylans (Matulewicz and Cerezo 1987, Mandal et al. 2009b, Mandal et al. 2009a), whereas *Rhodymenia palmata* synthesizes a linear xylan composed

of 1,3- and 1,4-linked Xyl*p* (Percival and Chanda 1950a). The red alga *Porphyra umbilicalis* may synthesize both 1,4-linked and 1,3-/1,4-linked xylans (Turvey and Williams 1970). The Rhodophytes and the green plant lineage have been estimated to have diverged about 1,500 million years ago (Yoon et al. 2004). Moreover, none of the genes encoding the red algal xylan synthases have been identified. Thus, it is not known whether the ability to synthesize xylans was inherited from an ancestor common to red algae and green plants or arose by convergent evolution. Linear xylans composed entirely of 1,3-linked β-D-Xyl residues are also synthesized by the chlorophyte Caulerpa (Atkins et al. 1969, Yamagaki et al. 1997). The chlorophyte algae and streptophyte lineage of green plants are estimated to have diverged between 725 and 1200 million years ago (Becker and Marin 2009), but again the lack of relevant genomic data limits our knowledge of the evolutionary relationship between the xylans synthesized by land plants and by the chlorophyta. The presence of small amounts 4-linked β-D-xylans in the cell walls of several evolutionarily advanced charophycean green algae has been inferred from data obtained using antibody-based glycome profiling and glycosyl-linkage composition analyses (Domozych et al. 2009, Sorensen et al. 2010). However, no detailed chemical and spectroscopic data has been published to show that these green algae synthesize xylans comparable with the xylans of land plants.

Our immunological, chemical and spectroscopic data provide evidence that the cell walls of *P. patens* contain glucuronoxylans with a 1,4-linked β-D-xylan backbone substituted with α-D-Glc*p*A sidechains. Thus, the *P. patens* glucuronoxylan is structurally similar to glucuronoxylans produced by vascular plants, but is distinguished from them by the absence of 4-*O*-Me-α-D-Glc*p*A sidechains, which are ubiquitous in the secondary cell wall glucuronoxylans of vascular

plants (Ebringerová et al. 2005). Our data also suggests that the *P. patens* glucuronoxylan backbone is occasionally substituted with an as yet unidentified pentosyl residue. However, further studies are required to substantiate this claim. The distribution of Glc*p*A sidechains in the moss glucuronoxylan is also unusual, with pairs of Glc*p*A sidechains separated by a single xylosyl residue. This branching pattern leads to the generation of oligosaccharide fragments bearing two Glc*p*A residues upon treatment with endoxylanase. Less densely substituted regions give rise to oligosaccharides bearing zero or one Glc*p*A side chain. Our data do not allow the overall distribution of these substitution patterns (*i.e.*, randomly distributed within the polymer, clustered in blocks, or separated in structurally distinct polymers) to be determined.

Our phylogenetic analysis suggests that the *P. patens* and *Selaginella moellendorfii* genomes include homologs of several genes that have been implicated in the biosynthesis of glucuronoxylans in angiosperms. The ability of *P. patens* to synthesize a glucuronoxylan that is structurally homologous to those produced by vascular plants supports the hypothesis that several *P. patens* genes are functional orthologs of their homologs in vascular plants. These observations are thus consistent with the notion that the ability to synthesize glucuronoxylan predated the appearance of vascular plants (Carafa et al., 2005). However, the absence of 4-*O*-Me-Glc*p*A in *P. patens* xylan and its ubiquitous presence in vascular plant xylans suggests that the ability to *O*-methylate glucuronic acid co-evolved with the vascular anatomy of tracheophytes.

We and others have suggested that the proteins encoded by *PARVUS, IRX8*, and *IRX7* are candidates for the glycosyltransferases involved in the synthesis of glucuronoxylan reducing end sequence (Peña et al. 2007). (Scheller and Ulvskov 2010) have extended this notion by

proposing that PARVUS transfers xylose to an as yet unidentified acceptor, that IRX8 is a xylose-specific galacturonosyltransferase, and that IRX7 is a rhamnose-specific xylosyltransferase. However, the land plant homologs of IRX7 fall into three distinct clusters, one cluster includes proteins from seedless plants, the second cluster includes proteins from grasses, and the third cluster includes proteins from gymnosperns and eudicots (**Figure 2.5**). Our data indicates that the xylans isolated from *P. patens*, *S. kraussiana*, *E. hyemale*, and the xylans of rice and other grasses (M.J. Peña, M.A. O'Neill, and W.S York, unpublished data) lack the Rha and GalA-containing glycosyl sequence at their reducing ends. Thus, we suggest that the ability to synthesize this oligosaccharide sequence coevolved with the ability to form secondary xylem and woody tissues. These observations also suggest that IRX7 homologs in grasses and seedless plants are not orthologous to Arabidopsis IRX7. Additional studies are required to determine if the reducing-end glycosyl sequence was lost when monocots and dicots diverged, during the evolution of monocots or when the poaceae diverged from the other monocots.

Secondary cell wall formation in vascular plants is accompanied by major changes in the pattern of gene expression (Aspeborg et al. 2005). These observations are consistent with the notion that the deposition of glucuronoxylan is associated with developmentally-related changes in wall composition and structure that facilitate the biological functions of specialized cells involved in mechanical support and water transport. Our immunolabeling data indicates cell-specific deposition of cell wall hemicelluloses in *P. patens*. For example, the xyloglucan-directed antibody, CCRC-M88, labels the walls of leaf and shoot cells, but weakly labels the axillary hair cells (**Figure 2.1e**). In contrast, the xylan-directed antibodies LM11, CCRC-M137 (**Figure 2.1**), CCRC-M147, CCRC-M154 and CCRC-M160 (**Figure S2.1**) preferentially label axillary hair

cell walls. In vascular plants, xyloglucan is present predominantly in primary cell walls but is only a quantitatively minor component of secondary cell walls. Thus, in *P. patens* and vascular plants the general pattern of hemicellulose deposition shows a striking resemblance in that both appear to deposit glucuronoxylan in specialized cells whose walls contain relatively small amounts of xyloglucan.

The presence of xylan in the axillary hair cells of *P. patens* may provide clues to its biological function in this species. It is likely that the pattern of gene expression is considerably different in axillary hair cells than in other *P. patens* cells and tissues (Hiwatashi et al. 2001). In vascular plants, secondary cell wall development is also accompanied by a major shift in the overall pattern of gene expression (Aspeborg et al. 2005). These observations are consistent with the notion that the deposition of glucuronoxylan is associated with major changes in the overall wall structure that are required for the biological function of certain specialized cells in both mosses and vascular plants. In vascular plants, these functions include mechanical support and water transport capabilities provided by secondary cell walls, which enable these plants to effectively colonize the terrestrial environment. It has been suggested that axillary hair cells of mosses also function in water relations by, for example, secreting mucilage to protect newly formed tissues from desiccation (Ligrone 1986). Further study is required to determine whether the presence of glucuronoxylans confers similar mechanical and physical properties to cell walls in the axillary hairs of mosses and the vascular tissues of tracheopytes. Nevertheless, it is tempting to speculate that these diverse tissues have a common evolutionary origin.

In summary, *P. patens* synthesizes a glucuronoxylan that is structurally homologous to glucuronoxylans in the secondary cell walls of vascular plants. However, the *P. patens*

glucuronoxylan differs from the secondary wall glucuronoxylans of gymnospersm and eudicots in that it lacks the distinctive oligosaccharide structure present at the reducing end and none of its Glc*p*A residues are *O*-methylated. Numerous genes encoding putative glycosyltransferase have been identified in *P. patens* and are likely to be orthologous to genes implicated in glucuronoxylan synthesis in vascular plants. The presence of these common glucuronoxylan structures and glycosyltransferase genes in both *P. patens* and vascular plants suggests that they have a common ancestry. It is likely that secondary cell walls, which are required for the formation of vascular tissues, evolved from such a common ancestor, which already possessed much of the cellular machinery required to synthesize the glucuronoxylan. Subsequent modification of glucuronoxylan structure during vascular plant evolution, including *O*-methylation of the Glc*p*A sidechains and the presence of the distinctive reducing end sequence, may be associated with a plants ability to form secondary xylem and woody tissues.

Materials and methods

Plant material

P. patens (Hedw.) B. S. G. (ecotype Gransden 2004) was grown under aseptic conditions with a 16 h light (50-70 μmol photons m^{-2} s^{-1}/8 h dark cycle at 23° C on solid modified Knop's medium (Fu et al. 2007). The gametophores (4-5 weeks old) were then transferred to liquid Knop's medium (200 mL) in 500 mL Erlenmeyer flasks and grown under 19 h light (50-70 μmol photons m^{-2} s^{-1}) at 24° C on a shaker (85-87 rpm). After 4-5 weeks, the moss gametophores were kept on a shaker (85 rpm) for 2 days in the absence of light to allow the tissues to metabolize starch. The gametophores were then washed with deionized water, and stored at -80° C.

Equisetum hyemale and *Selaginella kraussiana* were obtained from the Plant Biology greenhouse, the University of. Georgia. Aerial portions of the sporophyte generation of the plants were collected, rinsed with water and stored at -80° C.

Tissue fixation and immunolabeling

Physcomitrella gametophores (4-5 weeks old) were fixed for 2.5 h at room temperature in 25 mM Na phosphate buffer, pH 7.1, containing paraformaldehyde (1.6%; w/v) and glutaraldehyde (0.2%; w/v). The tissue was rinsed with 25 mM Na phosphate and water (twice for 15 minutes each) and then dehydrated using a graded ethanol series [20, 35, 50, 62, 75, 85, 95, 100, 100, 100% (v/v) EtOH, 30 minutes each step]. The dehydrated tissue was then infiltrated with LR White embedding resin (Ted Pella Inc., http://www.tedpella.com) [33% and 66% (v/v) resin in 100% EtOH, 24 h each, followed by 3 changes of 100% resin, also 24 h each]. The infiltrated tissue was transferred to gelatin capsules containing 100% resin for embedding, and the resin then polymerized by exposing the capsules for 48 h at 4 ºC to UV light (365 nm).

Semi-thin sections (250 nm) were cut using a Leica EM UC6 microtome (Leica Microsystems, http://www.leica-microsystems.com) and mounted on Colorfrost/Plus glass microslides (Fisher Scientific, http://www.fishersci.com). Immunolabelling was carried out at room temperature. Nonspecific antibody-binding sites were blocked by incubating the sections for 75 min with 3% (w/v) non-fat dry skim milk in 10 mM potassium phosphate, pH 7.1, containing 0.5 M NaCl (KPBS, 10 µL). The solution was removed and then KPBS (10 µL) added to the section for 5 min. The KPBS was removed and undiluted hybridoma supernatant (10 µL) added and then incubated for 120-150 min. Sections were washed with KPBS three

times for 5 minutes each, followed by incubation for 90-120 minutes with the secondary antibody. For the CCRC series of antibodies, we used goat anti-mouse conjugated to Alexa-fluor 488 (Invitrogen, http://www.invitrogen.com) diluted 1:100 in KPBS, and for the LM series of antibodies, we used goat anti-rat conjugated to Alexa-fluor 488 (Invitrogen) diluted 1:100 in KPBS. Sections were then washed with KPBS for 5 minutes, then with distilled water for 5 min. Prior to applying a cover slip, CITIFLUOR antifadant mounting medium AF1 (Electron Microscopy Sciences, http://www.emsdiasum.com) was applied.

Light microscopy was carried out using an Eclipse 80i microscope (Nikon, http://www.nikon.com/) equipped with differential interference contrast and epifluorescence optics. Images were captured with Nikon DS-Ri1 camera head (Nikon,) using NIS-Elements Basic Research software. A Nikon B-2E1C filter was used with excitation at 465-495 nm and emission at 515-555 nm. Images were assembled using Adobe Photoshop (Adobe, http://www.adobe.com/).

Electron microscopy

80 nm sections were cut using a Leica EM UC6 microtome (Leica Microsystems) and mounted on nickel grids (100 mesh). Sections were stained with 2% uranyl acetate (10 min) and lead citrate (2 min). Transmission electron microscopy (TEM) was carried out on a JEM 1210 high-resolution TEM (Jeol, http://www.jeol.com/) with digital imaging acquisition and archiving. Images were assembled using Adobe Photoshop.

Preparation of cell walls as their alcohol-inosluble residues (AIR)

Protonemal or leafy gametophore tissues of *Physcomitrella* (36 - 45 g) were frozen in liquid nitrogen and ground to a fine powder in a mortar and pestle. The powder was suspended in 50 mM sodium acetate, pH 5, containing 50 mM NaCl and 30 mM Na ascorbate (1 L). The suspension was filtered through nylon mesh and the insoluble residue then suspended in aq 80% (v/v) EtOH. The suspension was filtered through nylon mesh and the insoluble residue then suspended in absolute EtOH (1 g of tissue/ 6 – 7 mL of EtOH). The suspension was filtered through nylon mesh and the insoluble residue suspended in $CHCl_3$–MeOH (1:1 v/v, 1 g tissue/ml solvent) and kept overnight at room temperature. The residue was collected by filtration and washed with acetone (1 g of tissue/5 – 6 mL of acetone). The resulting AIR, which consists of cell-wall material along with starch, was vacuum dried at room temperature. We typically obtained a yield of 50 mg AIR from 1 g fresh weight of tissue.

Removal of starch from AIR

The AIR generated from *P. patens* tissues was found to contain large amounts of starch that had to be removed to obtain material suitable for isolation of cell wall polysaccharides. The AIR (1.0 g) was suspended in dimethyl sulfoxide (DMSO, 100 mL) and stirred for 24 h at room temperature (Carpita and Kanabus 1987). The suspension was filtered and the insoluble residue washed extensively with 50 mM NaOAc pH 7, (100 mL), containing 5% (v/v) DMSO. The washed residue was then suspended in 50 mM NaOAc pH 7, (100 mL), containing 5% (v/v) DMSO and 2.5 µL of α-amylase (5 units, *Bacillus* Type IIA, Sigma-Aldrich, http://www.sigmaaldrich.com) and kept at 37 °C overnight. The amylase-treated residue was collected by filtration and washed with water. Iodine staining [0.8% (w/v) Potassium iodide,

0.2% (w/v) Iodine] was used to visualize the residue under microscope and it indicated that ~80% of starch had been removed from the AIR.

Sequential extraction of P. patens de-starched AIR

The de-starched residue was suspended in 50 mM ammonium oxalate, pH 5 (0.1 g of AIR/10mL) and stirred overnight at room temperature. The suspension was then filtered through nylon mesh. The ammonium oxalate-treated residue was then suspended in 50 mM NaOAc, pH 5, (100 mL, 10 mg AIR/mL), containing 0.01% (w/v) thimerosal and treated with a xyloglucan-specific endoglucanase (XEG, 1 μL of enzyme/ 10 mL, Novozymes, http:// www.novozymes.com) as described (Pauly et al. 1999). The suspension was kept at room temperature for 24 h and then filtered through nylon mesh. The XEG treatment was repeated and then the soluble and insoluble material collected by filtration through nylon mesh.

The XEG-treated AIR was suspended in 1 M KOH (100 mL) containing 1% (w/v) $NaBH_4$ and kept for 24 h at room temperature. The suspension was filtered and the residue was suspended in 4 M KOH (100 mL) containing 1% (w/v) $NaBH_4$ for a further 24 h. Octanol (5 drops) was added to the 1 M and 4 M soluble extracts to avoid excessive foaming as they were neutralized with glacial acetic acid. After neutralization, the extracts were dialyzed (3500 MW cut – off tubing, Spectrum Laboratories, http://www.spectrumlabs.com) against repeated changes of deionized water and then lyophilized. Insoluble residue after 4M KOH extraction was treated with 100 mM sodium chlorite and 100 μL of glacial acetic acid (Wise et al. 1946, Ahlgren and Goring 1971). The solution was washed extensively with water and the insoluble residue was recovered by centrifugation. The residue was treated again with 4M KOH (post-chlorite 4M

KOH) to extract more material from the cell wall. The residue after post-chlorite 4M KOH was further treated with 5M KOH containing 4% (w/v) boric acid for 24 h at RT. The supernatent was collected and neutralized with glacial acetic acid and lyophilized for further analysis.

Sequential extraction of S. kraussiana *and* E. hyemale *AIR*

The AIR was extracted sequentially with 50 mM ammonium oxalate, 1 M and 4 M KOH as described above.

Total sugar estimation and ELISA

All soluble extracts of ammonium oxalate, 1M KOH, 4M KOH, chlorite, post chlorite 4M KOH and 5M KOH containing 4% (w/v) boric acid were dissolved in deionized water at a concentration of 0.2 mg/ mL. Phenol-sulfuric acid assay(Masuko et al. 2005) was used to estimate the total sugar contents in cell wall extracts. All extracts were diluted to same sugar concentration. ELISA plates (Costar 3598) were loaded with 50 µL of the diluted cell wall extracts (60 µg of sugar/ mL) and allowed to dry overnight at 37° C. ELISAs were performed as described (Pattathil et al. 2010). A series of monoclonal antibodies directed against structurally diverse plant cell wall carbohydrate epitopes were used (Pattathil et al. 2010). ELISA data are presented as a color-coded heat map with brightest yellow indicating the highest binding and black representing no binding (Pattathil et al. 2010).

Monoclonal Antibodies

CCRC, JIM, and MAC series of monoclonal antibodies used in this study were obtained as hybridoma cell culture supernatants from the Complex Carbohydrate Research Center collection

(available through CarboSource Services; http://www.carbosource.net). The LM series of antibodies were obtained from PlantProbes (Leeds, UK; http://www.plantprobes.net).

Endo-xylanase treatment of cell wall extracts and generation of xylan oligosaccharides

The 4 M KOH–soluble materials (~20 mg) was suspended in water and ethanol was added to a final concentration of 60% (v/v). The mixture was kept overnight at 4 °C. The insoluble material was collected by centrifugation (2800 g, 5 min) and lyophilised. The insoluble residue was further suspended in 50 mM ammonium formate, pH 5, (2.5mL), and treated for 24 h at 37° C with *Trichoderma viride* M1 endoxylanase (3.5 units, Megazyme, http://www.megazyme.com). The insoluble material was removed by centrifugation (2800 g, 5 min.) and the supernatant was collected. Ethanol was added to the supernatant to a final concentration of 60% (v/v), the mixture kept for 24h at 4 °C, and the precipitate that formed removed by centrifugation. The supernatant was purged with air to remove ethanol and the solution then lyophilized. Fractions enriched in the xylo-oligosaccharides were obtained by size-exclusion chromatography using a Dionex Ultimate 3000 LC (Dionex, *http://www.dionex.com*) and a Superdex SD75 HR10/30 column (GE Healthcare, http://www.gehealthcare.com) eluted with 50 mM ammonium formate, pH 5, at 0.5 ml/min. The eluant was monitored with a Shodex R101 refractive index detector (Shodex, http://www.shodex.net) and fractions collected manually.

Per-O-*methylation of the xylo-oligosaccharides*

Xylo-oligosaccharide-enriched material (~1 mg) was dissolved in dry DMSO (0.2 mL) and per-*O*-methylated as described (Mazumder and York 2010).

MALDI-TOF mass spectrometry

Positive ion MALDI-TOF mass spectra were recorded using a Bruker LT MALDI-TOF mass spectrometer interfaced to a Bruker biospectrometry workstation (Bruker Daltonics, http://www.bdal.com). Aqueous samples (1 μL of a mg/ml solution) were mixed with an equal volume of a matrix solution (0.1 M 2,5-dihydroxybenzoic acid in aq 50% (v/v) MeCN) and dried on the MALDI target plate. Typically, spectra from 200 laser shots were summed to generate a mass spectrum (Mazumder and York, 2010).

ESI mass spectrometry

Positive ion ESI mass spectra of the per-*O*-methylated oligosaccharides were obtained using a Thermo Scientific LTQ XL mass spectrometer (Thermo Scientific, http://www.thermoscientific.com) as described (Mazumder and York 2010).

^{1}H-NMR spectroscopy

Xylo-oligosaccharide-enriched material (~1 mg) was dissolved in D_2O (0.5-1.0 mL, 99.9%). ^{1}H-NMR spectra were recorded with a Varian Inova NMR spectrometer (Varian http://www.varianinc.com) operating at 600MHz. All two dimensional spectra were recorded using standard Varian pulse programs. Chemical shifts were measured relative to internal acetone at δ 2.225.

Glycosyl-linkage composition analyses

Glycosyl-linkage composition analysis was performed using a Hewlett Packard chromatograph (5890) coupled to a Hewlett Packard 5870 mass spectrometer (Agilent, http://www.home.agilent.com) as described (Mazumder and York 2010).

Phylogenetic analysis

There is one Pfam (Finn et al. 2006) domain model associated with the GT43 family, PF03360.8 (Glyco_transf_43), and one domain model associated with the GT47 family, PF03016.7 (Exostosin). We ran HMMer search (Eddy 1998) by querying these Hidden Markov Models (HMM) in ls mode (Eddy 1998) against the predicted open reading frames (translated peptides) of 16 plant and green algal genomes (see Table S1). An E-value cutoff $\leq 1e^{-5}$ was adopted to select significant protein homologs.

Multiple sequence alignments (MSAs) of the amino acid sequences were performed using MAFFT v6.603 (Katoh et al. 2005a) using L-INS-I (Ahola et al. 2006, Nuin et al. 2006). Maximum likelihood (ML) trees were built using PhyML v2.4.4 (Guindon and Gascuel 2003) with the JTT model, 100 replicates of bootstrap analyses, estimated proportion of invariable sites, four rate categories, estimated gamma distribution parameter, and an optimized starting BIONJ tree. The trees were visualized using MEGA version 4 (Tamura et al. 2007).

Supplementary data

Supplementary material for this article is available online at http://glycob.oxford journals.org.

Funding

This research was funded as part of the Bioenergy Science Center (DE-AC05-00OR22725), a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the U.S. Department of Energy Office of Science (grant DE-FG02-96ER20220), and by a US Department of Agriculture National Institute of Food and Agriculture National Research Initiative Competitive Grant (2007-35318-18389 to AWR). Support for infrastructure and analytical instrumentation was provided by the U.S. Department of Energy-funded Center for Plant and Microbial Complex Carbohydrates (grant DE–FG02–93ER20097). Generation of the CCRC-series of plant glycan-directed monoclonal antibodies used in this study was supported by the NSF Plant Genome Program (grant DBI-0421683).

Acknowledgements

We thank Katrina Saffold and Stefan Eberhard of the Complex Carbohydrate Research Center for recording ESI-MS spectra and for providing the *P. patens* cultures, respectively.

Conflict of interest statement

None declared

Abbreviations

AIR, alcohol insoluble residue; Ara*f*, α-L-arabinofuranosyl; DMSO, dimethyl sulfoxide; ESI-MS, electrospray-ionization mass spectrometry; *FRA*, *FRAGILE FIBER*; *GAUT*, *GALACTOSYLURONICACID TRANSFERASE*; GlcA, α-D-glucosyluronic acid residue; *GATL*, *GALACTOSYLURONICACID TRANSFERASE-LIKE*; 4-*O*-Me-GlcA, 4-O-methyl α-D-glucosyluronic acid; GT, glycosyltransferase; GUX, Glucuronic acid substitution of xylan; *GUT*, *GLUCURONOSYLTRANSFERASE*; *IRX, IRREGULAR XYLEM*; KPBS, 10 mM potassium phosphate, pH 7.1, containing 0.5 M NaCl; MALDI-TOF-MS, matrix-assisted laser-desorption ionization time-of-flight mass spectrometry; NMR, nuclear magnetIc resonance spectroscopy; SEC, size-exclusion chromatography; Xyl, 1,4-linked β-D-xylopyranosyl;

References

Ahlgren PA, Goring DAI. 1971. Removal of wood components during chlorite delignification of black spruce. *Can J Chem*. 49:1272-1275.

Ahola V, Aittokallio T, Vihinen M, Uusipaikka E. 2006. A statistical score for assessing the quality of multiple sequence alignments. *BMC Bioinformatics*. 7:484.

Aspeborg H, Schrader J, Coutinho PM, Stam M, Kallas A, Djerbi S, Nilsson P, Denman S, Amini B, Sterky F*, et al.* 2005. Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen. *Plant Physiol.* 137:983-997.

Atkins E, Parker K, Preston R. 1969. The Helical structure of the β-1,3-linked xylan in some siphoneous green algae. *Proc. Roy. Soc. B.* 173:209-221.

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

Bremner I, Wilkie K. 1966. The hemicelluloses of bracken:: Part I. An acidic xylan. *Carbohydr Res.* 2:24-34.

Brown DM, Zhang Z, Stephens E, Dupree P, Turner SR. 2009. Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *Plant J.* 57:732-746.

Carafa A, Duckett JG, Knox JP, Ligrone R. 2005. Distribution of cell-wall xylans in bryophytes and tracheophytes: new insights into basal interrelationships of land plants. *New Phytol.* 168:231-240.

Carpita NC, Kanabus J. 1987. Extraction of starch by dimethyl sulfoxide and quantitation by enzymatic assay. *Anal Biochem.* 161:132-139.

Cartmell A, McKee LS, Peña MJ, Larsbrink J, Brumer H, Kaneko S, Ichinose H, Lewis RJ, Vikso-Nielsen A, Gilbert HJ*, et al.* 2011. The structure and function of an arabinan-specific alpha-1,2-arabinofuranosidase identified from screening the activities of bacterial GH43 glycoside hydrolases. *J Biol Chem*. 286:15483-15495.

Dalessandro G, Northcote DH. 1977. Changes in enzymic activities of nucleoside diphosphate sugar interconversions during differentiation of cambium to xylem in pine and fir. *Biochem J*. 162:281-288.

Domozych DS, Sorensen I, Willats WG. 2009. The distribution of cell wall polymers during antheridium development and spermatogenesis in the Charophycean green alga, *Chara corallina*. *Ann. Bot.* 104:1045-1056.

Ebringerová A, Hromadkova Z, Heinze T. 2005. Hemicellulose. *Adv Polym Sci.* 186:1-67.

Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics.* 14:755-763.

Evert RF. 2006. Esau's plant anatomy: meristems, cells, and tissues of the plant body: their structure, function, and development, 3rd edition. Hoboken, NJ: John Wiley.

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R*, et al.* 2006. Pfam: clans, web tools and services. *Nucleic Acids Res.* 34:D247-251.

Fu H, Yadav MP, Nothnagel EA. 2007. *Physcomitrella patens* arabinogalactan proteins contain abundant terminal 3-*O*-methyl-L-rhamnosyl residues not found in angiosperms. *Planta.* 226:1511-1524.

Graham L, Cook M, Busse J. 2000. The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proc Natl Acad Sci USA*. 97:4535.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 52:696-704.

Hiwatashi Y, Nishiyama T, Fujita T, Hasebe M. 2001. Establishment of gene-trap and enhancer-trap systems in the moss *Physcomitrella patens*. *Plant J.* 28:105-116.

Izydorczyk M, Biliaderis C. 1995. Cereal arabinoxylans: advances in structure and physicochemical properties. *Carbohydr Polym.* 28:33-48.

Johansson M, Samuelson O. 1977. Reducing end groups in brich xylan and their alkaline degradation. *Wood Sci Technol.* 11:251-263.

Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33:511-518.

Kong Y, Zhou G, Avci U, Gu X, Jones C, Yin Y, Xu Y, Hahn MG. 2009. Two poplar glycosyltransferase genes, PdGATL1.1 and PdGATL1.2, are functional orthologs to PARVUS/AtGATL1 in Arabidopsis. *Mol Plant*. 2:1040-1050.

Kremer C, Pettolino F, Bacic A, Drinnan A. 2004. Distribution of cell wall components in Sphagnum hyaline cells and in liverwort and hornwort elaters. *Planta*. 219:1023-1035.

Lee C, Teng Q, Huang W, Zhong R, Ye Z. 2009. The poplar GT8E and GT8F glycosyltransferases are functional orthologs of Arabidopsis PARVUS involved in glucuronoxylan biosynthesis. *Plant Cell Physiol.* 50:1982.

Lee C, Teng Q, Huang W, Zhong R, Ye ZH. 2010. The Arabidopsis family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. *Plant Physiol.* 153:526-541.

Ligrone R. 1986. Structure, development and cytochemistry of mucilage-secreting hairs in the moss Timmiella-barbuloides (Brid) Moenk. *Ann Bot.* 58:859-868.

Mandal P, Pujol C, Damonte E, Ghosh T, Ray B. 2009. Xylans from Scinaia hatei: Structural features, sulfation and anti-HSV activity. *Int J Biol Macromol*. 46:173-178.

Masuko T, Minami A, Iwasaki N, Majima T, Nishimura S, Lee YC. 2005. Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. *Anal Biochem*. 339:69-72.

Matsunaga T, Ishii T, Matsumoto S, Higuchi M, Darvill A, Albersheim P, O'Neill M. 2004. Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. *Plant Physiol.* 134:339-351.

Matulewicz M, Cerezo A. 1987. Alkali-soluble polysaccharides from *Chaetangium fastigiatum*: Structure of a xylan. *Phytochemistry*. 26:1033-1035.

Mazumder K, York WS. 2010. Structural analysis of arabinoxylans isolated from ball milled switchgrass biomass. *Carbohydr Res* 345:2183-2193.

McCartney L, Marcus SE, Knox JP. 2005. Monoclonal antibodies to plant cell wall xylans and arabinoxylans. *J Histochem Cytochem.* 53:543-546.

Medina R, Lara F, Mazimpaka V, Shevock JR, and Garilleti R. 2011. *Orthotrichum pilosissimum* (Orthotrichaceae), a new moss from arid areas of Nevada with unique axillary hairs. *Bryologist*. 114:316-324.

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

Mishler B, Kelch D. 2009. Phylogenomics and early land plant evolution. In: Goffinet B, Shaw AJ, editors. *Bryophyte Biology*. Cambridge: University Press. p.173–197.

Moller I, Sorensen I, Bernal AJ, Blaukopf C, Lee K, Obro J, Pettolino F, Roberts A, Mikkelsen JD, Knox JP*, et al.* 2007. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *Plant J.* 50:1118-1128.

Niklas K. 1997. *The evolutionary biology of plants*. Chicago: University of Chicago Press.

Nuin PA, Wang Z, Tillier ER. 2006. The accuracy of several multiple sequence alignment programs for proteins. *BMC Bioinformatics*. 7:471.

Pattathil S, Avci U, Baldwin D, Swennes AG, McGill JA, Popper Z, Bootten T, Albert A, Davis RH, Chennareddy C*, et al.* 2010. A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol.* 153:514-525.

Pauly M, Albersheim P, Darvill A, York WS. 1999. Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. *Plant J*. 20:629-639.

Peña MJ, Zhong R, Zhou GK, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye ZH. 2007. Arabidopsis *irregular xylem8* and *irregular xylem9*: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell*. 19:549-563.

Peña MJ, Darvill AG, Eberhard S, York WS, O'Neill MA. 2008. Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology*. 18:891-904.

Percival E, Chanda S. 1950. The xylan of *Rhodymenia palmata*. *Nature*. 166:787-788.

Popper ZA. 2008. Evolution and diversity of green plant cell walls. *Curr Opin Plant Biol.* 11:286-292.

Popper Z. 2011. Evolution and diversity of plant cell walls: from algae to flowering plants. *Ann Rev Plant Biol*. 62:567-590.

Popper ZA, Fry SC. 2003. Primary cell wall composition of bryophytes and charophytes. *Ann Bot.* 91:1-12.

Popper ZA, Tuohy MG. 2010. Beyond the green: understanding the evolutionary puzzle of plant and algal cell walls. *Plant Physiol*. 153:373-383.

Qiu Y, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J. 2006. The deepest divergences in land plants inferred from phylogenomic evidence. *Proc Natl Acad Sci USA.* 103:15511-15516.

Scheller HV, Ulvskov P. 2010. Hemicelluloses. *Ann Rev Plant Biol.* 61:263-289.

Shimizu K, Ishihara M, Ishihara T. 1976. Hemicellulases of brown rotting fungus *Tyromyces palustris*. II. The oligosaccharides from the hydrolysate of a hardwood xylan by the intracellular xylanase. *Mokuzai Gakkaishi.* 22:618-625.

Smith B, Harris P. 1999. The polysaccharide composition of Poales cell walls:: Poaceae cell walls are not unique. *Biochem Syst Ecol.* 27:33-53.

Sorensen I, Domozych D, Willats WG. 2010. How have plant cell walls evolved? *Plant Physiol.* 153:366-372.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol.* 24:1596-1599.

Taylor T, Taylor E, Krings M. 2009. Paleobotany: The biology and evolution of fossil plants, 2nd edition, New York, NY: Academic Press.

Turvey JR, Williams EL. 1970. Structures of some xylans from red algae. *Phytochemistry*. 9:2383-2388.

Verbruggen MA, Spronk BA, Schols HA, Beldman G, Voragen AG, Thomas JR, Kamerling JP, Vliegenthart JF. 1998. Structures of enzymically derived oligosaccharides from sorghum glucuronoarabinoxylan. *Carbohydr Res.* 306:265-274.

Wise LE, Murphy M, Daddieco AA. 1946. Chlorite Holocellulose, Its Fractionation and Bearing on Summative Wood Analysis and on Studies on the Hemicelluloses. *Tech Assoc Pap*. 29:210-218.

Wu AM, Hornblad E, Voxeur A, Gerber L, Rihouey C, Lerouge P, 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 Physiol.* 153:542-554.

Wu AM, Rihouey C, Seveno M, Hornblad E, Singh SK, Matsunaga T, Ishii T, Lerouge P, Marchant A. 2009. The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J.*, 57:718-731.

Yamagaki T, Maeda M, Kanazawa K, Ishizuka Y, Nakanishi H. 1997. Structural clarification of Caulerpa cell wall β-1,3-xylan by NMR spectroscopy. *BioSci BioTechnol Biochem*. 61:1077-1080.

Yin Y, Chen H, Hahn MG, Mohnen D, Xu Y. 2010. Evolution and function of the plant cell wall synthesis-related Glycosyltransferase Family 8. *Plant Physiol.* 153:1729-1746.

Yoon H, Hackett J, Ciniglia C, Pinto G, Bhattacharya D. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol Biol Evol.* 21:809-818.

Young RE, McFarlane HE, Hahn MG, Western TL, Haughn GW, and Samuels AL. 2008. Analysis of the Golgi apparatus in *Arabidopsis* seed coat cells during polarized secretion of pectin-rich mucilage. *Plant Cell.* 20:1623-1638.

Zhong R, Peña MJ, Zhou GK, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WH, 3rd, Darvill AG, York WS, Ye ZH. 2005. Arabidopsis fragile fiber8, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. *Plant Cell*. 17:3390-3408.

Zhong R, Ye ZH. 2003. Unraveling the functions of glycosyltransferase family 47 in plants. *Trends Plant Sci.* 8:565-568.

Zhong R, Ye ZH. 2007. Regulation of cell wall biosynthesis. *Curr Opin Plant Biol*. 10:564-572.

Zhou GK, Zhong R, Richardson EA, Morrison WH, 3rd, Nairn CJ, Wood-Jones A, Ye ZH. 2006. The poplar glycosyltransferase GT47C is functionally conserved with Arabidopsis Fragile fiber8. *Plant Cell Physiol.* 47:1229-1240.

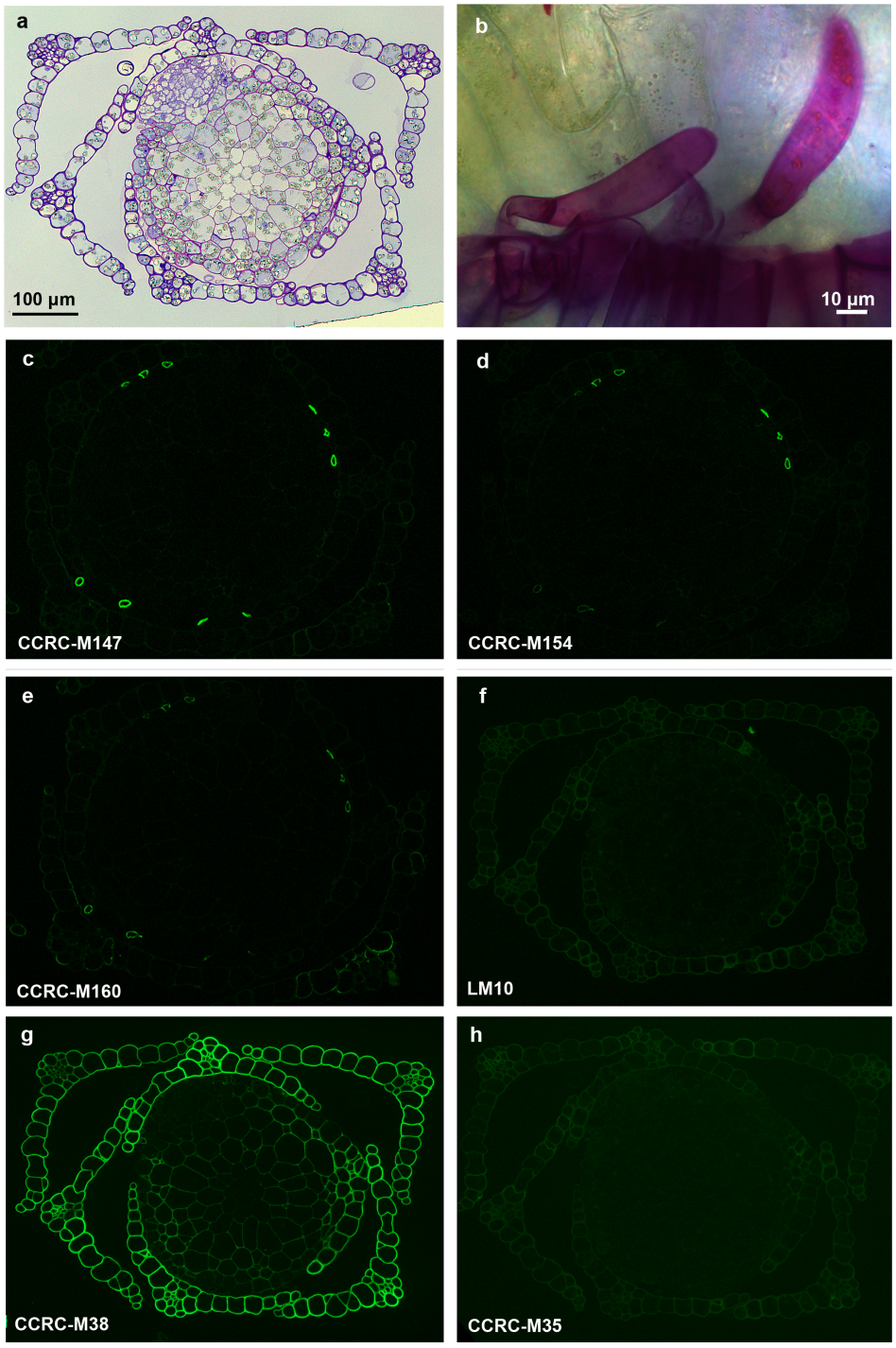
Supplemental information

Supplemental Table S2.1: The 16 plant species used in the phylogenetic study

Index	Abbr.	Clade	Species	Genome	Download
1	mpc	green algae	*Micromonas pusilla CCMP1545*	(Worden et al. 2009)	JGI v2.0
2	mpr	green algae	*Micromonas strain RCC299*	(Worden et al. 2009)	JGI v2.0
3	ol	green algae	*Ostreococcus lucimarinus*	(Palenik et al. 2007)	JGI v1.0
4	ot	green algae	*Ostreococcus tauri*	(Derelle et al. 2006)	JGI v1.0
5	cr	green algae	*Chlamydomonas reinhardtii*	(Merchant et al. 2007)	JGIv3.0
6	vc	green algae	*Volvox carteri f. nagariensis*	No	JGIv1.0
7	pp	moss	*P. patens ssp patens*	(Rensing et al. 2008)	JGI V1.1
8	sm	spike moss	*Selaginella moellendorffii*	No	JGI V1.0
9	pt	dicot	*Populus trichocarpa*	(Tuskan et al. 2006)	JGI V1.1
10	at	dicot	*Arabidopsis thaliana*	(ArabidopsisGenomeInitiative 2000)	ftp://ftp.arabidopsis.org/home/tair/Genes/ TAIR8_genome_release
11	vv	dicot	*Vitis vinifera*	(Jaillon et al. 2007)	http://www.genoscope.cns.fr/
12	mt	dicot	*Medicago truncatula*	No	http://www.medicago.org/genome/downloads/Mt2/
13	gm	dicot	*Glycine max*	No	JGI V1.0
14	os	monocot	*Oryza sativa*	(Goff et al. 2002, Yu et al. 2002)	ftp://ftp.tigr.org/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/version_6.0
15	sb	monocot	*Sorghum bicolor*	(Paterson et al. 2009)	JGI V1.0
16	zm	monocot	*Zea mays*	No	http://ftp.maizesequence.org/release-3a.50

Supplemental Table S2.2: Homologs of xylan biosynthesis genes in *P. patens* and *S. moellendorfii*

Arabidopsis gene/ID	**Identified in Species**	**Gene Identifier**	**Protein IDs**
IRX8 At5g54690	*P. patens*	estExt_gwp_gw1.C_160220	176352
	P. patens	fgenesh1_pg.scaffold_18000059	68313
	P. patens	estExt_gwp_gw1.C_100063	175027
	S. moellendorfii	estExt_fgenesh1_pm.C_590046	268781
PARVUS At1g19300	*P. patens*	estExt_Genewise1.C_2810001	224206
	P. patens	e_gw1.43.57.1	123164
	P. patens	fgenesh1_pg.scaffold_229000012	92683
	P. patens	estExt_Genewise1.C_1290049	216448
	P. patens	fgenesh1_pg.scaffold_4000222	65107
	S. moellendorfii	e_gw1.78.38.1	125336
IRX 9 At2g37090 & IRX9-Like At1g27600	*P. patens*	fgenesh1_pm.scaffold_52000018	57016
	P. patens	estExt_ gwp_gw1.C_10403	173857
	S. moellendorfii	e_gw1.2.354.1	78265
IRX14 At4g36890 & IRX14-Like At5g67230	*P. patens*	gw1.19.152.1	32848
	P. patens	e_gw1.78.95.1	129721
	P. patens	gw1.248.48.1	31864
	S. moellendorfii	gw1.45.105.1	12085
	S. moellendorfii	gw1.32.71.1	12082
IRX7/ Fra8, At2g28110	*P. patens*	e_gw1.315.12.1	150374
	P. patens	e_gw1.13.109.1	115417
	P. patens	estExt_gwp_gw1.C_2170048	194463
	S. moellendorfii	gw1.34.15.1	2308
IRX10 GUT2, At1g27440	*S. moellendorfii*	estExt_fgenesh2_pg.C_210321	442111
IRX10-Like GUT1, At5g61840	*P. patens*	estExt_Genewise1.C_70337	202121



Supplemental Figure S2.1: Bright-field and immunofluorescence light microscope images of cross sections of P.patens leafy gametophores. Epitopes in serial cross-sections were visualized

by their interactions with specific monoclonal antibodies. The scale indicated by the bar in panels A, C – H is 100 μm. The scale indicated by the bar in panel B is 10 μm.

(A) Bright-field image of toluidine blue stained cross section.

(B) Bright-field microscope image showing toluidine blue stained axillary hairs.

(C) CCRC-M147 recognizes xylan and specifically labeled axillary hairs.

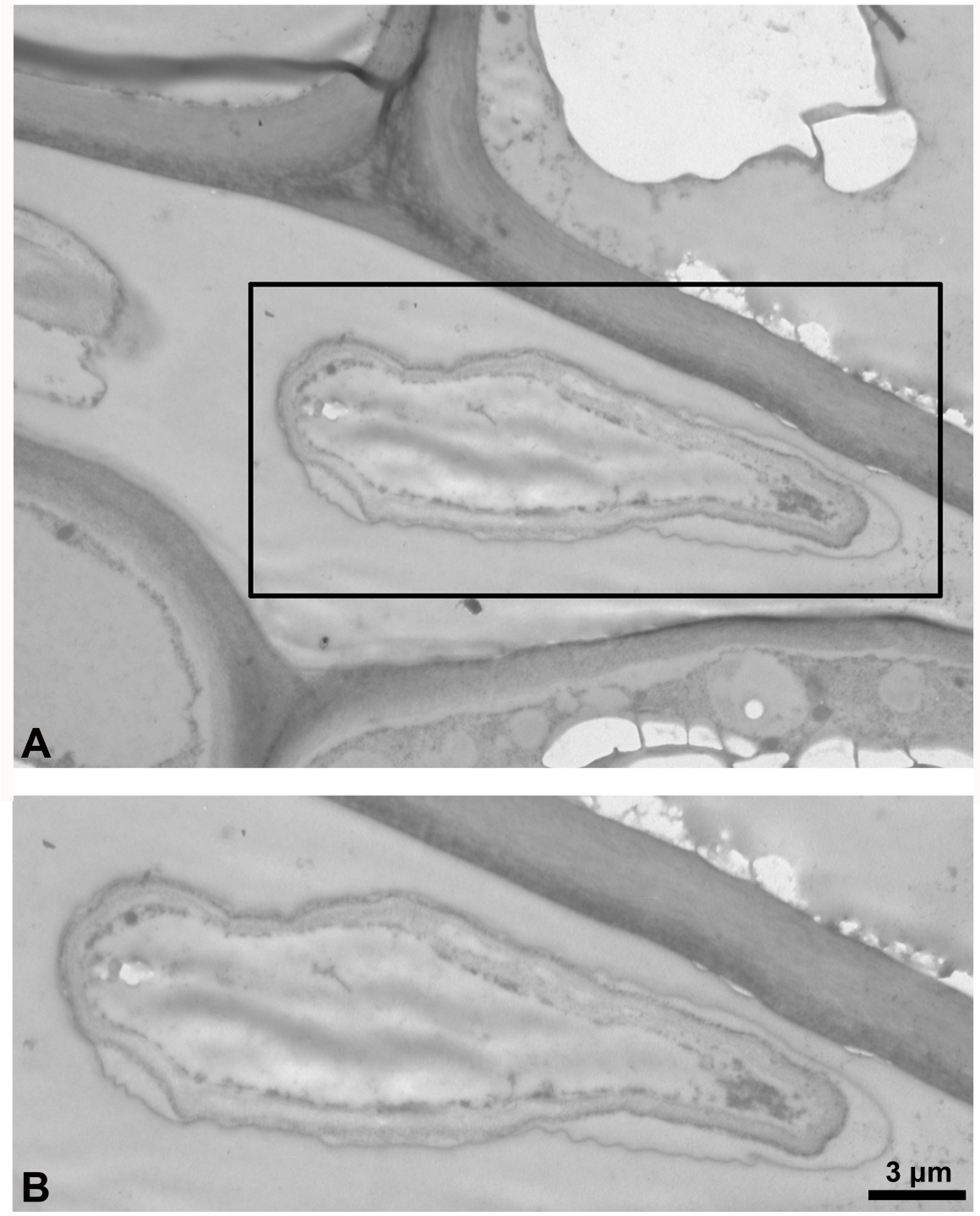
(D) CCRC-M154 recognizes xylan and labeled axillary hairs.

(E) CCRC-M160 recognizes xylan and labeled axillary hairs more strongly than leaf cells.

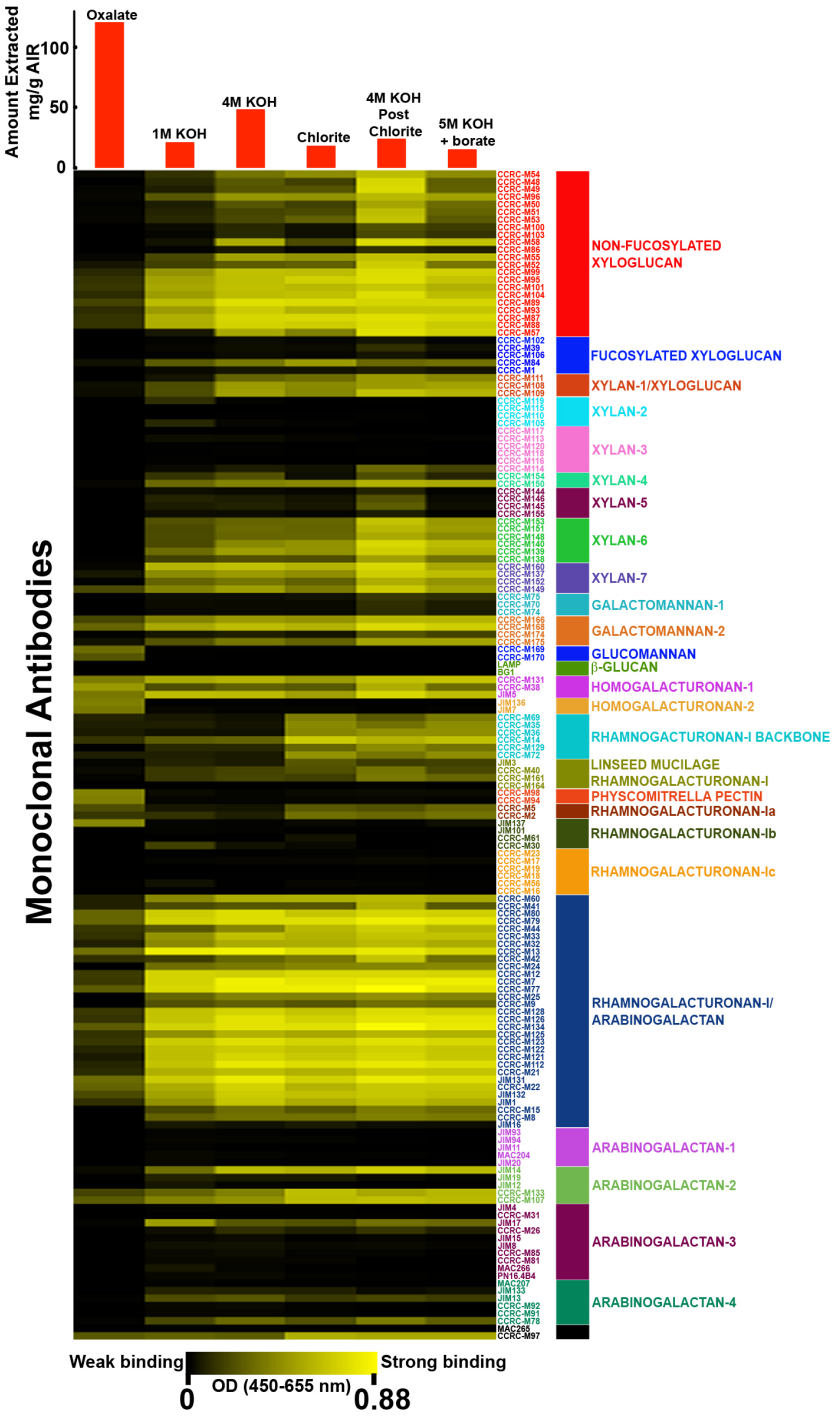
(F) LM10 specific for unsubstituted xylan, weakly labeled leaf cell walls.

(G) CCRC-M38, recognizing de-esterified pectin, labeled leaf cell walls more strongly than leafy shoot cell walls. No labeling of axillary hairs was detected.

(H) CCRC-M35, recognizing epitopes on the unbranched rhamnogalacturonan I backbone, weakly labeled the leaf cell walls and no labeling of axillary hairs was detected.

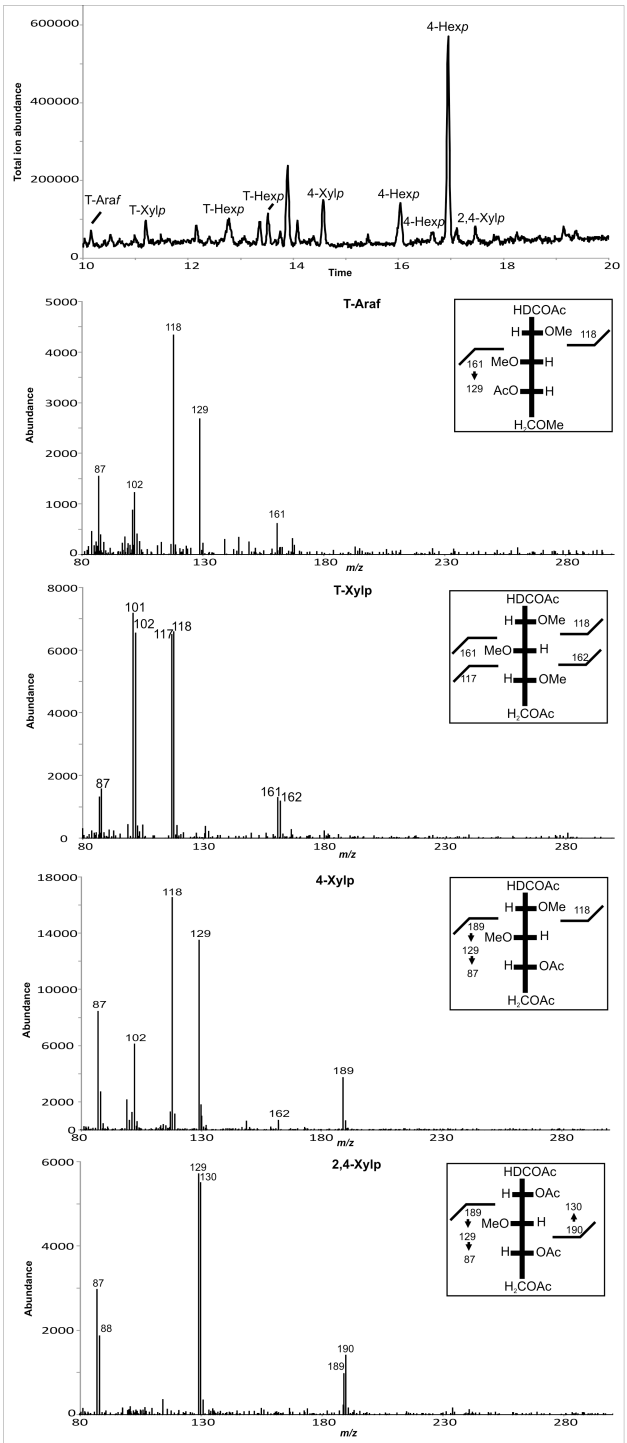


Supplemental Figure S2.2: Transmission electron microscopy image of the axillary hair cell. The scale indicated by the bar is 2 µm.

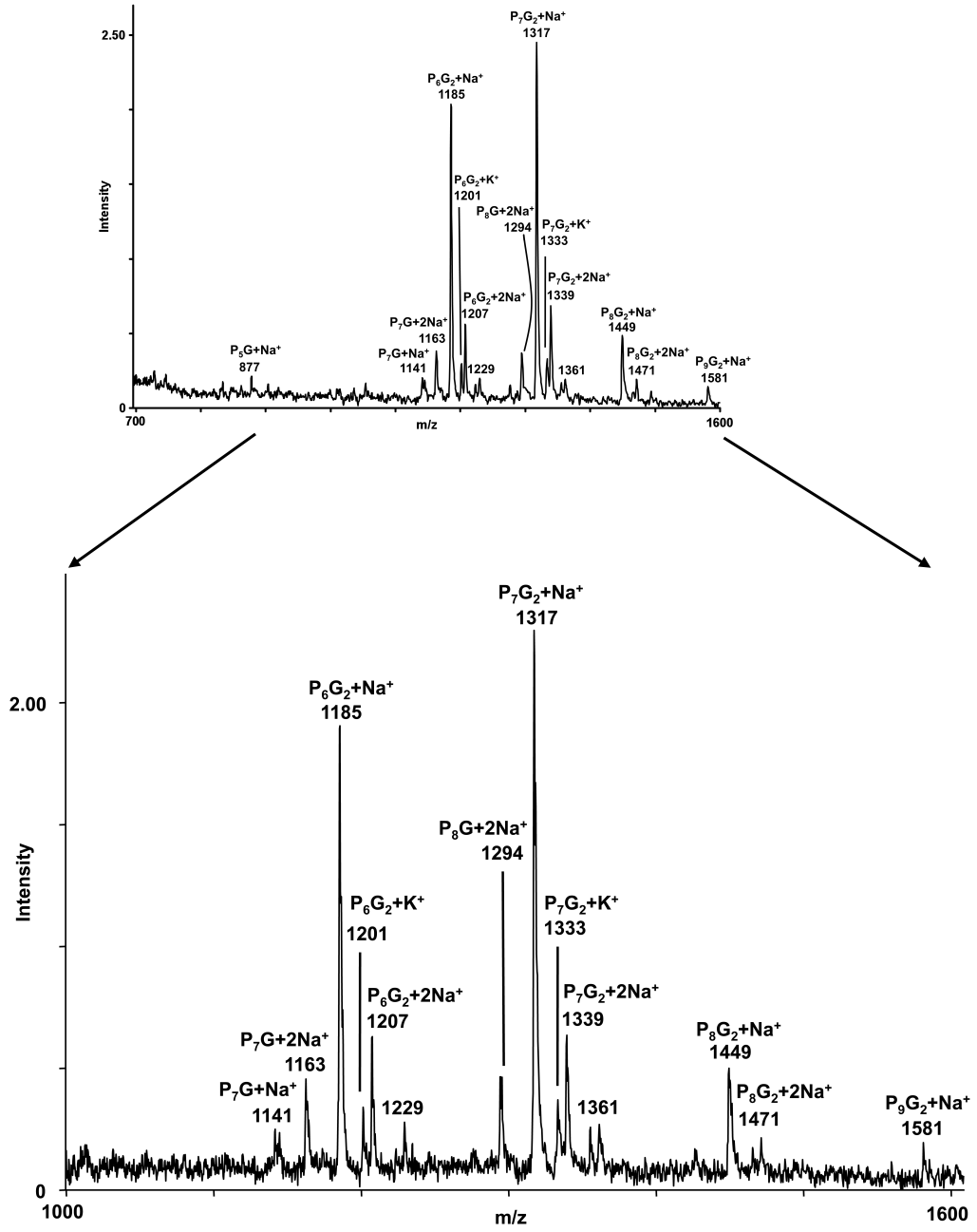


Supplemental Figure S2.3: Glycome profiling of P. patens cell wall extracts. The extracted materials released from de-starched P. patens AIR by various reagents (as labeled at the top of the figure) were loaded onto ELISA plates and were screened against an array of plant glycan-

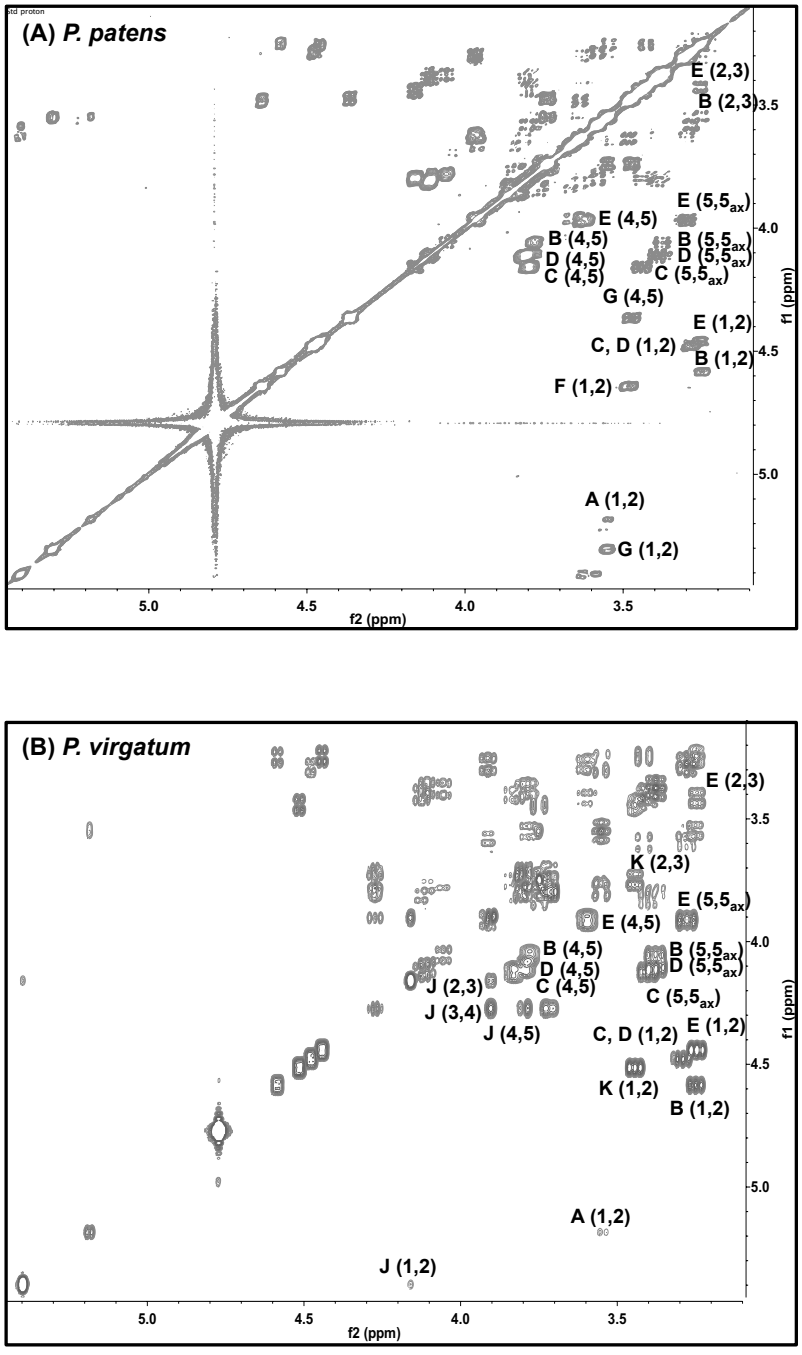
directed monoclonal antibodies (Pattathil et al. 2010) asdescribed in Methods. The panel on the right of the figure lists the individual antibodies used, grouped according to the polysaccharides predominantly recognized by these antibodies. Antibody binding is represented as a colored heat map, with black signifying no binding, and bright yellow representing the strongest binding. The bar graphs at the top indicate the amount of solubilized material recovered at each extraction step per gram of AIR.



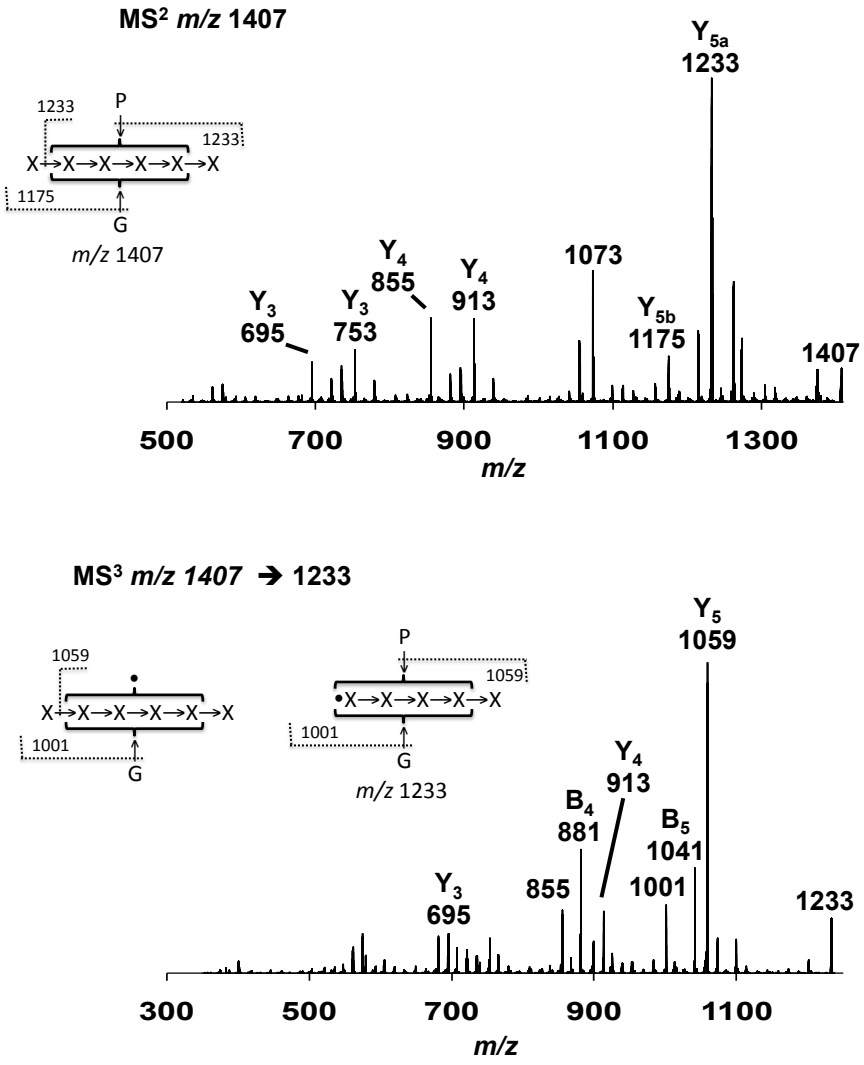
Supplemental Figure S2.4: GC–MS of the partially methylated alditol acetate derivatives prepared by sequentially methylating, hydrolyzing, reducing and acetylating the 4MKOH-soluble fraction from P. patens.



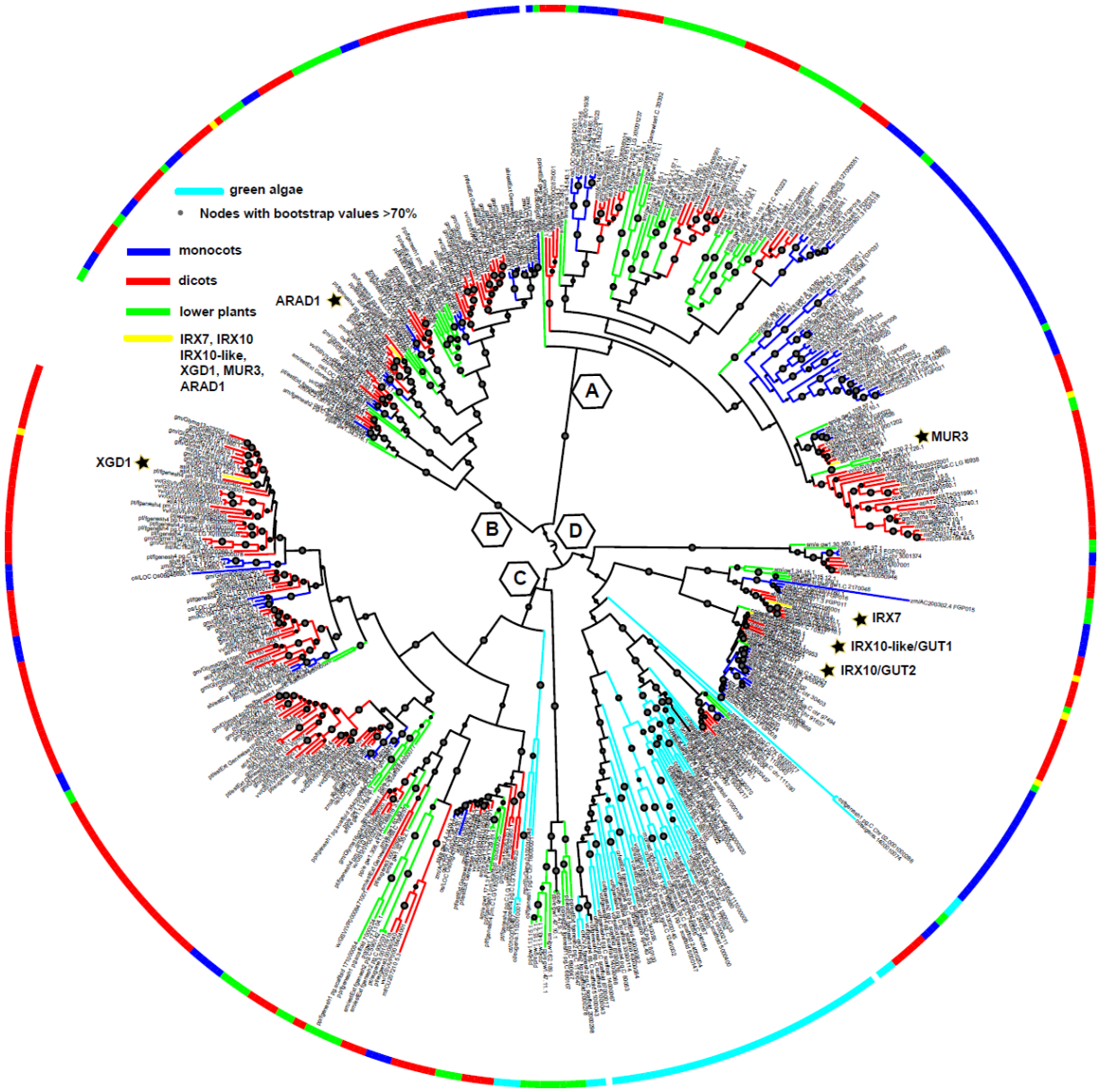
Supplemental Figure S2.5: MALDI-TOF spectrum of the purified xylo-oligosaccharides generated by xylanase treatment of the 4M KOH extract of *P. patens* AIR. Ions are labeled with the most likely ([M + Na]+, [M - H + 2Na]+, [M + K]+) assignments, but ions with more than two sodium atoms are also possible.



Supplemental Figure S2.6: Partial gCOSY spectra of xylo-oligosaccharides isolated from *P. patens and P. virgatum* (switchgrass), illustrating lack of signals corresponding to α-LAraf sidechains (residues J and K) in the P. patens xylo-oligosaccharides.



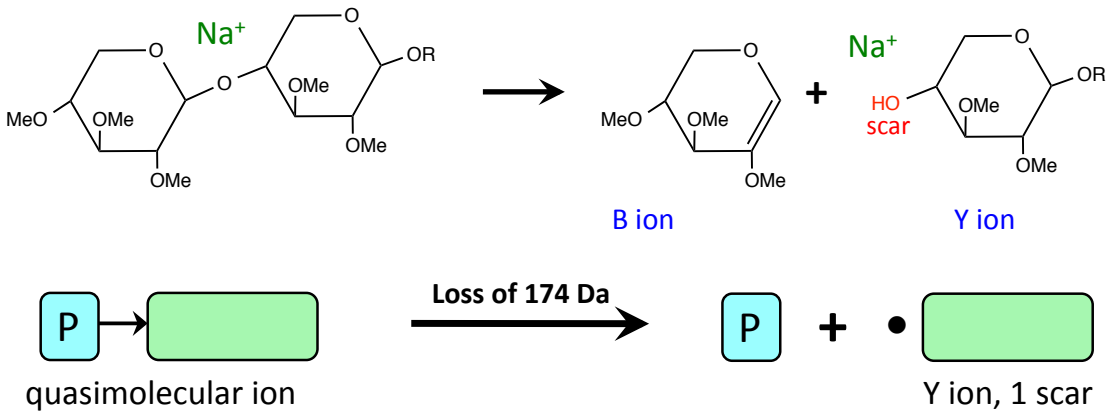
Supplemental Figure S2.7: ESI-MSn of per-O-methylated oligosaccharides (precursor ion at m/z 1407) generated by xylanase treatment of the 4M KOH extract of P. patens AIR. The fragmentation pathway (m/z 1407 – 1233 – 1059), which occurs by two sequential losses of 174 Da, is illustrated. This pathway is consistent with the presence of oligosaccharides having a pentosyl sidechain.



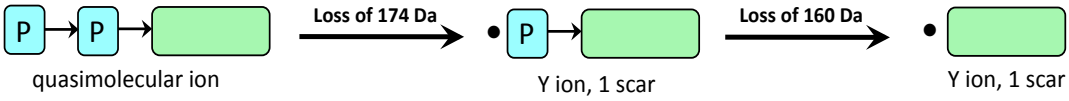
Supplemental Figure S2.8: The maximum likelihood phylogeny of GT47 family proteins from 16 plant genomes. Multiple sequence alignment (MSA) of the conserved Pfam GT47 domain was performed using MAFFT v6.603 (Katoh et al. 2005) using L-INS-I. The phylogeny was reconstructed using the PhyML v2.4.4.

ESI-MS^n of xylo-oligosaccharides

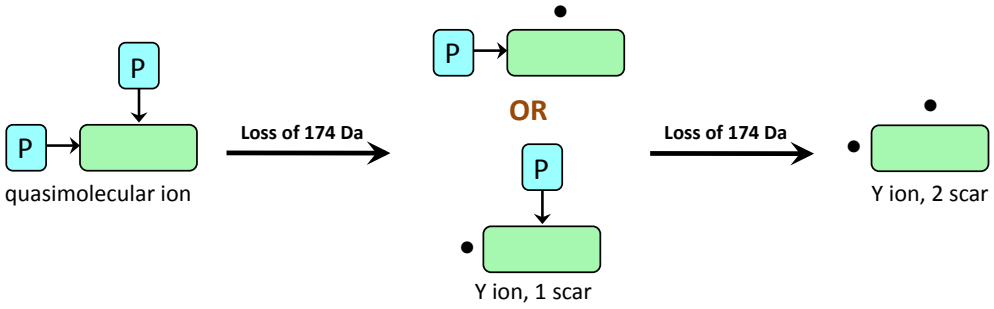
We recently demonstrated (Mazumder and York 2010) that the ESI-MS^n spectra of per-*O*-methylated xylo-oligosaccharides, including those that contain 4-*O*-Me-GlcpA sidechains, are dominated by B and Y fragment ions, which provide robust sequence information (Domon and Costello 1988, Mazumder and York 2010) Y ions contain the reducing terminus of the oligosaccharide, whereas B ions contain the non-reducing terminus (Domon and Costello 1988). Methylation of the oligosaccharides prior to MS^n analysis increases the sensitivity of the method considerably (Domon and Costello 1988) and introduces mass markers that can be used to identify glycosidic bonds that are cleaved during MS^n analysis, which reveals diagnostic, sequence-specific *fragmentation pathways* for the methylated oligosaccharides. For example, a Y ion generated in MS^2 by loss of a terminal pentosyl residue (i.e., loss of 174 Da from the quasimolecular ion) has a so-called "scar" (unmethylated hydroxyl group) exposed by cleavage of the glycosidic bond attaching this pentosyl residue to oligosaccharide. This scar can be recognized by its characteristic 14 Da mass difference, relative to a methylated site. This is illustrated below, chemically and diagramatically. In the diagram, a scar is represented by a black dot.



There are two possible ways for such a Y ion to lose a second pentosyl residue (after its selection for fragmentation in MS^3) to form another Y ion: If the resulting Y ion contains a *single* scar (i.e., loss of 160 Da), then the pentosyl residue lost during MS^3 contains the scar generated in MS^2. That is, the initial scar is lost at the same time a new scar is formed, resulting in an ion with only one scar, indicating that the two pentosyl residues lost thus far are linked to each other as a terminal disaccharide in the original structure, as illustrated below.



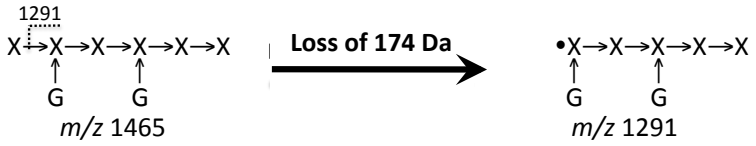
In contrast, if the resulting Y ion contains two scars (i.e., it arises via another loss of 174 Da), the pentosyl residue that is lost during MS^3 does *not* contain the scar generated in MS^2. This can occur only if the original structure has sidechain that is terminated by a pentosyl residue.



In fact, MS^n of the low abundance quasimolecular ion at *m/z* 1407 (Figure S6) in the spectrum of the *P. patens* xylo-oligosaccharides showed exactly this pattern, indicating the presence of a pentosyl sidechain on at least one of the oligosaccharides that give rise to this ion.

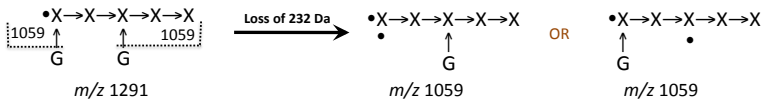
In general, if the number of scars in a Y ion is equal to the number of scars in its immediate precursor, the cleavage leading to the Y ion is accompanied by the loss of a scar that is present in the precursor ion. Thus, key sequence information can often be unambiguously deduced if the Y ion contains only one scar or it is formed by loss of a single residue from its immediate precursor ion. In combination with other structural information (*e.g.*, the *P. patens* heteroxylan backbone is composed of 1,4-linked β-D-Xyl*p* residues) obtained by orthogonal methods such as NMR, application of this rule to fragmentation pathways observed during MS^n analysis allows the location of specific sidechain structures in the xylo-oligosaccharides to be determined.

The quasimolecular ion at *m/z* 1465 (consisting of six pentosyl and two hexuronosyl residues) in the MS^2 psectrum of the per-*O*-methylated xylo-oligosaccharides was selected and subjected to ESI-MS^2 (Figure 3). Diagnostic B and Y ions in the resulting spectrum indicated the presence of at least two oligosaccharides that differed in the location of the Glc*p*A residues. The fragmentation pathway (illustrated in Figure 3 and described next) establishes the glycosyl sequence of one of the major components. Loss of a terminal pentosyl residue from the quasimolecular (*m/z* 1465) ion during MS^2 generated a Y ion at *m/z* 1291 (P_5G_2 with one scar).

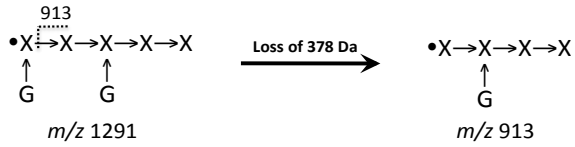


This ion was selected for fragmentation in MS^3, where loss of a hexuronosyl residue was accompanied by generation of an additional scar to form a Y ion at *m/z* 1059 (P_5G with two

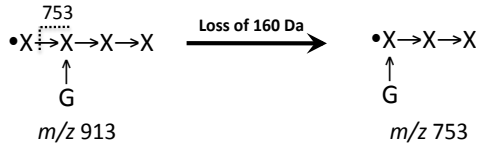
scars), indicating the presence of a hexuronosyl sidechain, which is consistent with NMR analysis.



The *m/z* 1291 precursor ion also lost the combined masses of a hexuronosyl residue and a pentosyl residue to form a Y ion at *m/z* 913 (P_4G with one scar), establishing the presence of a disaccharide consisting of these two residues at the non-reducing end of at least one of the *m/z* 1291 ions fragmented in MS^3.

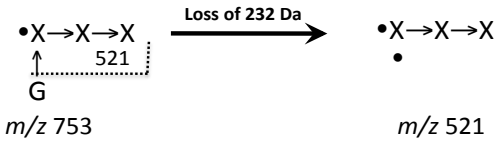


That is, oligosaccharides giving rise to the pathway (*m/z* 1465 – 1291 – 913) have a hexuronosyl sidechain attached to the penultimate backbone residue of the initial structure. The ion at *m/z* 913 was selected for fragmentation in MS^4, generating a Y ion at *m/z* 753 (P_3G with one scar) by loss of a single pentosyl residue.

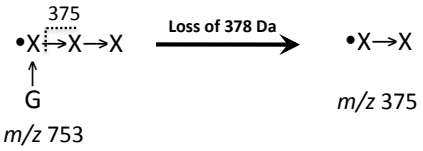


Thus, oligosaccharides giving rise to the (*m/z* 1465 – 1291 – 913 - 753) pathway do not have a sidechain on the next residue in the backbone (i.e., the residue to which the penultimate

backbone residue is linked). This MS^4 spectrum (Figure 3) also contained Y ions at *m/z* 375 (P_2 with one scar) and *m/z* 681 (P_4 with two scars) along with B ions at *m/z* 721 (P_4G with two scars) and *m/z* 561 (P_2G with two scars). (Although the "reducing end" of the B ion does not have a hydroxyl group, it is counted as a type of scar in this discussion.) This combination of fragments in the MS^4 spectrum is consistent with the presence of a uronosyl sidechain on the third backbone residue of the original oligosaccharide. Selection of the Y ion at *m/z* 753 and fragmentation in MS^5 generated a Y ion at *m/z* 521 (P_3 with two scars) by loss of a hexuronosyl residue, indicating presence of a hexuronosyl sidechain on the precursor ion.



This MS^5 spectrum also contained a Y ion at *m/z* 375 (P_3 with one scar), which can be formed only by loss of a disaccharide containing one hexuronosyl and one pentosyl residue from the non-reducing end of the *m/z* 753 precursor ion.



This confirms that a hexuronosyl sidechain is linked to the third backbone residue of the oligosaccharide that gives rise to the (*m/z* 1465 – 1291 – 913 – 753 – 375) pathway, which completely defines the glycosyl sequence of this oligosaccharide. This conclusion is supported by the presence of a B ion at *m/z* 401 (PG with two scars) in the MS^5 spectrum. When

interpreted in light of the results of NMR analysis, which establish a backbone of 1,4-linked xylosyl residues with glucuronic acid sidechains, and the observation that oligosaccharides with branched reducing end residues are not generated by the endoxylanase (Correia et al. 2011), the (*m/z* 1465 – 1291 – 913 – 753 – 375) pathway provides very strong evidence for presence of the doubly-branched structure illustrated in Figure 3.

References:

ArabidopsisGenomeInitiative. 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature*, 408:796-815.

Correia MAS, Mazumder K, Bras JLA, Firbank SJ, Zhu Y, Lewis RJ, York WS, Fontes CMGA, Gilbert HJ. 2011. The structure and function of an arabinoxylan-specific xylanase. *J Biol Chem.* 286:22510-22520.

Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S, Partensky F, Degroeve S, Echeynie S, Cooke R, et al. 2006. Genome analysis of the smallest free-living eukaryote Ostreococcus tauri unveils many unique features. *Proc Natl Acad Sci U S A*. 103:11647-11652.

Domon B, Costello CE. 1988. A Systematic Nomeclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconj. J.*, 5:397-409.

Goff SA, Ricke D, Lan T-H, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, et al. 2002. A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. japonica). *Science*, 296:92-100.

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C, et al. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, 449:463-467.

Mazumder K, York WS. 2010. Structural analysis of arabinoxylans isolated from ball milled switchgrass biomass. *Carbohydr Res.* 345:2183-2193.

Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Marechal-Drouard L, et al. 2007. The Chlamydomonas genome reveals the evolution of key animal and plant functions. *Science*, 318:245-250.

Palenik B, Grimwood J, Aerts A, Rouze P, Salamov A, Putnam N, Dupont C, Jorgensen R, Derelle E, Rombauts S, et al. 2007. The tiny eukaryote Ostreococcus provides genomic insights into the paradox of plankton speciation. *Proc Natl Acad Sci U S A*. 104:7705-7710.

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, et al. 2009. The Sorghum bicolor genome and the diversification of grasses. *Nature*, 457:551-556.

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud P-F, Lindquist EA, Kamisugi Y, et al. 2008. The Physcomitrella Genome Reveals Evolutionary Insights into the Conquest of Land by Plants. *Science*, 319:64-69.

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, et al. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, 313:1596-1604.

Worden AZ, Lee JH, Mock T, Rouze P, Simmons MP, Aerts AL, Allen AE, Cuvelier ML, Derelle E, Everett MV, et al. 2009. Green Evolution and Dynamic Adaptations Revealed by Genomes of the Marine Picoeukaryotes Micromonas. *Science (New York, N.Y*, 324:268-272.

Yu J, Hu S, Wang J, Wong GK-S, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, et al. 2002. A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. indica). *Science*, 296:79-92.

CHAPTER 3

COMPARISON OF ARABINOXYLAN STRUCTURE IN BIOENERGY AND MODEL GRASSES[1]

[1]Ameya R. Kulkarni, Sivakumar Pattathil, Michael G. Hahn, William S. York, and Malcolm A. O'Neill. (2012). Comparison of Arabinoxylan Structure in Bioenergy and Model Grasses. Industrial Biotechnology. 8(4): 222-229. doi:10.1089/ind.2012.0014. Reprinted here with permission of the publisher.

Abstract

Heteroxylans were solubilized from the alcohol insoluble residue of switchgrass, rice, brachypodium, miscanthus, foxtail millet and poplar with 1 M KOH. A combination of enzymatic, chemical, NMR and mass spectroscopic, and immunological techniques indicated that grass arabinoxylans have comparable structures. The grass arabinoxylans contain no discernible amounts of the reducing end sequence present in dicot glucuronoxylan. Our data suggest that rice, brachypodium and foxtail millet are suitable model plants for developing technologies to modify arabinoxylan structure and biosynthesis that can be extended to monocot energy crops.

Introduction

Perennial grasses including switchgrass (*Panicum virgatum*) and Miscanthus (*Miscanthus x giganteus)* have considerable potential as energy crops for the production of liquid transportation fuels (Lewandowski et al. 2003). However, improving the field performance and biomass quality of these plants is challenging (Heaton et al. 2008). Switchgrass is self-incompatible which complicates forward or reverse genetics and other genetics related issues (Doust et al. 2009). Miscanthus is a sterile allotriploid and its genome has not been sequenced. Nevertheless, there is considerable ongoing effort to develop genetic and molecular tools to improve the agronomic qualities of these plants. Such efforts together with an increased understanding of grass cell wall structure and biosynthesis will facilitate the development of energy crops with biomass that is less recalcitrant to conversion to fermentable sugars.

The model plant *Arabidopsis thaliana* has been used extensively to identify and functionally characterize genes involved in plant cell wall biosynthesis (Liepman et al. 2010). The availability of a large number of mutant lines affected in cell wall-related genes has also provided an opportunity to determine the roles of cell wall polysaccharides in plant growth and development (Liepman et al. 2010). Nevertheless, Arabidopsis is a dicot and last shared a common ancestor with the grasses (Poaceae) ~200 million years ago. Moreover, the cell walls of dicots and grasses differ (Vogel 2008). Thus, Arabidopsis is not a good model plant for studying genetic traits or cell wall structure and synthesis in grasses.

A good model grass should be easy to grow, have a short life cycle, and reproduce by self-fertilization. Its genome should be sequenced and well annotated. The plant should be easily,

rapidly and stably transformed, and mutant lines should be available and accessible. Moreover, the composition and structure of the biomass from the model plant and the energy crop should be comparable. Three plants, *Oryza sativa* (rice), *Brachypodium distachyon* (brachypodium), and *Setaria italica* (foxtail millet), have been promoted as model plants for understanding growth, photosynthetic efficiency, and cell wall biosynthesis in grasses. Rice and brachypodium are diploid C3 plants that have had their genomes sequenced and in-bred and mutant lines are available. Brachypodium with its small size and fast generation time is a good model for functional genomic studies and for the dissection of more highly conserved traits in bioenergy grasses including genes involved in cell wall (Mur et al. 2011). Foxtail millet is a diploid C4 plant with a sequenced genome and is closely related to switchgrass. Thus, foxtail millet, is a good model plant for the genetic dissection of traits in bioenergy grasses that carry out C4 photosynthesis, as well as for studies of plant architecture and cell wall synthesis (Doust et al. 2009).

Arabinoxylan accounts for ~20% of grass biomass. Some of the arabinose side chains contain ester-linked ferulic and coumaric acids that participate in the cross-linking of arabinoxylans and in the formation of ferulate-polysaccharide-lignin complexes that cross-link the cell wall (de O. Buanafina 2009). Such ferulate-polysaccharide-lignin complexes are hypothesized to function in numerous processes in plants (de O. Buanafina 2009). These complexes together with the interactions between arabinoxylans and cellulose are believed to be the important factors in the recalcitrance of biomass to conversion to sugars (Himmel et al. 2007). In addition, degradation of xylan and other hemicelluloses during biomass pretreatment may result in the formation of compounds that are toxic to the microorganisms used to

deconstruct the biomass and covert the released sugars to liquid fuels (Mosier et al. 2005). Thus, there is considerable interest in understanding arabinoxylan structure and how it can be altered to improve the bioprocessing characteristics of grass biomass. This paper focuses on the comparison of structures of the carbohydrate moieties of grass arabinoxylans and does not address the structures of the phenolic esters linked to this polysaccharide.

Arabinoxylans have a backbone composed of 1→4-linked β-D-xylopyranosyl (Xyl*p*) residues, which are often substituted at *O*-3 with Ara*f* residues and to a much lesser extent with glucuronic acid (GlcA) or 4-O-methyl GlcA residues (MeGlcA) (Ebringerová et al. 2005). Here we have compared the glycosyl sequences of the alkali-soluble arabinoxylans from the biomass of three model grasses (rice, brachypodium, and foxtail millet) and two potential grass energy crops (switchgrass and miscanthus). Our data show that these arabinoxylans have comparable structures and likely lack the reducing end glycosyl sequence of dicot glucuronoxylans. Rice, brachypodium, and foxtail millet are suitable model grasses for developing technologies to modify arabinoxylan structure and biosynthesis that can be directly extended to grass energy crops.

Materials and Methods:

Plant Material

Stems and leaves of mature *Miscanthus x giganteus, B. distachyon*, and *S. italica* were obtained from the Plant Biology greenhouse at the University of Georgia (Athens, GA). Rice straw was obtained from Dr. Wayne Parrot at the University of Georgia. The air-dried stems and leaves were Wiley-milled (-20/+80 mesh) and stored at room temperature. Milled poplar and

switchgrass were obtained from the National Renewable Energy Laboratory, Golden, CO. The switchgrass (*P. virgatum* cultivar Alamo) was grown in Ardmore Oklahoma (The Samuel Roberts Noble Foundation), milled using a Hammer mill with a 1" screen and then ground in a Wiley mill using a 1mm screen. The milled material was sieved to -20/+80 mesh. A hybrid poplar (*Populus trichocarpa* x *deltoides*) tree was harvested at the Oak Ridge National Laboratory (Oak Ridge, TN) and cut into short logs. The logs were debarked, split with an axe, and chipped using a Yard Machines 10HP chipper. The chips were then milled in a Thomas Scientific Wiley Mill (Model 4) using a 1mm screen size. All the material was then sieved (-20/+80 mesh) using a W.S. Tyler Sieve (Model: RX-29 Type: Rotap).

Preparation of cell walls as their alcohol-insoluble residues (AIR)

Cell walls of switchgrass, rice, miscanthus, brachypodium and foxtail millet were prepared as their alcohol-insoluble residue (AIR) (Mazumder and York 2010). The Wiley-milled biomass from all five grasses was ball-milled for 16 h at 4 °C and 90-100 rpm in aq. 80% EtOH (v/v) using ¼ inch zirconium grinding media (U.S. Stoneware, East Palestine, OH). The insoluble residues were washed with acetone and dried under vacuum.

Sequential extraction of cell wall glycans from AIR

AIR (0.5g) was sequentially extracted with 50 mM ammonium oxalate (35 mL), 1 M KOH containing 1% (w/v) $NaBH_4$, and 4 M KOH containing 1% (w/v) $NaBH_4$. The insoluble residue remaining after treatment with 4 M KOH was then delignified using sodium chlorite and acetic acid (Selvendran et al. 1975). The delignified AIR was then extracted with 4 M KOH containing 1% (w/v) $NaBH_4$ (Selvendran and O'Neill 1987). The 1 M KOH, 4 M KOH and post-chlorite 4 M

KOH extracts were neutralized with glacial acetic acid if required, dialyzed against deionized water (3500 molecular weight cut off tubing), and lyophilized.

Glycosyl-linkage composition analysis by generation of partially methylated alditol acetates

Oligosaccharides (~1 mg) were methylated as described (Ciucanu and Kerek 1984, Mazumder and York 2010). The methylated oligosaccharides (~200μg) were then hydrolyzed for 2 h at 120 °C with 2 M TFA . The released glycoses were converted to their corresponding alditols by treatment with $NaBD_4$ and then treated for 20 min at 50 °C with acetic anhydride-TFA (1:1 v/v) to generate the partially methylated alditol acetate derivatives. The partially methylated alditol acetates were separated and quantified using a Hewlett Packard 5890 gas chromatograph coupled to a Hewlett Packard 5870 mass selective detector (Agilent, Santa Clara, CA) as described (Kulkarni et al. 2012b).

Generation of xylo-oligosaccharides by endo-xylanase treatment of the 1 M KOH-soluble material

Suspensions of the 1 M KOH-soluble materials (~20 mg) from the grasses and from poplar in 50 mM ammonium formate, pH 5, were treated for 24 h at 37 °C with *endo*-xylanse (3.5 unit, *Trichoderma viride* M1, Megazyme, Wicklow, Ireland)(Kulkarni et al. 2012b). The insoluble residues were removed by centrifugation and EtOH then added to the supernatant to a final concentration of 60% (v/v). The mixture was kept for 24 h at 4 °C. The precipitate that formed was removed by centrifugation and the soluble material lyophilized Material enriched in the xylo-oligosaccharides were obtained by fractionating the soluble material on a Superdex 75

HR10/300 size-exclusion chromatography column eluted with 50 mM ammonium formate, pH 5, at 0.5 mL/min and with refractive index detection.

MALDI-TOF mass spectrometry

Positive ion matrix-assisted laser-desorption ionization-time of flight (MALDI-TOF) mass spectra of xylo-oligosaccharides were recorded using a Bruker LT MALDI-TOF mass spectrometer interfaced to a Bruker Biospectrometry workstation (Bruker Daltonics, Billerica, MA). Aqueous samples (1 μL of a 1 mg/mL solution) were mixed with an equal amount of a matrix solution (0.1 M 2,5-dihydroxybenzoic acid in aq. 50% MeCN) and dried on the MALDI target plate. Spectra from 200 laser shots were summed to generate each mass spectrum (Kulkarni et al. 2012b).

^{1}H-NMR spectroscopy

Xylo-oligosaccharides from the five grasses and poplar (~3 mg) were dissolved in D_2O (~0.5 mL, 99.9%) and ^{1}H NMR spectra were recorded with a 600 MHz with a Varian Inova NMR spectrometer (Agilent, Santa Clara, CA). 2D spectra were recorded using standard Varian pulse programs. Chemical shifts were measured relative to internal acetone at δ 2.225.

Glycome profiling

Glycome Profiling of cell wall extracts was performed as described (Pattathil et al. 2012), using a collection of plant cell wall glycan-directed mAbs (Pattathil et al. 2010) (see detailed listing in Supplemental Table 1).

Results and discussion

To compare the arabinoxylans from the five grasses, the alcohol insoluble residue (AIR) was prepared from mature leaves and stem tissue and then sequentially extracted with ammonium oxalate, 1 M KOH containing $NaBH_4$, and 4 M KOH containing $NaBH_4$. The insoluble residues were delignified with chlorite-acetic acid and then extracted with 4 M KOH. Glycosyl residue composition and glycosyl-linkage composition analyses established that most of the solubilized xylan was present in the 1 M KOH extract (**Table 3.1**). Indeed, 1,4 linked and 1,3,4 linked xylosyl residues were predominant in the 1 M and 4 M KOH extracts from the five grasses suggesting the presence of branched arabinoxylans in these wall extracts.

Table 3.1. Glycosyl-linkage compositions of material solubilized by 1 and 4 M KOH treatments of the grass biomass.

Linkage	Brachypodium		Miscanthus		Foxtail millet		Switchgrass		Rice	
	1M KOH	4M KOH	1M KOH	4M KOH	1M KOH	4M KOH	1M KOH	4M KOH	1M KOH	4M KOH
	mol%									
T-Ara*f*[1]	8	10	9	11	9	9	10	8	11	8
T-Xyl*p*	2	3	3	4	4	4	3	2	3	3
4-Xyl*p*	59	61	65	66	61	66	57	58	44	40
3,4-Xyl*p*	11	11	11	10	11	9	11	13	14	10
2,3,4-Xyl*p*	tr	tr	tr	tr	tr	2	tr	tr	2	tr

[a]T-Ara*f* is non-reducing terminal arabinofuranosyl.

Comparison of the structures of the 1 M KOH-soluble arabinoxylans from grasses

To facilitate comparison of arabinoxylan structures, the 1 M KOH-soluble arabinoxylans from each grass were treated with an endo-b-1,4-xylanase to generate xylo-oligosaccharides. These oligosaccharides were isolated by size-exclusion chromatography (SEC) and then structurally characterized using MALDI-TOF-MS and one and two dimensional ^{1}H NMR spectroscopy. The glucuronoxylans from dicots including Arabidopsis (York and O'Neill 2008) and poplar (Lee et al. 2011b) have the glycosyl sequence [,4)-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-Gal*p*A-(1,4)-D-Xyl] at their reducing end. This reducing end sequence has been implicated in controlling the degree of polymerization (DP) of dicot glucuronoxylans (York and O'Neill 2008) and is a target for modifying xylan structure to improve the bioprocessing characteristics of woody biomass. It is not known if this glycosyl sequence is present in grass arabinoxylans. Thus, xylo-oligosaccharides were generated from the 1 M KOH-soluble glucuronoxylan from poplar wood biomass to allow us to compare the reducing end sequences of grass and dicots xylans.

MALDI-TOF-MS showed that the most abundant xylo-oligosaccharides generated from the grasses and poplar have molecular masses between 700 Da and 1400 Da (**Figure. 3.1**). The predominant ions ($[M+Na]^+$ and ($[M-H+2Na]^+$) in the grass spectra were consistent with the presence of oligosaccharides composed mainly of pentosyl (Xyl and Ara) residues (**Figure 3.1A-E**). In grasses, the major ion at *m/z* 701 corresponds to $[M+Na]^+$ of an oligosaccharide composed of five pentosyl residues, consistent with the presence of arabinoxylan. The most abundant poplar oligosaccharides are substituted with a MeGlcA residue to give monoisotopic $[M+Na]^+$ ions at *m/z* 759, 891 and 1023 and monoisotopic $[M-H+2Na]^+$ ions at *m/z* 781, 913 and 1045

(**Figure 3.1F**). The monoisotopic ions at *m/z* 759 and 891 corresponds to oligosaccharides that are composed of four xylosyl residues substituted with a MeGlcA and five xylosyl residues substituted with a MeGlcA respectively. These ions also correspond to monoisotopic $[M+Na]^+$ ion of the sequence [β-D-Xyl*p*-(1,4)]-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-Gal*p*A-(1,4)-D-Xyl that has been observed at the reducing end of glucuronoxylans from several dicots (Peña et al. 2007). Weak signals that correspond to oligosaccharides composed of pentose residues along with a single GlcA or MeGlcA residue were observed in the spectra of the endo-xylanase-generated fragments of the grass arabinoxylans (**Figure 3.1 A-E**). However, due to mass degeneracy, the MALDI spectra do not provide sufficient information to determine whether the reducing end sequence characteristic of dicot glucuronoxylan is present in the grass xylans.

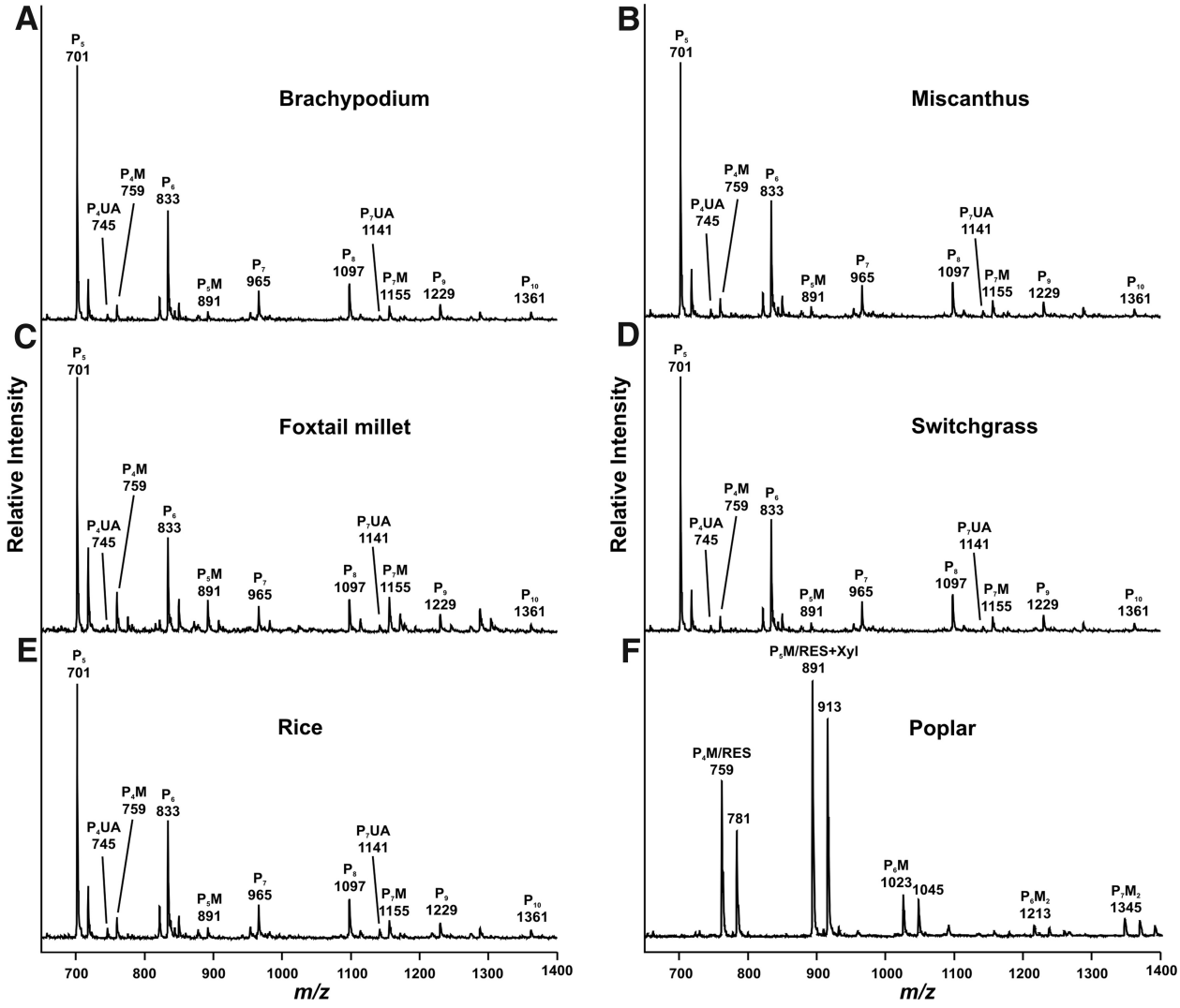


Figure 3.1: MALDI-TOF spectrum of the purified xylo-oligosaccharides generated by xylanase treatment of the 1 M KOH extracts grasses **(A-E);** and poplar **(F)**. Ions are labeled with the most likely ([M+Na]+) assignments, but ions with more than two sodium atoms are also possible. In poplar, the reducing end sequence is represented as RES.

To confirm and extend the results obtained by MALDI-TOF-MS, the endo-xylanase-generated xylo-oligosaccharides from the grasses and from poplar were further characterized using ^{1}H-NMR spectroscopy. The chemical shifts and scalar coupling constants for the protons of each glycose were assigned based on 2D gCOSY analyses (**Figure 3.2**) and published data (Verbruggen et al. 1998, Peña et al. 2007, Mazumder and York 2010, Kulkarni et al. 2012b). Signals were assigned for the anomeric and ring proton resonances for the terminal non-reducing β-D-xylosyl, internal 4-linked β-D-xylosyl, internal 2,4-linked β-D -xylosyl, internal 3,4-linked β-D-xylosyl and reducing xylosyl residues as well as for the terminal a-L-Ara*f* and a-L-Ara*f*-(1,2)-a-L-Ara*f* side chains at *O*-3 of the xylosyl backbone and α-D-Glc*p*A residues and 4-O-Me GlcpA residues linked to O2 of the xylosyl backbone (residues A-L, **Table 3.2, Figure 3.2**). The spectra of all of the grass xylo-oligosaccharide samples are remarkably similar, showing only minor difference in the cross-peak patterns (**Figure 3.2A-E**). However, the poplar xylo-oligosaccharide spectrum is completely different, with relatively few cross-peaks in common with the grass xylo-oligosaccharide spectra (**Figure 3.2F**). These spectra confirmed that the arabinoxylans from different grass species have very similar structures that are distinct from the structure of poplar glucuronoxylan. However, the gCOSY spectra of the grass arabinoxylans are not identical. Specifically, nearly all of the glucuronic acid residues in the Miscanthus and switchgrass (**Figure 3.2B and 3.2D**) arabinoxylans are methylated at O4 (residue I) while some of the glucuronic acid residues in the Brachypodium, foxtail millet and rice arabinoxylans (**Figure 3.2A, 3.2C, and 3.2E**) are not methylated (residue G). Although this structural discrepancy may be species-specific, it may also reflect developmental differences in the specific tissues that were analyzed. Such relatively minor structural differences are unlikely to preclude using Brachypodium, foxtail

millet and rice as model species to gain insight into the biosynthesis of arabinoxylans in other grasses, including the biomass crops Miscanthus and switchgrass.

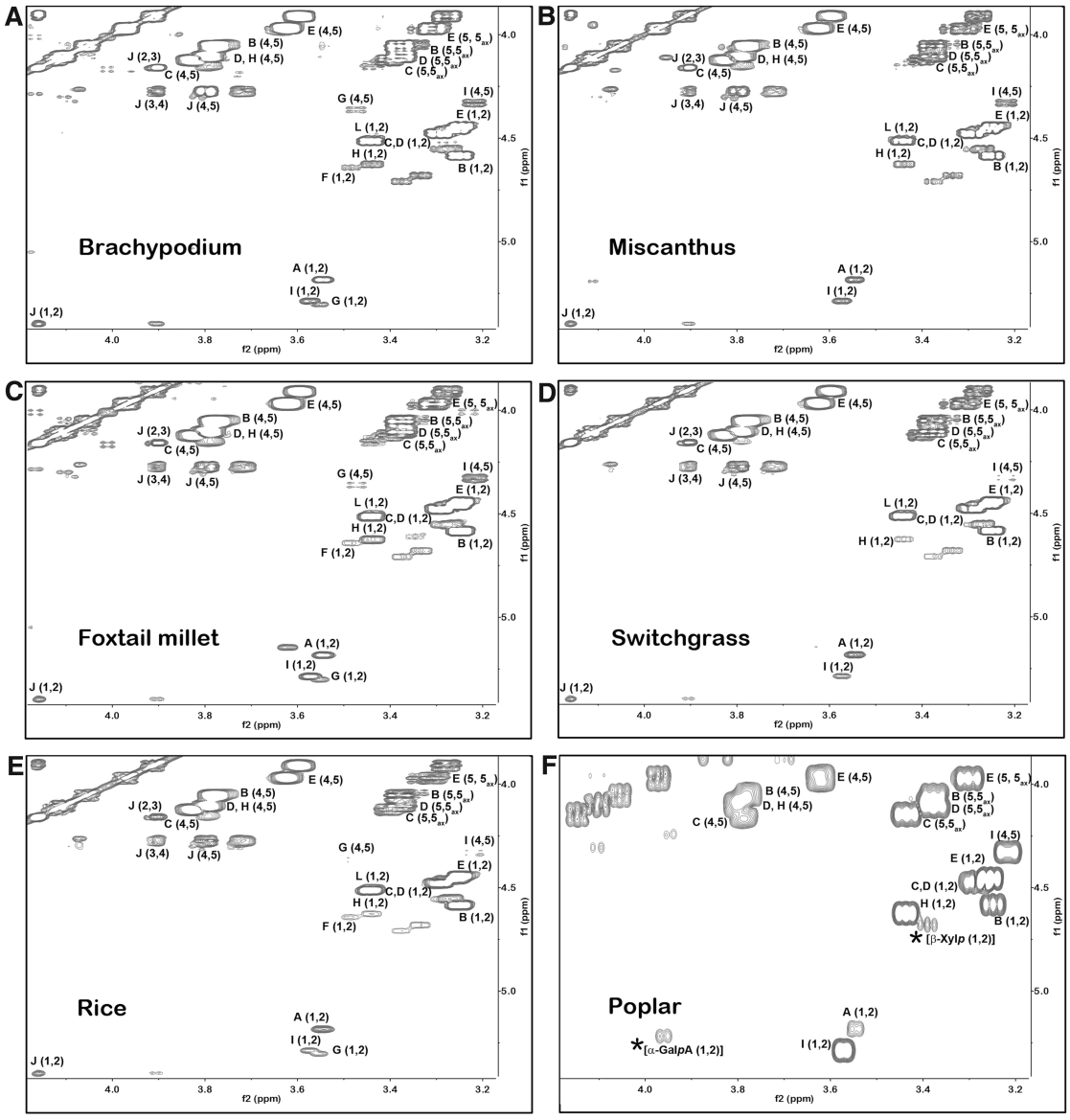


Figure 3.2. Partial 600-MHz gCOSY NMR spectrum of purified xylo-oligosaccharides generated by β-endoxylanase digestion of the 1 M KOH extracts of AIR from grasses **(A-E)**; and poplar **(F)**. Crosspeak assignments are indicated using an uppercase letter to indicate the glycosyl residue that contains the protons (*Table 2*) and numbers indicating the position of the protons in the residue. Resonances by crosspeaks (*) due to the presence of the reducing end sequence of the glucuronoxylan from poplar are also indicated.

Table 3.2. ^{1}H NMR assignments of the xylo-oligosaccharides generated by endoxylanase treatment of the 1 M KOH extract of AIR from five grasses and poplar.

Key	Residue	H1	H2	H3	H4	H5	$H5_{ax}$
A	α-Xyl*p* (reducing)	5.183-5.185	3.542-3.544	3.77			
B	β-Xyl*p* (reducing)	4.583-4.585	3.251-3.253	3.545-3.548	3.77	3.374-3.376	4.055-4.057
C	β-1,4-Xyl*p* (internal) (major)	4.483	3.285-3.287	3.585-3.586	3.827	4.412	3.418
D	β-1,4-Xyl*p* (internal) (minor)	4.470	3.282	3.56	3.79	3.37	4.10
E	β-Xyl*p* (terminal)	4.455-4.457	3.255-3.257	3.427-3.431	3.620-3.622	3.970-3.97	3.301-3.303
F	β-1,2,4-Xyl*p* (α-Glc*p*A)	4.643-4.646	3.480-3.485	3.625-3.635	3.798-3.810	3.388-3.393	4.107-4.109
G	α-Glc*p*A	5.302-5.305	3.552-3.554	3.738-3.745	3.461-3.470	---	---
H	β-1,2,4-Xyl*p* (α-4Me-Glc*p*A)	4.624-4.629	3.436-3.439	3.62	3.80	3.383-3.390	4.102-4.108
I	α-4Me-Glc*p*A	5.286-5.290	3.571-3.573	3.758-3.765	3.216-3.220	4.328-4.332	---
J	α-L-Ara*f* (terminal)	5.398	4.158	3.910	4.298	3.815	3.7
K	α-L-Ara*f*	5.540	4.269	4.073	4.308	3.818	3.7
L	β-1,3,4-Xylp (α-L-Ara*f*)	4.515-4.522	3.443-3.446	3.770-3.783	3.85	4.140-4.152	3.420-3.428

Chemical shifts are reported in ppm relative to internal acetone, δ 2.225. β-Xyl (α-Glc*p*A) is a β-linked xylosyl residue that bears a Glc*p*A sidechain at O-2. β-Xyl (α-L-Ara*f*) is a β-linked xylosyl residue that bears a α-L-Ara*f* sidechain at O-3. H-4 and H-5 of the reducing α-xylose were not assigned. The number of decimal places for each chemical shift reflects the accuracy of the measurement. Residues are indicated by an uppercase letter as a key for cross referencing with Figure 2.

The signals for 4)-β-D-Xylp-(1,4)-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-Gal*p*A-(1,4)-D-Xyl*p* (the reducing end sequence characteristic of dicot glucuronoxylan) (Peña et al. 2007, York and O'Neill 2008) are clearly visible in the NMR spectra of poplar xylo-oligosaccharides (**Figure 3.2F**). By contrast, these signals are not discernible in the NMR spectra of the grass xylo-oligosaccharides implying that this reducing end sequence, if present at all, is a very minor component in the arabinoxylans present in grass cell walls. As this reducing end sequence has been implicated in controlling DP of dicot glucuronoxylans (York and O'Neill 2008), its apparent absence in grass arabinoxylans suggests that xylan DP is controlled by different mechanisms in grasses and in dicots.

Glycome profiling of grass biomass using a suite of cell wall polysaccharide-directed monoclonal antibodies

Glycome Profiling utilizing a suite of monoclonal antibodies (mAbs) that recognize diverse cell wall polysaccharide epitopes was used to determine the major types of polysaccharides present in the grass cell walls and how readily they are solubilized by aqueous buffer and by alkali before and after delignification. The AIR isolated from switchgrass, rice, brachypodium, miscanthus and foxtail millet was sequentially extracted with 50 mM ammonium oxalate, sodium carbonate, 1 M KOH, 4 M KOH, acidified-chlorite, and finally (post-chlorite) 4 M KOH. Most of the solubilized polysaccharide material was present in the 1 M KOH and 4 M KOH extracts (**Figure 3.3A**).

Glycome profiling showed strong binding of all of the xylan-3 and xylan-4 groups of mAbs and some of the antibodies in the xylan-2 mAbs to the 1 M KOH, 4 M KOH, chlorite-

acetic acid and post chlorite 4 M KOH extracts. These results are consistent with published reports (Ebringerová et al. 2005, Faik 2010, Scheller and Ulvskov 2010) that the cell walls of grasses contain arabinoxylan. In addition, strong binding of the mAb, BG1, which binds to (1→3,1→4)β-glucans (Meikle et al. 1994), was observed in the alkali and acidified-chlorite extracts. Miscanthus and rice also showed binding of the mAb, LAMP, which recognizes (1→3)β-glucans (callose) (Meikle et al. 1991), in these same extracts. The presence of the mixed-linkage β-glucan in the cell walls of mature grass tissue was somewhat unexpected, as this polysaccharide is typically associated with the cell walls of rapidly growing tissues (Carpita 1996). Likewise, the presence of detectable amounts of callose in the extracts from some grasses was unexpected, as callose deposition is typically associated with wound responses or pathogenesis (Jacobs et al. 2003b).

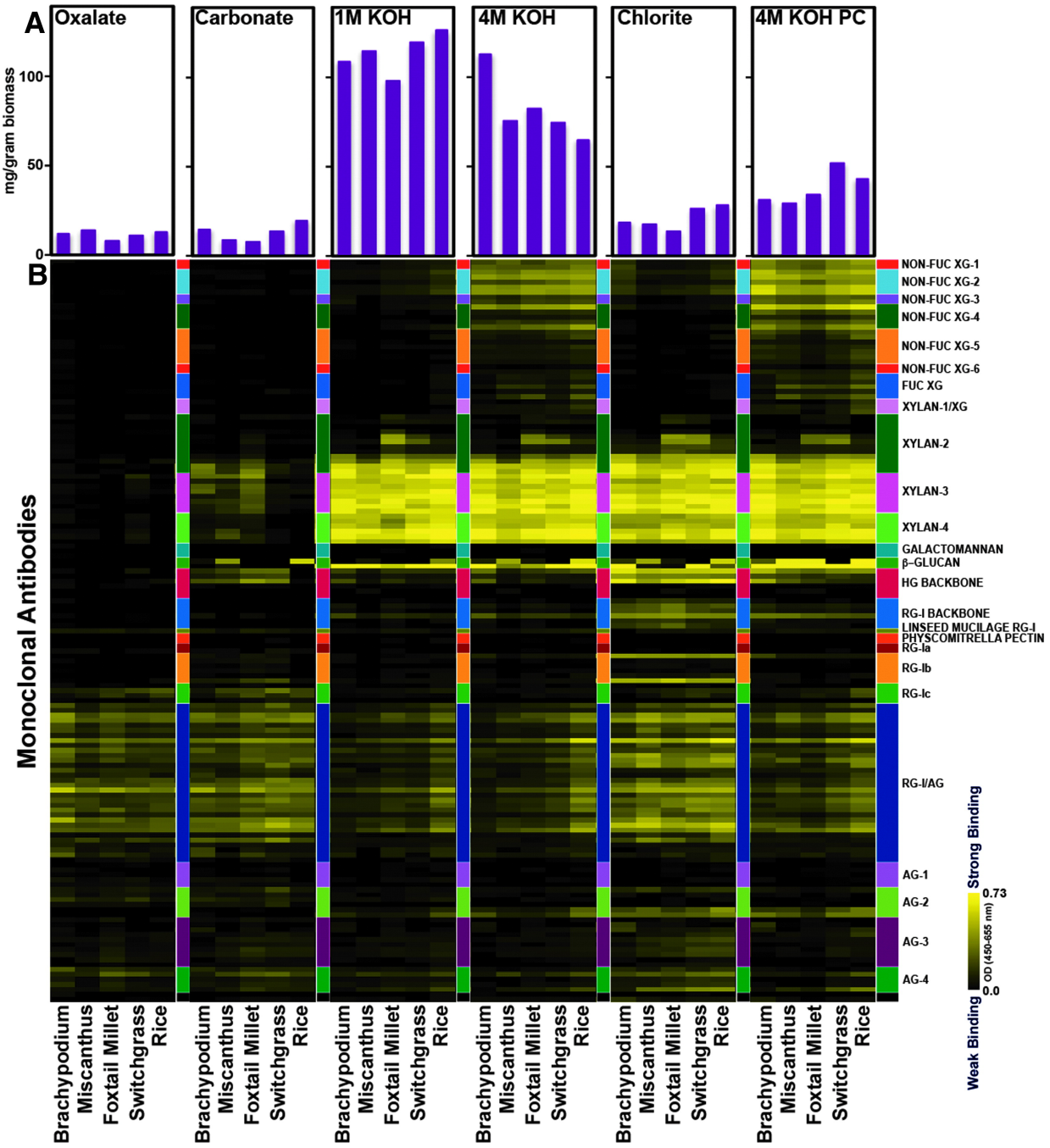


Figure 3.3. Glycome profiling of five grass cell wall extracts. The extracted AIR by various reagents were loaded onto ELISA plates and were screened against an array of plant glycan-directed monoclonal antibodies.16 The panel on the right of the figure lists the individual antibodies used, grouped according to the polysaccharides predominantly recognized by these

antibodies. Antibody binding is represented as a colored heat map, with black signifying no binding and bright yellow representing the strongest binding. The bar graphs at the top indicate the amount of solubilized material recovered at each extraction step per gram of AIR.

In all five grasses the binding of mAbs that recognize epitopes other than xylan and β-glucan was comparable. For example, xyloglucan epitopes were more abundant than pectic/arabinogalactan epitopes (HG backbone, RG-I backbone, RG-Ia, -Ib, -Ic, RG-I/AG, AG-1, AG-2, AG-3 and AG-4 groups of mAbs) in 4 M KOH and post chlorite 4 M KOH extracts. The opposite was true in the acidified chlorite extracts. Interestingly, xylan, pectin and pectic arabinogalactan epitopes were abundant in the acidified chlorite extract suggesting, that some of the xylan and pectin is associated with lignin in grass biomass. Together, our results suggest that the cell walls of the five grasses studied contain similar polysaccharides, although there are subtle differences in their glycan epitope compositions.

In conclusion, we have used a combination of chemical, spectroscopic and immunological techniques to show that similar pectic and hemicellulosic polysaccharides are present in biomass from switchgrass, miscanthus, rice, brachypodium, and foxtail millet. The quantitatively predominant hemicellulose is arabinoxylan. Some subtle structural differences including the degree of backbone branching of the xylan backbone and the extent of binding of xylan-directed antibodies were observed. Our study provides evidence that grass arabinoxylans lack discernible amounts of the distinct reducing end sequence of glycoses that is present in glucuronoxylans from hardwoods and softwoods. Nevertheless, the high degree of structural similarity for arabinoxylans from rice, brachypodium and foxtail millet suggests that these three plants provide useful models for identifying fundamental aspects of xylan biosynthesis in commercial grass bioenergy crops such including Miscanthus and switchgrass.

Acknowledgements

This research was funded by the Bio-Energy Science Center (BESC) administered by Oak Ridge National Laboratory and funded by a grant (DE-AC05-00OR22725) from the Office of Biological and Environmental Research, Office of Science, U.S. Department of Energy. The generation of the CCRC series of plant cell wall glycan-directed monoclonal antibodies used in this work was supported by the NSF Plant Genome Program (DBI-0421683). We would like to acknowledge Maria J. Peña of the Complex Carbohydrate Research Center for discussions with respect to NMR spectroscopy.

Disclosure Statement

No competing financial interests exist

References:

Ahlgren PA, Goring DAI. 1971. Removal of wood components during chlorite delignification of black spruce. *Can. J. Chem* 49:1272-1275.

Ahola V, Aittokallio T, Vihinen M, Uusipaikka E. 2006. A statistical score for assessing the quality of multiple sequence alignments. *BMC Bioinformatics.* 7:484.

Aspeborg H, Schrader J, Coutinho PM, Stam M, Kallas A, Djerbi S, Nilsson P, Denman S, Amini B, Sterky F*, et al.* 2005. Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen. *Plant Physiol.* 137:983-997.

Atkins E, Parker K, Preston R. 1969. The Helical structure of the -1, 3-linked xylan in some siphoneous green algae. *Proc. Roy. Soc. B.* 173:209-221.

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

Bremner I, Wilkie K. 1966. The hemicelluloses of bracken:: Part I. An acidic xylan. *Carbohydr Res.* 2:24-34.

Brown DM, Zhang Z, Stephens E, Dupree P, Turner SR. 2009. Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *Plant J.* 57:732-746.

Carafa A, Duckett JG, Knox JP, Ligrone R. 2005. Distribution of cell-wall xylans in bryophytes and tracheophytes: new insights into basal interrelationships of land plants. *New Phytol.* 168:231-240.

Carpita NC. 1996. Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47:445-476.

Carpita NC, Kanabus J. 1987. Extraction of starch by dimethyl sulfoxide and quantitation by enzymatic assay. *Anal. Biochem.* 161:132-139.

Cartmell A, McKee LS, Peña MJ, Larsbrink J, Brumer H, Kaneko S, Ichinose H, Lewis RJ, Vikso-Nielsen A, Gilbert HJ*, et al.* 2011. The structure and function of an arabinan-specific alpha-1,2-arabinofuranosidase identified from screening the activities of bacterial GH43 glycoside hydrolases. *J. Biol. Chem.* 286:15483-15495.

Ciucanu I, Kerek F. 1984. A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res.* 131:209-217.

Correia MAS, Mazumder K, Bras JLA, Firbank SJ, Zhu Y, Lewis RJ, York WS, Fontes CMGA, Gilbert HJ. 2011. The structure and function of an arabinoxylan-specific xylanase. *J. Biol. Chem.* 286:22510-22520.

Dalessandro G, Northcote DH. 1977. Changes in enzymic activities of nucleoside diphosphate sugar interconversions during differentiation of cambium to xylem in pine and fir. *Biochem. J.* 162:281-288.

de O. Buanafina MM. 2009. Feruloylation in Grasses: Current and Future Perspectives. *Mol. Plant* 2:861-872.

Domon B, Costello CE. 1988. A Systematic Nomeclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconj. J.* 5:397-409.

Domozych DS, Sorensen I, Willats WG. 2009. The distribution of cell wall polymers during antheridium development and spermatogenesis in the Charophycean green alga, Chara corallina. *Ann. Bot.* 104:1045-1056.

Doust AN, Kellogg EA, Devos KM, Bennetzen JL. 2009. Foxtail millet: a sequence-driven grass model system. *Plant Physiol.* 149:137-141.

Ebringerová A, Hromadkova Z, Heinze T. 2005. Hemicellulose. *Adv. Polym. Sci.* 186:1-67.

Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics.* 14:755-763.

Evert RF. 2006. Esau's plant anatomy: meristems, cells, and tissues of the plant body: their structure, function, and development Wiley-Liss.

Faik A. 2010. Xylan biosynthesis: News from the grass. *Plant Physiol.* 153:396-402.

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R*, et al.* 2006. Pfam: clans, web tools and services. *Nucleic Acids Res.* 34:D247-251.

Fu H, Yadav MP, Nothnagel EA. 2007. *Physcomitrella patens* arabinogalactan proteins contain abundant terminal 3-O-methyl-L: -rhamnosyl residues not found in angiosperms. *Planta.* 226:1511-1524.

Graham L, Cook M, Busse J. 2000. The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proc. Natl. Acad. Sci. USA* 97:4535.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52:696-704.

Heaton EA, Dohleman FG, Long SP. 2008. Meeting US biofuel goals with less land: the potential of Miscanthus. *Global Change Biol.* 14:2000-2014.

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315:804-807.

Hiwatashi Y, Nishiyama T, Fujita T, Hasebe M. 2001. Establishment of gene-trap and enhancer-trap systems in the moss *Physcomitrella patens*. *Plant J.* 28:105-116.

Izydorczyk M, Biliaderis C. 1995. Cereal arabinoxylans: advances in structure and physicochemical properties. *Carbohydr. Polym.* 28:33-48.

Jacobs AK, Lipka V, Burton RA, Panstruga R, Strizhov N, Schulze-Lefert P, Fincher GB. 2003. An Arabidopsis callose synthase, GSL5, is required for wound and papillary callose formation. *The Plant Cell Online* 15:2503-2513.

Johansson M, Samuelson O. 1977. Reducing end groups in brich xylan and their alkaline degradation. *Wood Sci. Technol.* 11:251-263.

Katoh K, Kuma K, Toh H, Miyata T. 2005a. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33:511-518.

Katoh K, Kuma K, Toh H, Miyata T. 2005b. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic acids research* 33:511-518.

Kong Y, Zhou G, Avci U, Gu X, Jones C, Yin Y, Xu Y, Hahn MG. 2009. Two poplar glycosyltransferase genes, PdGATL1.1 and PdGATL1.2, are functional orthologs to PARVUS/AtGATL1 in Arabidopsis. *Mol. Plant* 2:1040-1050.

Kremer C, Pettolino F, Bacic A, Drinnan A. 2004. Distribution of cell wall components in Sphagnum hyaline cells and in liverwort and hornwort elaters. *Planta* 219:1023-1035.

Kulkarni AR, Pena MJ, Avci U, Mazumder K, Urbanowicz BR, Pattathil S, Yin Y, O'Neill MA, Roberts AW, Hahn MG*, et al.* 2012. The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes. *Glycobiology* 22:439-451.

Lee C, Teng Q, Huang W, Zhong R, Ye Z. 2009. The poplar GT8E and GT8F glycosyltransferases are functional orthologs of Arabidopsis PARVUS involved in glucuronoxylan biosynthesis. *Plant Cell Physiol.* 50:1982.

Lee C, Teng Q, Huang W, Zhong R, Ye ZH. 2010. The Arabidopsis family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. *Plant Physiol.* 153:526-541.

Lee C, Teng Q, Zhong R, Ye ZH. 2011. Molecular dissection of xylan biosynthesis during wood formation in poplar. *Mol. Plant* 4:730-747.

Lewandowski I, Scurlock JMO, Lindvall E, Christou M. 2003. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. *Biomass Bioenergy* 25:335-361.

Liepman AH, Wightman R, Geshi N, Turner SR, Scheller HV. 2010. Arabidopsis - a powerful model system for plant cell wall research. *Plant J.* 61:1107-1121.

Ligrone R. 1986. Structure, development and cytochemistry of mucilage-secreting hairs in the moss Timmiella-barbuloides (brid) moenk. *Ann. Bot.* 58:859-868.

Mandal P, Pujol C, Damonte E, Ghosh T, Ray B. 2009a. Xylans from Scinaia hatei: structural features, sulfation and anti-HSV activity. *Int. J. Biol. Macromol.*

Mandal P, Pujol C, Damonte E, Ghosh T, Ray B. 2009b. Xylans from Scinaia hatei: Structural features, sulfation and anti-HSV activity. *Int. J. Biol. Macromol.* 46:173-178.

Masuko T, Minami A, Iwasaki N, Majima T, Nishimura S, Lee YC. 2005. Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. *Anal. Biochem.* 339:69-72.

Matsunaga T, Ishii T, Matsumoto S, Higuchi M, Darvill A, Albersheim P, O'Neill M. 2004. Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. *Plant Physiol.* 134:339-351.

Matulewicz M, Cerezo A. 1987. Alkali-soluble polysaccharides from *Chaetangium fastigiatum*: Structure of a xylan. *Phytochemistry* 26:1033-1035.

Mazumder K, York WS. 2010. Structural analysis of arabinoxylans isolated from ball milled switchgrass biomass. *Carbohydr. Res.* 345:2183-2193.

McCartney L, Marcus SE, Knox JP. 2005. Monoclonal antibodies to plant cell wall xylans and arabinoxylans. *J. Histochem. Cytochem.* 53:543-546.

Meikle PJ, Bonig I, Hoogenraad NJ, Clarke AE, Stone BA. 1991. The location of (1-->3)-β-glucans in the walls of pollen tubes of *Nicotiana-alata* using a (1-->3)-β-glucan-specific monoclonal-antibody. *Planta* 185:1-8.

Meikle PJ, Hoogenraad NJ, Bonig I, Clarke AE, Stone BA. 1994. A (1-->3,1-->4)-β-glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1-->3,1-->4)-β-glucans. *Plant J.* 5:1-9.

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

Mishler B, Kelch D. 2009. Phylogenomics and early land plant evolution. *Bryophyte Biology*:173–197.

Moller I, Sorensen I, Bernal AJ, Blaukopf C, Lee K, Obro J, Pettolino F, Roberts A, Mikkelsen JD, Knox JP*, et al.* 2007. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *Plant J.* 50:1118-1128.

Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.* 96:673-686.

Mur LA, Allainguillaume J, Catalan P, Hasterok R, Jenkins G, Lesniewska K, Thomas I, Vogel J. 2011. Exploiting the Brachypodium Tool Box in cereal and grass research. *New Phytol.* 191:334-347.

Niklas K. 1997. The evolutionary biology of plants University of Chicago Press.

Nuin PA, Wang Z, Tillier ER. 2006. The accuracy of several multiple sequence alignment programs for proteins. *BMC Bioinformatics* 7:471.

Pattathil S, Avci U, Baldwin D, Swennes AG, McGill JA, Popper Z, Bootten T, Albert A, Davis RH, Chennareddy C*, et al.* 2010. A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol.* 153:514-525.

Pattathil S, Avci U, Miller JS, Hahn MG. 2012. Immunological approaches to plant cell wall and biomass characterization: glycome profiling. In: Himmel ME editor. Biomass Conversion: Methods and Protocols: Humana Press.

Pauly M, Albersheim P, Darvill A, York WS. 1999. Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. *Plant J* 20:629-639.

Peña MJ, Darvill AG, Eberhard S, York WS, O'Neill MA. 2008. Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology* 18:891-904.

Peña MJ, Zhong R, Zhou GK, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye ZH. 2007. Arabidopsis irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19:549-563.

Percival E, Chanda S. 1950. The xylan of *Rhodymenia palmata*.

Popper Z. 2011. Evolution and diversity of plant cell walls: from algae to flowering plants. *Ann. Rev. Plant Biol.* 62:567-590.

Popper ZA. 2008. Evolution and diversity of green plant cell walls. *Curr. Opin. Plant Biol.* 11:286-292.

Popper ZA, Fry SC. 2003. Primary cell wall composition of bryophytes and charophytes. *Ann. Bot.* 91:1-12.

Popper ZA, Tuohy MG. 2010. Beyond the green: understanding the evolutionary puzzle of plant and algal cell walls. *Plant Physiol.* 153:373-383.

Qiu Y, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J. 2006. The deepest divergences in land plants inferred from phylogenomic evidence. *Proc. Natl. Acad. Sci. U S A* 103:15511-15516.

Scheller HV, Ulvskov P. 2010. Hemicelluloses. *Ann. Rev. Plant Biol.* 61:263-289.

Selvendran RR, Davies AMC, Tidder E. 1975. Cell wall glycoproteins and polysaccharides of mature runner beans. *Phytochemistry* 14:2169-2174.

Selvendran RR, O'Neill MA. 1987. Isolation and analysis of cell walls from plant material. *Methods Biochem. Anal.* 32:25-152.

Shimizu K, Ishihara M, Ishihara T. 1976. Hemicellulases of brown rotting fungus Tyromyces palustris. II. The oligosaccharides from the hydrolysate of a hardwood xylan by the intracellular xylanase. *Mokuzai Gakkaishi*:618-625.

Smith B, Harris P. 1999. The polysaccharide composition of Poales cell walls:: Poaceae cell walls are not unique. *Biochem. Syst. Ecol.* 27:33-53.

Sorensen I, Domozych D, Willats WG. 2010. How have plant cell walls evolved? *Plant Physiol.* 153:366-372.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24:1596.

Taylor T, Taylor E, Krings M. 2009. Paleobotany: The biology and evolution of fossil plants Academic Pr.

Turvey JR, Williams EL. 1970. Structures of some xylans from red algae. *Phytochemistry* 9:2383-&.

Verbruggen MA, Spronk BA, Schols HA, Beldman G, Voragen AG, Thomas JR, Kamerling JP, Vliegenthart JF. 1998. Structures of enzymically derived oligosaccharides from sorghum glucuronoarabinoxylan. *Carbohydr. Res.* 306:265-274.

Vogel J. 2008. Unique aspects of the grass cell wall. *Curr. Opin. Plant Biol.* 11:301-307.

Wise LE, Murphy M, Daddieco AA. 1946. Chlorite Holocellulose, Its Fractionation and Bearing on Summative Wood Analysis and on Studies on the Hemicelluloses. *Tech Assoc Pap* 29:210-218.

Wu AM, Hornblad E, Voxeur A, Gerber L, Rihouey C, Lerouge P, 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 Physiol.* 153:542-554.

Wu AM, Rihouey C, Seveno M, Hornblad E, Singh SK, Matsunaga T, Ishii T, Lerouge P, Marchant A. 2009. The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J.* 57:718-731.

Yamagaki T, Maeda M, Kanazawa K, Ishizuka Y, Nakanishi H. 1997. Structural clarification of Caulerpa cell wall -1, 3-xylan by NMR spectroscopy. *BioSci., BioTechnol., and Biochem* 61:1077-1080.

Yin Y, Chen H, Hahn MG, Mohnen D, Xu Y. 2010. Evolution and function of the plant cell wall synthesis-related Glycosyltransferase Family 8. *Plant Physiol.*

Yoon H, Hackett J, Ciniglia C, Pinto G, Bhattacharya D. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 21:809.

York WS, O'Neill MA. 2008. Biochemical control of xylan biosynthesis - which end is up? *Curr. Opin. Plant Biol.* 11:258-265.

Zhong R, Peña MJ, Zhou GK, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WH, 3rd, Darvill AG, York WS, Ye ZH. 2005. Arabidopsis fragile fiber8, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. *Plant Cell* 17:3390-3408.

Zhong R, Ye ZH. 2003. Unraveling the functions of glycosyltransferase family 47 in plants. *Trends Plant Sci.* 8:565-568.

Zhong R, Ye ZH. 2007. Regulation of cell wall biosynthesis. *Curr. Opin. Plant Biol.* 10:564-572.

Zhou GK, Zhong R, Richardson EA, Morrison WH, 3rd, Nairn CJ, Wood-Jones A, Ye ZH. 2006. The poplar glycosyltransferase GT47C is functionally conserved with Arabidopsis Fragile fiber8. *Plant Cell Physiol.* 47:1229-1240.

Chapter 4

STRUCTURAL DIVERSITY OF THE REDUCING END SEQUENCES OF HETEROXYLANS PRESENT IN THE CELL WALLS OF MONOCOTYLEDENOUS PLANTS

[1]Ameya R. Kulkarni, Malcolm A. O'Neill, Maria .J. Peña, and William S. York, and. (2013). Diversity in the structures and reducing end sequences of heteroxylans present in the cell walls of monocotyledonous plants. To be submitted to Plant Physiology.

Abstract

Knowledge of the molecular and biochemical factors that control heteroxylan synthesis and structure in woody plants and grasses is required to facilitate the development of technologies that improve the economic value of biomass derived from these plants. The characteristic glycosyl sequence -4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-Gal*p*A-(1,4)-D-Xyl (i.e., Sequence 1) is present at the reducing end of glucuronoxylans in the cell walls of several eudicot and gymnosperm species, Sequence 1 is required for normal xylan synthesis, growth and development in these plants. However, the reducing end glycosyl sequences of heteroxylans in the cell walls of monocots have not been characterized. Here, we show that the heteroxylans isolated from the cell walls of selected members of the Acorales, Alismatales, Dioscoreales, Pandanales, Liliales, Arecales, Commelinales, and Zingerberales are similar to eudicot and gymnosperm heteroxylans in that they have Sequence 1 at the reducing end. By contrast, the arabinoxylans isolated from members of the Poales (Poaceae, Cyperaceae, and Bromeliaceae) and Asparagales lack sequence 1. Two reducing end sequences were identified in the xylans isolated from switchgrass, miscanthus, brachypodium, rice, and foxtail millet stem cell walls and structurally characterized. In one sequence the reducing xylose is substituted at *O*-2 with 4-*O*-methyl glucuronic acid and in the second sequence the reducing xylose is substituted at *O*-3 with Ara*f*. Most of the reducing end xylose is also substituted at *O*-2 with 4-*O*-methyl glucuronic acid in the glucuronoxylans isolated from members of the Asparagales. Thus, sequence 1 is absent in the cell walls of some but not all Poales species, indicating that it was not eliminated when monocots and dicots diverged.

Introduction

Monocots are a diverse group of land plants that include species of economic and ecological importance (Bremer *et al.* 2009, Givnish *et al.* 2010). Monocots comprise of approximately 65,000 species that are distributed in 12 orders and 82 families (Bremer, et al. 2009, Givnish, et al. 2010). The Poales, which includes the Poaceae (grasses), Cyperaceae (sedges) and Bromeliaceae (bromeliads), are considered to be the most diverse order in the monocots. These plants together with the Dasypogonaceae, Zingiberales, Commelinales, and Arecales (palms) constitute the commelinind clade within the monocots, with the Arecales being most closely related to non-commelinids (Givnish, et al. 2010). Non-commelinind monocots are a paraphyletic group consisting of the Asparagales, Liliales, Pandanales, Dioscoreales, Petrosaviales, Alismatales, and Acorales. The Acorales are believed to be sister to all other monocots (Bremer, et al. 2009, Chase *et al.* 2006, Givnish, et al. 2010). With the exception of the grasses, very little information is available about the structure of the heteroxylans present in the cell walls of the monocots (Buchala and Meier 1972, Carnachan and Harris 2000, Smith and Harris 1999, Smith and Harris 1995).

Arabinoxylans (AX) account for up to 30% of the mass of the cell walls of grasses (Kulkarni *et al.* 2012a). These xylans have a backbone composed of 1,4-linked β-D-xylopyranosyl (Xyl*p*) residues (**Figure 4.1**). Some of the Xyl*p* residues are substituted at *O*-3 with α-L-Ara*f* and α-L-Ara*f*-(1,2)-α-L-Ara*f* (Anders *et al.* 2012). The Ara*f* residues may themselves be esterified with ferulic or coumaric acids, a feature considered to be a unique characteristic of grass AX (Faik 2010). Small amounts of Glc*p*A or 4-*O*-methyl GlcpA have also been shown to be linked to the backbone of grass arabinoxylans (Faik 2010, Kulkarni, et al. 2012a). The limited amount of published data suggest that the walls of other commelinids also

contain arabinoxylan, whereas the walls of the non-commelinids are likely to contain glucuronoxylans (GX)(Buchala and Meier 1972, Carnachan and Harris 2000, Gibeaut *et al.* 2005, Jacobs *et al.* 2003, Simas-Tosin *et al.* 2013, Smith and Harris 1999). The backbone residues of such GXs, which are abundant in the secondary walls of dicots, are substituted at *O*-2 with α-D-glucosyluronic acid (α-D-Glc*p*A) or 4-O-methyl α-D-glucosyluronic acid (4-O-Me-Glc*p*A) (Ebringerová *et al.* 2005, Scheller and Ulvskov 2010).

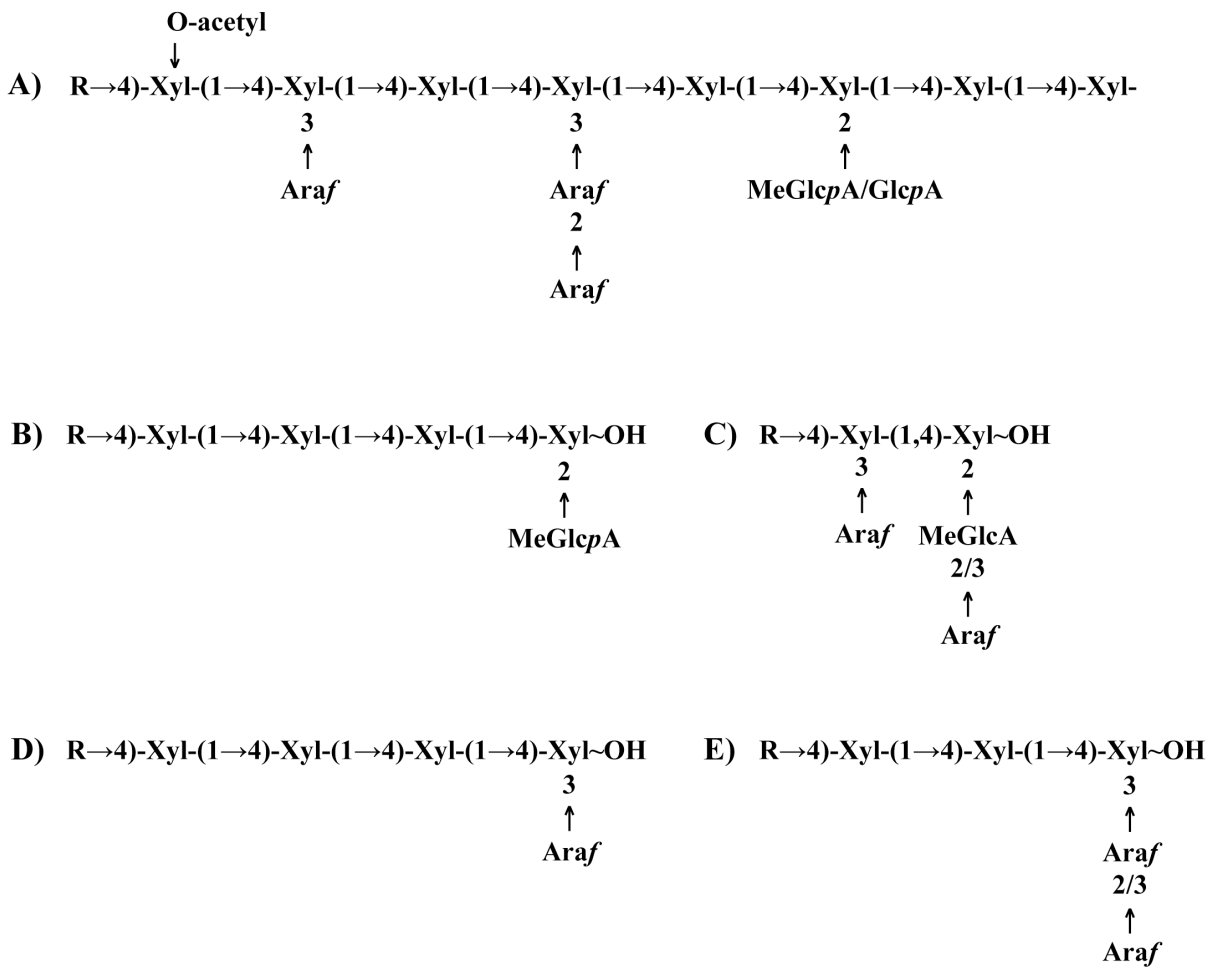


Figure 4.1: Structure of glycosyl sequences in grasses.

(A) General structure of glucuronoarabinoxylan in grasses.

(B) The glycosyl sequence ARK-1 at the reducing end of grasses.

(C) The glycosyl sequence ARK-2 at the reducing end of grasses. This structure was found to be a minor component of the cell walls.

(D) The glycosyl sequence ARK-3 at the reducing end of grasses.

(E) The glycosyl sequence ARK-4 at the reducing end of grasses. This structure was found to be a minor component of the cell walls.

The xylitol of the ARK sequences was formed because the free reducing ends from the grasses were labeled with 2-AB by a secondary amination reaction. The glycosyl sequences were isolated from endo-xylanase digests by size exclusion chromatography and solid phase extraction.

Several eudicot and gymnosperm GXs have been shown to have a characteristic glycosyl sequence 4-β-D-Xylp-(1,4)-β-D-Xylp-(1,3)-α-L-Rhap-(1,2)-α-D-GalpA-(1,4)-D-Xylp (i.e., sequence **1**) at the reducing end (Johansson and Samuelson 1977, Peña *et al.* 2007). Studies with Arabidopsis mutants led to the proposal that this glycosyl sequence regulates xylan chain length either by acting as a primer or terminator of backbone elongation (Peña, et al. 2007, York and O'Neill 2008). Several putative glycosyltransferases, including FRA8, IRX8, and PARVUS have been implicated in the biosynthesis of sequence **1**. Nevertheless, the function of sequence **1** has not been determined (Peña, et al. 2007, Scheller and Ulvskov 2010, York and O'Neill 2008).

Structural comparison of arabinoxylans isolated from the cell walls of switchgrass, rice, brachypodium, miscanthus and foxtail millet led us to conclude that these polysaccharides lack discernible amounts of sequence **1** at the reducing end (Kulkarni, et al. 2012a). However, the actual glycosyl sequence(s) at the reducing end of grass cell wall AXs have not been reported nor are there published data describing the sequence of glycosyl residues at the reducing end of heteroxylans present in the walls of non-grass monocots.

Here, we report the structural characterization of the heteroxylans isolated from plants from 10 of the 12 monocot orders. Our data support the notion that the AX is the predominant xylan in the walls of the commelinids, whereas GX is the predominant xylan in the walls of non-commelinds. The walls of the Arecales, the basal commelinid, likely contain both AX and GX. The heteroxylans of the Poales and Asparagales contains no discernible amounts of sequence **1**, rather most of the reducing end residue is a xylose substituted at *O*-2 with 4-*O*-Me-Glc*p*A or at O-3 with Ara*f*. We refer to these sequences as Arabinoxylan Reducing end K-mers (ARK)-1, ARK-2, ARK-3 and ARK-4. By contrast, the heteroxylans synthesized by the other orders of

monocots have reducing end sequences comparable to those typically found in eudicots and gymnosperms.

Results

Structural Diversity in the Glycosyl Sequences at the Reducing End of Heteroxylans Isolated from the Cell Walls of Monocotyledonous Plants

To obtain insight into the structures of the heteroxylans in the cell wall of monocots, the depectinated alcohol insoluble residues (AIR) of plants representing 10 of the 12 orders of monocots (See Materials and Methods) were extracted with 1 M KOH to solubilize material enriched in heteroxylans (Kulkarni, et al. 2012a, Kulkarni *et al.* 2012b, Mazumder and York 2010, Peña, et al. 2007). The 1 M KOH soluble polysaccharides were then fragmented with endo-xylanase and the resulting xylo-oligosaccharides were separated by size exclusion chromatography. These xylo-oligosaccharides were then structurally characterized using matrix assisted laser-desorption time-of-flight mass spectrometry (MALDI-TOF-MS) and by 1-D and 2-D ^{1}H nuclear magnetic resonance (NMR) spectroscopy (**Figure 4.2 and 4.3**).

The most abundant xylo-oligosaccharides generated from monocots have masses between 700 Da and 1300 Da (**Figure 4.2**). The MALDI-TOF-MS showed the presence of three ion series, each consisting of $[M+Na]^+$ and $[M-H+2Na]^+$ ions. The most abundant series consisted of monoisotopic $[M+Na]^+$ ions at *m/z* 759, 891, and 1023 and monoisotopic $[M-H+2Na]^+$ ions at *m/z* 781, 913 and 1045. These correspond to oligosaccharides with 4 to 6 pentosyl residues substituted with a 4-O-MeGlcA residue as well as to sequence **1 -** β-D-Xylp-(1,4)-[β-D-Xylp-(1,4)]$_{0\text{-}2}$-β-D-Xylp-(1,3)-α-L-Rhap-(1,2)-α-D-GalpA-(1,4)-D-Xyl - which is present at the reducing end of glucuronoxylans from dicots (Peña, et al. 2007). Thus, these MALDI-TOF mass spectra alone do not provide sufficient information to distinguish between reducing end

sequences and acidic xylo-oligosaccharides substituted with MeGlc*p*A. In the MALDI-TOF spectra of xylan fragments from commelinid species, an additional series of monoisotopic $[M+Na]^+$ ions consistent with the presence of neutral arabinoxylan fragments (i.e., *m/z* 701, 833, 965 and 1097) was present along with ions corresponding to the acidic oligosaccharides.

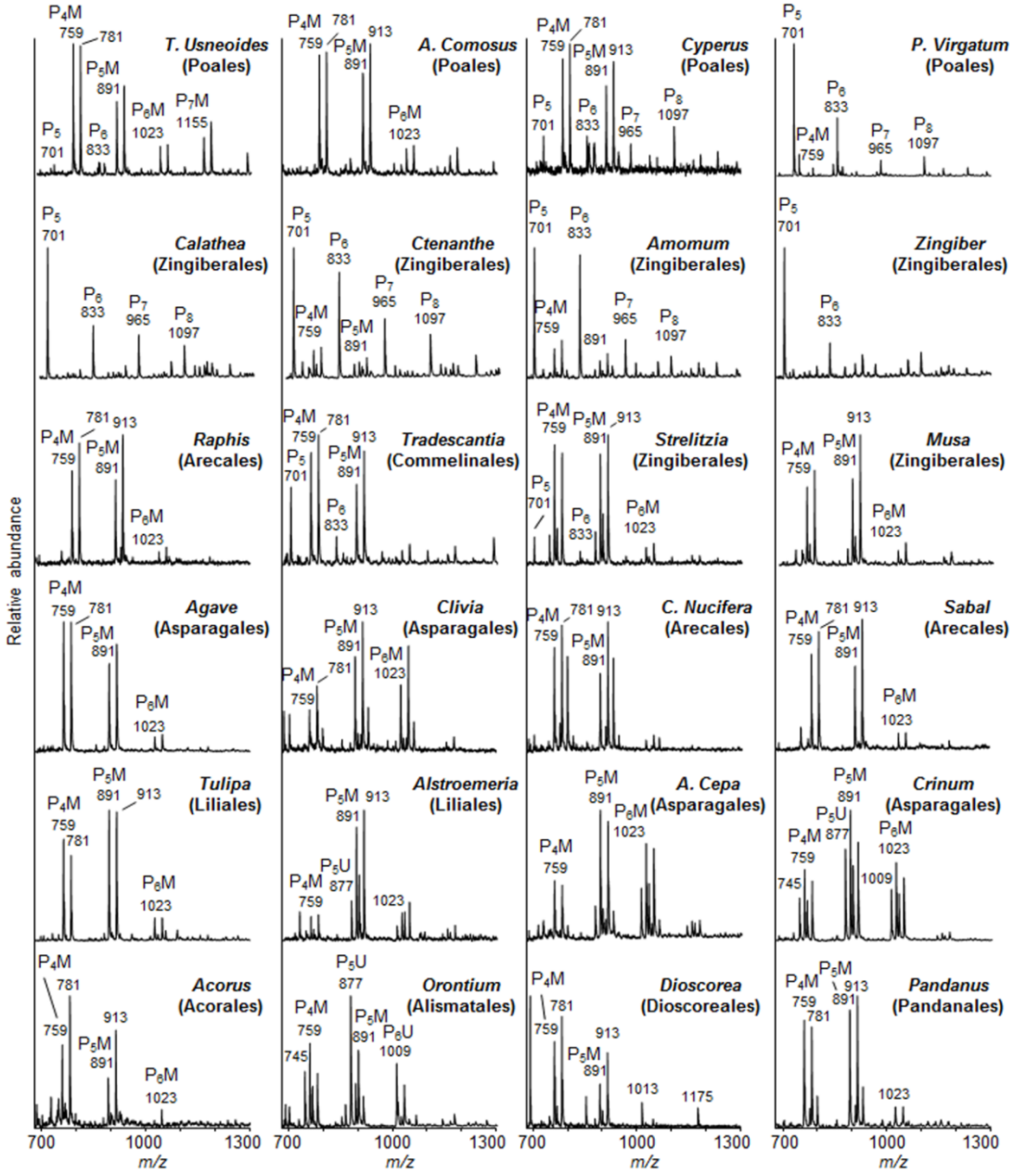


Figure 4.2: MALDI-TOF spectrum of the isolated xylo-oligosaccharides generated by xylanase treatment of the 1 M KOH extracts of AIR from monocot plant species. Ions are labeled with the most likely ([M + Na]+, [M + 2Na]+) assignments, but ions with more than two sodium atoms are also possible.

To obtain additional information on the structures of the monocot heteroxylans, we used NMR spectroscopy to characterize the xylo-oligosaccharides generated from selected plants. Signals characteristic of 4-*O*-Me-Glc*p*A and Glc*p*A (**Figure 4.3**) in varying proportions were present in the NMR spectra of all the orders analyzed. GX was present predominantly in the non-commelinid monocots, whereas GAX was present in the commelinids. Indeed, signals for 3-linked-Ara*f* were present in all the commelinid heteroxylans analyzed. However, signals corresponding to 2-linked-Ara*f* were only observed in the AX from members of the Poales. Thus 2-linked-Ara*f* is a characteristic of Poales AX, which suggests that the gene(s) encoding the glycosyltransferases responsible for synthesizing side chains containing this linkage co-evolved with the Poales (**Figure 4.4**). Notably, the abundance of MeGlc*p*A and Glc*p*A was highest and the abundance of Ara*f* was lowest in *Coccus nucifera*, a member of the Arecales. Thus, Arecales produces xylans that are the most similar to non-commelinid xylans, consistent with the assignment of Arecales as a basal commelinid that may represent a transitional stage between commelinids and non-commelinids. All but two of the orders of monocots had signals characteristic for the Gal*p*A and Rha*p* residue that are present in Sequence **1**. However, no signals for Sequence **1** were observed in the NMR spectra of the xylo-oligosaccharides generated from the heteroxylans of the Poales and Asparagales (**Figure 4.3 and 4.4**). The NMR spectra of the xylo-oligosaccharides generated from some of the plants, including crinum, onion, dioscorea and golden club, included a signal at δ 5.41, (**Figure 4.3**), previously attributed to H1 of 4-*O*-Me-Glc*p*A that is itself substituted with a galactosyl residue (Reis *et al.* 2005, Shatalov *et al.* 1999).

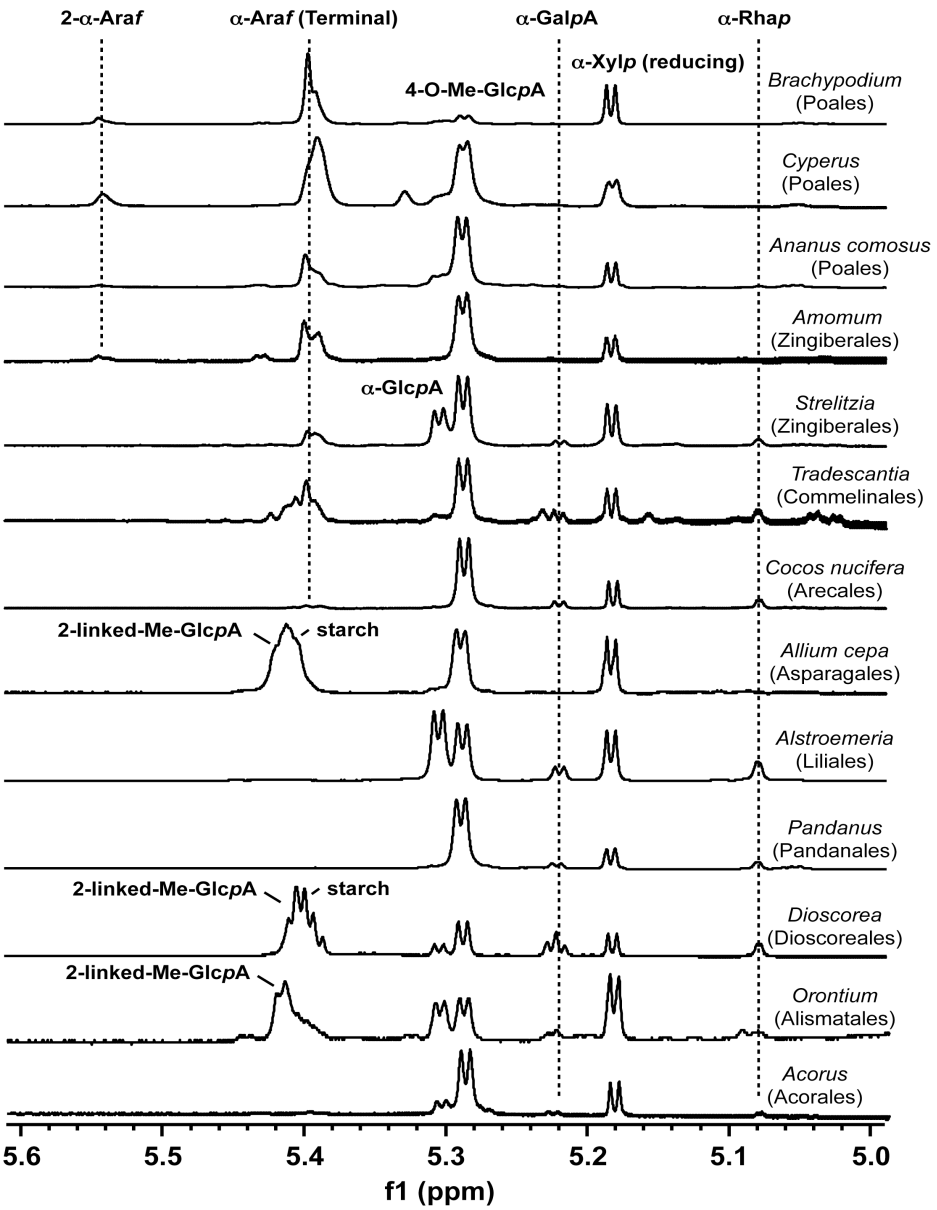
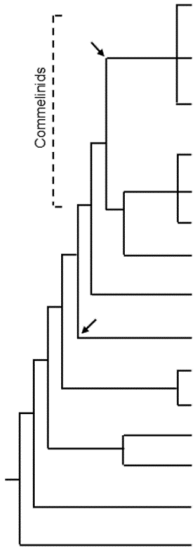


Figure 4.3: Trends in the anomeric region of the 600-MHz ^{1}H-NMR spectra of xylo-oligosaccharides generated by endo-xylanase treatment of 1 M KOH soluble extract from the monocots. H1 resonances of reducing xylose, MeGlc*p*A, Glc*p*A, 3-linked Ara*f*, and 2-linked Ara*f* residues are shown. H1 resonances of α-D-Gal*p*A and α-L-Rha*p* residues of sequence 1 are also shown. The resonances were consistent with previously published literature (Peña et al. 2007, Kulkarni et al. 2012a, Kulkarni et al. 2012b).



Family	Order	Plant names	Type of xylan	Sequence 1	Side chain			
					4-MeGlc*p*A	Glc*p*A	3-linked-Ara*f*	2-linked-Ara*f*
Poaceae	Poales	*Brachypodium*, *Panicum virgatum* (switchgrass), *Oryza* (rice), *Miscanthus*, *Setaria italica* (foxtail millet)	GAX	**ND**	+	**minor**	+	+
Cyperaceae		*Cyperus*	GAX	**ND**	+	**minor**	+	+
Bromeliaceae		*Ananas comosus* (pineapple), *Tillandsia usneoides* (spanish moss)	GAX	**ND**	+	**minor**	+	**minor**
Zingiberaceae	Zingiberales	*Zingiber* (Ginger), *Amomum* (cardamom), *Hedychium*	GAX	+	+	+	+	**ND**
Musaceae		*Musa*, banana	GAX	+	+	+	+	**ND**
Strelitziaceae		*Strelitzia*	GAX	+	+	+	+	**ND**
Commelinaceae	Commelinales	*Tradescantia*	GAX	+	+	**minor**	+	**ND**
Arecaceae	Arecales	*Cocos nucifera* (coconut), *Sabal*, *Howea*	GAX/GX	+	+	**minor**	**minor**	**ND**
Amaryllidaceae	Asparagales	*Allium cepa* (onion), *Crinum*, *Clivia*, *Agapanthus*	GX	**ND**	+	+	**ND**	**ND**
Alstroemeriaceae	Liliales	*Alstroemeria* (lily)	GX	+	+	+	**ND**	**ND**
Liliaceae		*Tulipa* (tulip)	GX	+	+	**minor**	**ND**	**ND**
Pandanaceae	Pandanales	*Pandanus*	GX	+	+	**ND**	**ND**	**ND**
Dioscoreaceae	Dioscoreales	*Dioscorea*	GX	+	+	**minor**	**ND**	**ND**
Araceae	Alismatales	*Orontium* (golden club), *Elodea*	GX	+	+	+	**ND**	**ND**
Acoraceae	Acorales	*Acorus*	GX	+	+	+	**ND**	**ND**

Figure 4.4: Summary of the trends in the xylan structures along with the phylogenetic tree of monocots (adapted from Chase *et al*. 2006). A complete list of the plant species used in the analyses is also mentioned. Arrows indicate the point in evolution where sequence **1** was lost.

The Reducing end Xylose of Poales and Asparagales heteroxylans is substituted at O-2 with 4-O-Methyl Glucuronic Acid or at O-3 with Araf

We previously provided evidence that the arabinoxylans isolated from different grass cell walls lack discernible amounts of sequence **1** at their reducing end (Kulkarni, et al. 2012a). Here, we show that the glucuronoxylan isolated from the cell walls of *A. cepa*, *Crinum*, *Agapanthus*, and *Asparagus* from Asparagales order also lacks sequence **1** at the reducing end (**Figure 4.3**). Our chemical evidence shows that in grass and Asparagales heteroxylans most of the reducing end xylose is substituted at *O*-2 with 4-*O*-MeGlcA and/or at *O*-3 with Ara*f* (**Figure 4.5, 4.6, 4.7**).

To identify the glycosyl sequence at the reducing end of grass arabinoxylans, the cell walls from switchgrass, brachypodium, foxtail millet, miscanthus, rice, and crinum were suspended in DMSO-HOAC (70:30 v/v) and reacted with 2-aminobenzamide (2-AB) in the presence of sodium cyanoborohydride. This reductive amination procedure converts each exposed reducing end glycose residue to a 2-AB-labeled derivative (**Supplemental Figure. S4.1**). The labeled cell walls were then sequentially extracted with ammonium oxalate and 1 M KOH (containing 1% v/v $NaBH_4$). The 1 M KOH-soluble extracts, which contained most of the xylan, were fragmented with an endo-xylanase to generate xylo-oligosaccharides. Some of the xylo-oligosaccharides had 2-AB at the reducing end, and these were enriched by a combination of size exclusion chromatography and Solid Phase Extraction (SPE) using graphitized carbon. These 2-AB labeled oligosaccharides were then structurally characterized using MALDI-TOF-MS, ESI-MS^n, and 1D and 2D NMR spectroscopy.

Extensive 1D and 2D (TOCSY, HSQC, ROESY, HMBC) NMR spectroscopy experiments were performed on the 2-AB labeled xylo-oligosaccharides generated from foxtail

millet (*Setaria italica*) to obtain complete chemical shift and scalar coupling information for each glycosyl residue (**Figure 4.5A and 4.5B, Table 4.1**). Our values are consistent with previously published data for 2AB-labled xylo-oligosaccharides (Ishii *et al.* 2008). The ^{1}H and the ^{13}C resonances of the terminal non-reducing β-D xylosyl, internal 4-linked β-D xylosyl, internal 2,4-linked β-D xylosyl, 3,4-linked β-D-xylosyl and reducing xylosyl residues were fully assigned. The resonances of terminal 4-*O*-MeGlc*p*A residues linked to *O*-2 of backbone xylosyl residues and to the xylitol-2AB were also identified (**Figure 4.5A, Table 4.1**). ROESY and HMBC experiments were performed to determine the point of attachment of MeGlc*p*A to the reductively aminated xylitol, which was labeled due to its location at the reducing end of the intact polysaccharide. Dipolar interactions between H-1 of one of the terminal 4-O-Me-Glc*p*A residues and H-2 of this aminated xylitol were observed in the ROESY spectrum of the 2-AB labeled xylo-oligosaccharides (**Figure 4.6, Table 4.2**). ROE cross peaks were also observed between H-1 of β-D-Xylp and H-4 of the aminated xylitol. Together these data provide strong evidence that the 4-linked xylose at the reducing end of foxtail millet arabinoxylan is substituted at *O*-2 with MeGlc*p*A.

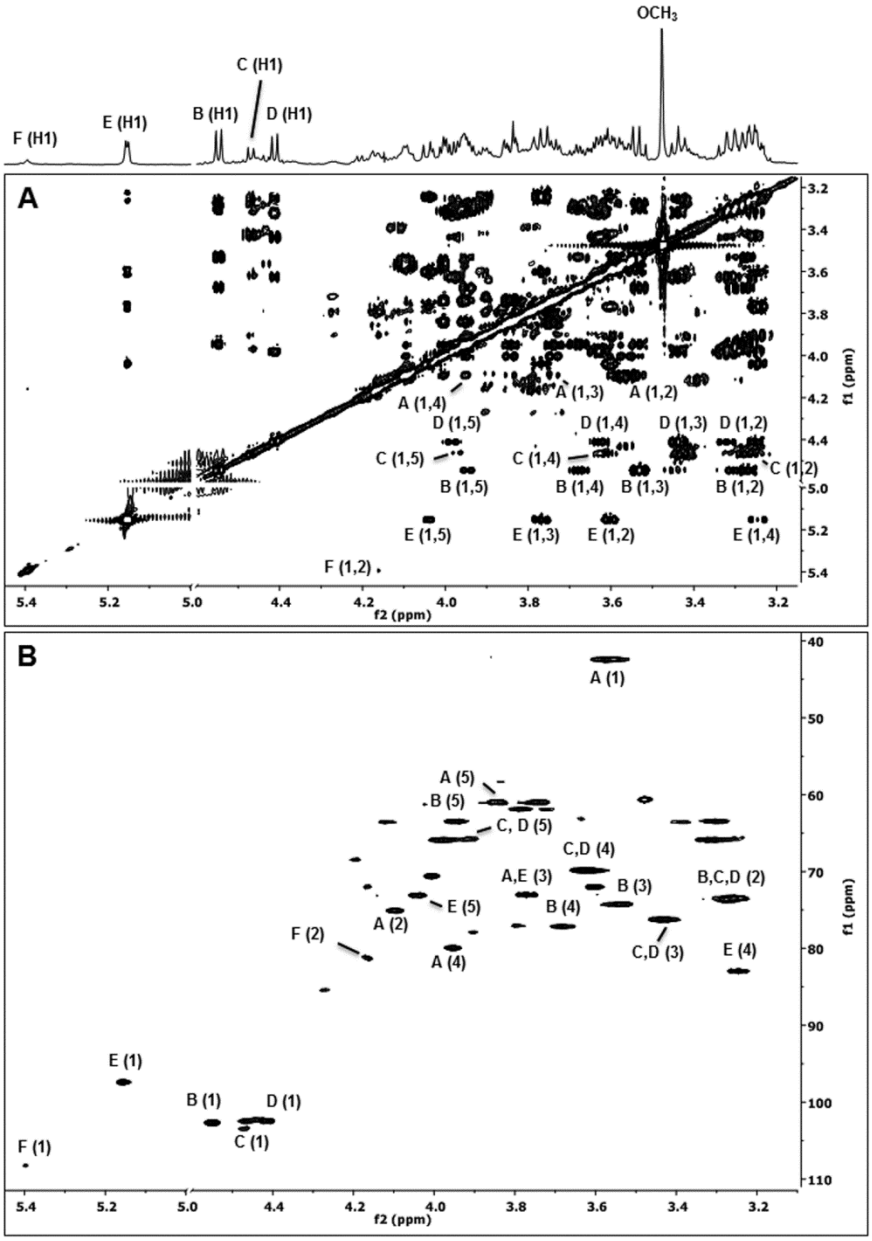


Figure 4.5: Partial 600-MHz NMR spectra, (A) TOCSY and (B) HSQC, of purified 2-AB xylo-oligosaccharides generated by β-endoxylanase digestion of the 1 M KOH extracts of AIR from *S. italica*. Crosspeak assignments of protons and carbons of each glycosyl residues are indicated using an uppercase letter (Table 4.1) and numbers indicating the position of the protons and carbons in the residue.

Table 4.1. ^{1}H and ^{13}C NMR assignments of the 2-AB-labeled xylo-oligosaccharides obtained by endoxylanase treatment of the 1 M KOH extract from the five grasses

Key	Residue	H1/C1	H2/C2	H3/C3	H4/C4	H5/C5
A	D-Xylitol [Xyl-2AB]	3.56/42.54	4.09/75.11	3.75/73.03	3.95/79.98	3.84/61.00
B	β-1,4-Xyl*p* (internal)	4.55/102.59	3.28/73.55	3.53/74.27	3.68/77.19	3.95/63.47
C	β-1,3,4-Xylp (α-L-Ara*f*)	4.46/103.40	3.27/73.59	3.42/76.26	3.61/69.83	3.96/65.87
D	β-Xyl*p* (terminal)	4.41/102.39	3.25/73.55	3.44/76.22	3.63/69.86	3.98/65.92
E	α-4Me-Glc*p*A	5.15/97.47	3.60/73.05	3.77/73.04	3.25/82.91	4.04/73.13
F	α-L-Ara*f* (terminal)	5.40/108.24	4.16/81.37			

Chemical shifts are reported in ppm relative to internal acetone, δ_H 2.225 and δ_C 30.89. β-Xyl*p* (α-L-Ara*f*) is a β-linked xylosyl residue that bears a α-L-Ara*f* sidechain at *O*-3. Signals for Xylitol are actually for xylitol that has 2-AB attached to it. See supplemental figure S1 for the reaction.

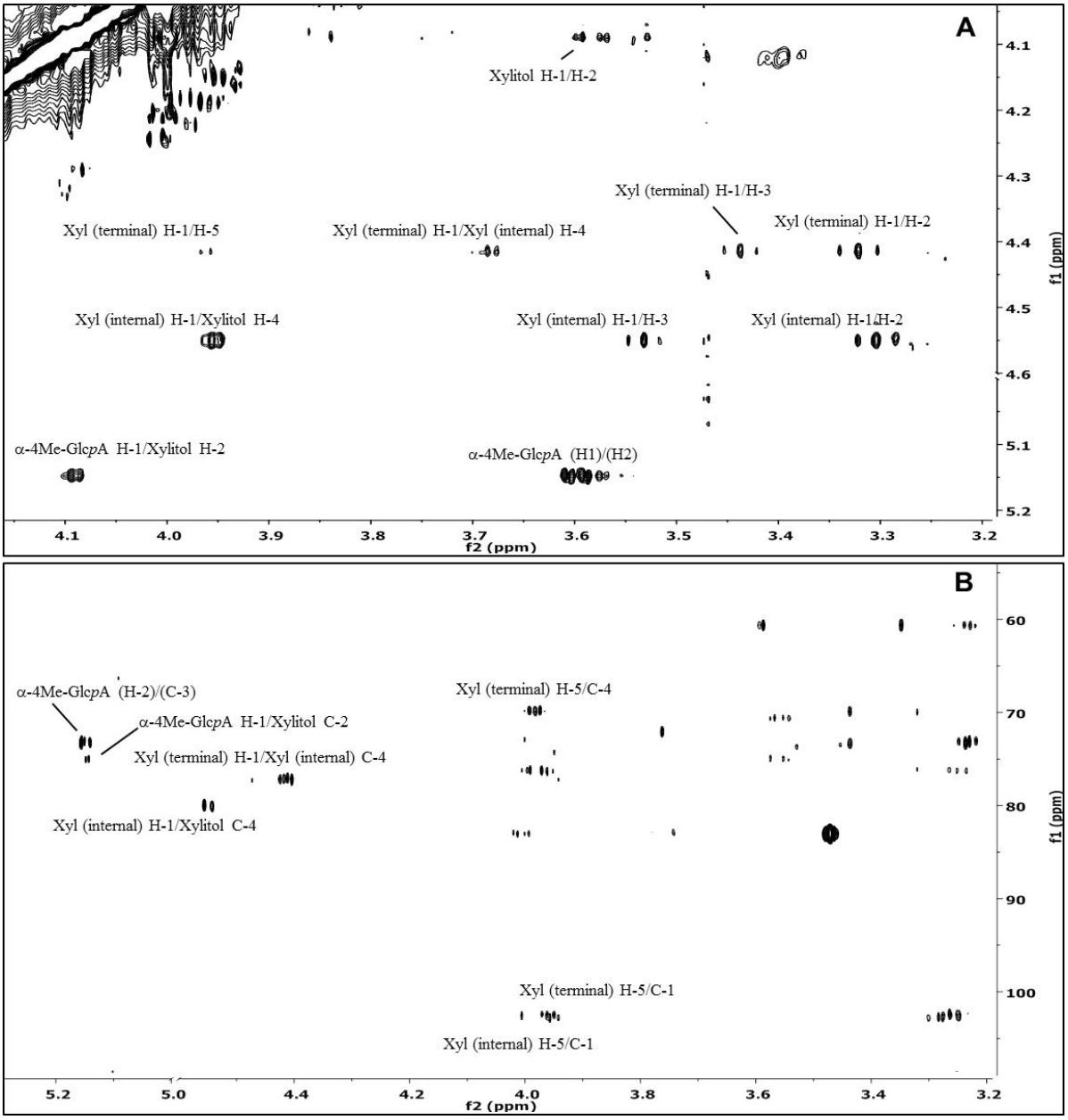


Figure 6: Partial 600-MHz NMR spectrum, (A) ROESY and (B) HMBC, of purified 2-AB xylo-oligosaccharides generated by β-endoxylanase digestion of the 1 M KOH extracts of AIR from *S. italica*. Crosspeak assignments of protons and carbons of each glycosyl residues are indicated in table 4.2.

Table 4.2. ROE connectivities of the sugar residues in ARK sequences

Residue	δH	Residue	δH	ROE connectivity
Xylitol H-1	3.58	Xylitol H-2	4.09	1,2 intra-ring correlation
Xyl (internal) H-1	4.55	Xylitol H-4	3.95	β-D-Xyl*p*-(1,4)-β-D-Xyl*p* linkage
Xyl (internal) H-1	4.55	Xyl (internal) H-5	3.95	1,5 intra-ring correlation
Xyl (internal) H-1	4.55	Xyl (internal) H-3	3.53	1,3 intra-ring correlation
Xyl (internal) H-1	4.55	Xyl (internal) H-2	3.29	1,2 intra-ring correlation
Xyl (internal) H-1	4.41	Xyl (terminal) H-4	3.68	β-D-Xyl*p*-(1,4)-β-D-Xyl*p* linkage
Xyl (terminal) H-1	4.41	Xyl (terminal) H-5	3.98	1,5 intra-ring correlation
Xyl (terminal) H-1	4.41	Xyl (terminal) H-3	3.44	1,3 intra-ring correlation
Xyl (terminal) H-1	4.41	Xyl (terminal) H-2	3.25	1,2 intra-ring correlation
α-4Me-Glc*p*A H-1	5.15	Xylitol H-2	4.09	α-4Me-Glc*p*A-(1,2)-β-D-Xyl*p* linkage
α-4Me-Glc*p*A H-1	5.15	α-4Me-Glc*p*A H-2	3.60	1,2 intra-ring correlation

The results described in the last paragraph did not provide a complete description of the distribution of MeGlc*p*A in the reducing end labeled xylo-oligosaccharides nor were they sufficient to locate the small amounts Ara*f* present. Thus, to more fully characterize these structures, the 2-AB labeled xylo-oligosaccharides were per-*O*-methylated and analyzed by ESI-MS^n. 2-AB labeled xylo-oligosaccharides were also isolated from switchgrass, miscanthus, brachypodium, rice and crinum and per-*O*-methylated and subjected to ESI-MS. The ESI-MS^n spectra of such oligosaccharides are dominated by Y and B ions, which provide robust sequence information (Domon and Costello 1988, Kulkarni, et al. 2012b, Mazumder and York 2010). Y ions correspond to fragments generated by loss of residues from the non-reducing end whereas B ions correspond to fragments generated by loss of residues from the reducing terminus (Domon and Costello 1988, Kulkarni, et al. 2012b).

The full MS of per-*O*-methylated 2-AB labeled xylo-oligosaccharides from foxtail millet (**Supplemental Figure S4.2**) and from the other grasses (Data not shown as all other grasses had similar spectra) contained a series of monoisotopic $[M+Na]^+$ ions at *m/z* 769, 929, 1089, 1249, which correspond to 2-AB labeled oligosaccharides with two to five pentosyl (xylose or arabinose) residues substituted with a single MeGlc*p*A (**Supplemental Figure S4.2**). Another fraction from foxtail millet was observed to be enriched in monoisotopic $[M+Na]^+$ ions at *m/z* 711, 871, 1031, and 1191. These ions were consistent with 2-AB oligosaccharides composed of three to six pentosyl (xylose or arabinose) residues. The precursor ions at *m/z* 929, 1089, and 1031 were selected for further fragmentation.

The quasimolecular ion at *m/z* 929 corresponds to a methylated 2-AB-labeled oligosaccharides composed of three pentosyl residues and one MeGlc*p*A. The ESI-MS^n spectra (**Figure 4.7, Table 4.3**) together with the results of NMR spectroscopy are consistent with the presence of

oligosaccharides with different substitution patterns. For example, the series of fragment ions at *m/z* 929-755-595/523 likely arise from Xyl-Xyl-(MeGlcA)-Xyl-2AB and/or (Ara)-Xyl-(MeGlcA)-Xyl-2AB (ARK-1). The high abundance of the diagnostic Y ion at *m/z* 595 in both the MS^2 and MS^3 spectra confirms that 4-O-MeGlc*p*A is attached to the 2-AB-labeled xylitol. The relatively low abundance of the fragment ions at *m/z* 581 and 363 indicated that a small portion of the MeGlc*p*A is itself substituted with a pentosyl residue. Thus the fragmentation pathway 929-755-581-363 likely arise from Xyl-(Ara-MeGlcA)-Xyl-2AB (ARK-2).

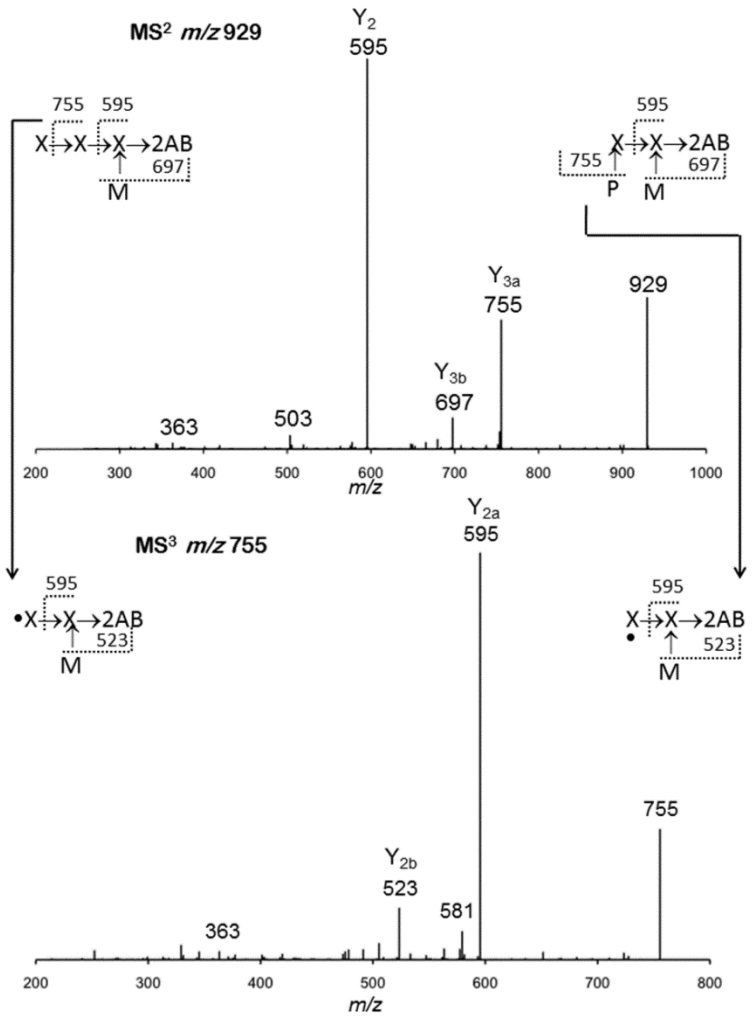


Figure 4.7: Precursor ion at *m/z* 929 was selected and subjected to ESI-MS^n. Xylo-oligosaccharides were generated by endo-xylanase treatment of 1 M KOH soluble extract of *S. italica* 2-AB labeled AIR and per-*O*-methylated. Same structures were found to be present in *Miscanthus*, *P. virgatum*, *Oryza*, and *Brachypodium*. Fragmentation events leading to the generation of Y-ions are shown in each spectrum. The fragmentation pathway (929 – 755 – 595/523) was seen to be consistent with the structures Xyl-Xyl-(MeGlcA)-Xyl-2AB and (Pentose)-Xyl-(MeGlcA)-Xyl-2AB.

The quasimolecular ion at *m/z* 1089 corresponding to a methylated 2-AB-labeled oligosaccharide composed of four pentosyl residues and one Glc*p*A. ESI-MS^n indicated that at least three oligosaccharides with different substitution patterns were present (**Supplemental Figure 4.3, Table 4.3**). For instance, the series of fragment ions at *m/z* 1089-915-755-595/523 is consistent with the sequence Xyl-Xyl-Xyl-(MeGlcA)-Xyl-2AB (ARK-1), whereas the series of fragment ions at m/z 1089-915-741-595 provides evidence that sequence Xyl-(Ara)-Xyl-(MeGlcA)-Xyl-2AB is present. The high abundance of diagnostic Y ion at *m/z* 595 in MS^2, MS^3, and MS^4 spectra again provides strong evidence that 4-O-MeGlcpA is linked to the formerly reducing end xylosyl. Low abundance of ions at *m/z* 683 and 581 showed the presence of small amounts of oligosaccharides with a substituted MeGlcA at the reducing end. Thus the fragmentation pathway 1089-915-755/683-595/581 arise from Xyl-Xyl-(Ara-MeGlcA)-Xyl-2AB (ARK-2). A detailed summary of fragmentation pathways leading to other possible structures is provided in Table 4.3.

The quasimolecular ion at *m/z* 1031 corresponds to a methylated 2-AB-labeled oligosaccharide composed of five pentosyl residues. ESI-MS^n indicated that both a linear and several branched sequences are present (**Supplemental Figure S4.4**). The high abundance of fragment ions at *m/z* 683, 523, and 363 in MS^2, MS^3, and MS^4 and the fragment ion at *m/z* 857 in the MS^3 spectra provides evidence that a pentosyl (arabinosyl) residue is linked to the xylose backbone. The major fragment ion at *m/z* 523 in the MS^4 spectrum of *m/z* 697, indicate that either one of xylose residue at the reducing end or the adjacent xylosyl residue is branched. However, the major fragment ion at *m/z* 363 in the MS^5 spectrum of *m/z* 537 establishes that a pentosyl residue is also linked to the xylose residue at the reducing end Xyl-Xyl-Xyl-(Pentose)-Xyl-2AB (ARK-3). The fragmentation pathway at *m/z* (1031-857-697-377) arises from a linear

oligosaccharide with structure Xyl-Xyl-Xyl-Xyl-Xyl-2AB, thus indicating the presence of less abundant oligosaccharides in which the reducing end xylosyl residue is not branched.

A combination of ESI-MS and NMR spectroscopy has provided evidence that the heteroxylans isolated for five different grasses contain a branched xylose residue at their reducing ends. It was observed that the branched xylose was always substituted either with a MeGlc*p*A or with an Ara*f* residue. Small amounts of unbranched reducing end xylose residues were also found in the grasses. However, the possibility of other glycosyl sequences being present at the reducing ends cannot be neglected.

No discernible amounts of sequence 1 were detected by NMR analyses of the xylo-oligosaccharides generated from selected members of the Asparagales. To determine the glycosyl sequence at the reducing end of these heteroxylans, the AIR from *Crinum* was reacted with 2-AB and the heteroxylan then extracted with 1 M KOH (containing 1% v/v $NaBH_4$). Xylo-oligosaccharides were generated by endo-xylanase treatment of the 1 M KOH-soluble material and enriched using size-exclusion chromatography. The 2-AB labeled xylo-oligosaccharides were per-*O*-methylated and subjected to ESI-MS^n. Fragmentation of the precursor ion at *m/z* 1089 confirmed that ARK-1 and ARK-2 sequences were present in the crinum also (**Supplemental Figure S4.5**). The precursor ion at *m/z* 1031 was also subjected to ESI-MS^n analysis (data not shown) confirming the presence of ARK-3, and ARK-4 sequences at the reducing ends of crinum. This data suggest that the ARK sequences may have an important role in xylan synthesis in plants that lack sequence 1.

Table 4.3. Fragmentation pathways from ESI-MSn of the precursor ions leading to the sequence information of the structures present in grasses

Precursor ion	Fragmentation pathway	Structures
929	929→755→595	Xyl→Xyl→Xyl-2AB ↑ MeGlcA
	929→755→595/523	Xyl → Xyl-2AB ↑ ↑ Pentose MeGlcA
	929→755→581	Xyl→Xyl-2AB ↑ MeGlcA ↑ Pentose
1089	1089→915→755→595/523	Xyl→Xyl→Xyl→Xyl-2AB ↑ MeGlcA
	1089→915→741→595	Xyl→Xyl→Xyl-2AB ↑ ↑ Pentose MeGlcA
	1089→915→755→581	Xyl→Xyl→Xyl-2AB ↑ MeGlcA ↑ Pentose
1031	1031→857→697→377	Xyl→Xyl→Xyl→Xyl→Xyl-2AB
	1031→857→697→537→377	Xyl→Xyl→Xyl→Xyl-2AB ↑ Pentose
	1031→857→683→537→377	Xyl→Xyl→Xyl→Xyl-2AB ↑ Pentose
	1031→857→697→537→363	Xyl→Xyl→Xyl→Xyl-2AB ↑ Pentose
	1031→857→683→537/509→363	Xyl→Xyl→Xyl-2AB ↑ ↑ Pentose Pentose
	1031→857→683→523→363	Xyl→Xyl→Xyl-2AB ↑ Pentose ↑ Pentose

The arrows indicate the point of generation of a scar during ESI-MS. Pentose indicate the presence of either a Araf or a Xylp side chain since they have the same mass in ESI-MS.

Discussion

Our data provide compelling evidence for the presence of the glycosyl sequences, ARK-1 to ARK-4, at the reducing ends of the heteroxylans produced by Poales and Asparagales, indicating that the ARK sequences are characteristic of species that do not produce detectable amounts of **1**. Because of their conservation in these species, we hypothesize that these reducing-end sequences play an important role in the biosynthesis of xylan in some orders of monocots.

Structural Analyses of Xylans from Diverse Monocot Species Reveal Trends that can be correlated to Taxonomy

Most studies of monocot cell walls have emphasized the grasses. Other than Poaceae, very little information is available about xylan structure in the cell walls of monocots. Studies with cyperus, pineapple, flax, palms showed that the walls of these plants contain glucuronoarabinoxylan (Bendahou *et al.* 2007, Buchala and Meier 1972, Jacobs, et al. 2003, Scheller and Ulvskov 2010, Simas-Tosin, et al. 2013). A comparison of xyloglucan and pectins from the cell walls of monocotyledons provided evidence for structural diversity across the taxonomy of monocots (Jarvis et al. 1988, Hsieh and Harris 2009). However there are no reports on the complete characterization of xylan across different orders of monocots. Structural analysis of xylo-oligosaccharides prepared from different orders of monocot species showed several distinct trends (**Figure 4.3 and 4.4**). It was previously proposed that changes in cell wall polysaccharide structures can provide insight into the events that occurred during the evolution of land plants (Hörnblad *et al.* 2013, Kulkarni, et al. 2012b). Our analyses of structural diversity in heteroxylan structure in diverse monocot species support this notion.

Plants forms two types of cell walls, primary and secondary cell walls that differ in function and composition. Primary walls surround growing and expanding cells, whereas secondary cell walls are deposited once the cell has ceased to grow. Primary cell walls in flowering plants are broadly divided into two categories, namely, type I and type II cell walls (Carpita 1996, Carpita and Gibeaut 1993, Pauly and Keegstra 2008, Vogel 2008). Type I primary cell walls are common features of dicots and noncommelinid monocots (e.g. acorales, asparagales, and arecales) with xyloglucan (XyG) being the major hemicellulosic component in the cell walls. All dicots and non-commelinid monocots that we have examined deposit GX into their secondary cell walls, suggesting this type of heteroxylan is conserved in plants that produce type I cell walls. On the contrary, type II cell walls, which are present in commelinids (e.g. poales, commelinales, zingiberales), contain glucuronoarabinoxylan (GAX) as the major hemicellulose (Pauly and Keegstra 2008, Vogel 2008). Our data establishes that the GAX is present in diverse commelinids, thus being consistent with the published literature. Both GX and GAX were present in Arecales consistent with the idea that this basal order preserves structural features that coexisted at the time commelinids diverged from non-commelinids. 2-linked Ara*f* residues were present in the Poales order only indicating that the genes encoding enzymes that catalyze the attachment of 2-linked Ara*f* residues evolved when Poales diverged from other monocots. It was also observed that Sequence **1** was present at the reducing ends of heteroxylans produced by all other monocots orders with the exception of Poales and Asparagales. It is likely that sequence **1** was conserved during the evolution of non-commelinid monocotyledons from more basal angiosperm taxa.

ARK Sequences are Present in Monocot Species that Lack Sequence 1

Analysis of monocots species indicated the absence of sequence 1 in grasses and Asparagales orders of monocots. AIR from the grasses and crinum were reacted with 2-AB, thus attaching this fluorescent label to the free reducing ends of the native polysaccharide in situ. The labeled polysaccharides were then extracted with alkali and fragmented with an endoxylanase. Comprehensive complete structural characterization of the resulting 2-AB labeled xylo-oligosaccharides by NMR and mass spectroscopy revealed the presence of ARK sequences at their reducing ends. Poales and Asparagales orders possess ARK sequences at their reducing ends, even though these two orders do not comprise a paraphyletic group. One possible explanation is that parallel evolution of these species independently resulted in new mechanisms for xylan synthesis in which ARK sequences functionally replace sequence **1**.

Models for the synthesis of GAX in monocots

Previously published literature indicates that evolutionarily diverse plants, such as gymnosperms and angiosperms, possess glycosyl sequence **1** at the reducing ends of GXs, suggesting that they synthesize GX using similar mechanisms (Johansson and Samuelson 1977, Lee *et al.* 2011, Peña, et al. 2007, Shimizu *et al.* 1978, York and O'Neill 2008). In dicots, sequence **1** has been proposed to play an important role in xylan biosynthesis. Studies with Arabidopsis mutants suggest that this glycosyl sequence has a role in regulating xylan chain length by acting either as a primer or as a terminator (Peña, et al. 2007, York and O'Neill 2008). Two mechanisms for the xylan biosynthesis have been proposed. In the first, sequence **1** acts as a primer to which xylose residues are sequentially added at the non-reducing end, extending and

elongating the backbone. In the second, sequence **1** acts as a terminator that whose addition prevents the addition of xylose residues to the reducing end of the nascent chain, (Peña, et al. 2007, York and O'Neill 2008). Absence of sequence **1** in heteroxylans in grasses and Asparagales and presence of ARK sequences at the reducing end of these polysaccharides suggests that these plants have developed different mechanisms for controlling xylan chain length, and that the ARK sequences might play an important role in xylan synthesis when sequence **1** is absent. In grasses, ARK sequences might be acting as a primer to which xylose residues are sequentially added at the non-reducing end to generate the xylan backbone. Alternatively, generation of an ARK sequence at the reducing end may be a signal for termination backbone synthesis. This termination process could be initiated, for example, as a result of the action of a specific hydrolase (or a transglycosylase with some hydrolytic activity) that cleaves the glycosidic bond that connects a branched xylosyl residue in the nascent xylan to a carrier molecule. Our observations thus serve as the basis for testable models for evolution of xylan in monocots.

In summary, our results showed that grasses synthesize ARK sequences at their reducing ends. These sequences may play a key role in xylan biosynthesis by acting as either a primer or a terminator sequence. The presence of the ARK sequences in xylans from both Asparagales (crinum) and Poales, but their absence in other monocot species, suggests that these distinct structures may have arisen by parallel evolution and provides the basis for specific hypotheses as to how this might have occurred. Analysis of xylans in monocots thus establishes that changes in the structure of cell walls in land plants can be correlated to taxonomy, making it possible to develop models that predict specific types of structural diversity in xylans across the various

orders of monocots. These correlations can also provide insight into the mechanism of xylan synthesis in grasses, enabling technologies for altering the structure of xylans to develop and/or otherwise manipulate advanced biomass crops with economically desirable properties.

Materials and Methods

Plant Material

Stems and leaves of mature *Miscanthus x giganteus, B. distachyon*, and *S. italica* along with Acorales (Acorus), Alismatales (Elodea), Dioscoreales (Dioscorea), Pandanales (Pandanus), Asparagales (Agapanthus, Asparagus, Clivia, Crinum, Iris, Ludisia, Onion), Arecales (Caryoyta, Howea, Rapis, Sabal, Coconut), Commelinales (Tradescantia, Eichornia, Spiderwort), Zingiberales (Cardamon, Ginger, Hedychium, Strelitzia), Bromeliaceae (Pineapple, Tillandsia), and Cyperaceae (Cyperus) were obtained from the Plant Biology greenhouse at the University of Georgia (Athens, GA). Rice straw was obtained from Wayne Parrot at the University of Georgia. The air-dried stems and leaves were Wiley-milled (-20/+80 mesh) and stored at room temperature. Milled poplar and switchgrass were obtained from the National Renewable Energy Laboratory, Golden, CO. The switchgrass (*P. virgatum* cultivar Alamo) was obtained from the Samuel Roberts Noble Foundation (Ardmore Oklahoma). It was milled using a Hammer mill with a 1" screen and then ground in a Wiley mill using a 1mm screen. The milled material was sieved to -20/+80 mesh (Kulkarni et al. 2012a). *A. comosus, tulipa, alstroemeria* leaves and stems were obtained from Kroger, a retailer.

Preparation of Cell Walls

Cell walls from grasses were prepared as their alcohol-insoluble residue (AIR) (Mazumder and York 2010, Kulkarni et al. 2012a). Biomass from the five grasses were ball-milled for 16 h at 4 °C and 90-100 rpm in aq. 80% EtOH (v/v) using ¼ inch zirconium grinding media (U.S. Stoneware, East Palestine, OH). The insoluble residues were washed with acetone and dried under vacuum. AIR from other monocots species investigated were prepared by homogenizing suspensions of the biomass in 80% (v/v) EtOH with a Polytron tissue disruptor (Kinematica, Switzerland). The insoluble residues were then washed with 100% (v/v) EtOH, followed by MeOH:$CHCL_3$ (1:1) and acetone and then dried at room temperature under vaccum.

2-Aminobenzamide Labeling of the Reducing Glycoses Present in AIR

Suspensions of the AIR from the five grasses and from Crinum in DMSO: Acetic acid (70:30, v/v) containing sodium cyanoborohydride (1M) and 2-AB (0.5M) were kept overnight at 65°C (Bigge *et al.* 1995). The suspensions were then dialyzed against deionized water (3500 molecular weight cut off tubing) and lyophilized. Treating the AIR with 2-AB results in the labeling of the native reducing ends by reductive amination.

Sequential Extraction of AIR

AIR (0.5g) from monocots, and the 2-AB labeled grasses and Crinum was sequentially extracted with 50 mM ammonium oxalate (35 mL), and 1 M KOH containing 1% (w/v) $NaBH_4$. 1 M KOH soluble extracts were further neutralized and dialyzed (3500 molecular weight cut off tubing) against deionized water and lyophilized (Kulkarni et al. 2012a, Kulkarni et al. 2012b).

Generation of Xylo-Oligosaccharides by Endo-Xylanase Treatment of the 1 M KOH-Soluble Material

The 1 M KOH-soluble materials from the grasses and other monocots (~ 20 mg) were suspended in 50 mM ammonium formate, pH 5, and treated for 24 h at 37 °C with *endo*-xylanse (3.5 unit, *Trichoderma viride* M1, Megazyme, Wicklow, Ireland)(Kulkarni et al. 2012b). The insoluble residues were removed by centrifugation and EtOH then added to the supernatant to a final concentration of 60% (v/v). The mixture was kept for 24 h at 4 °C. The precipitate that formed was removed by centrifugation and the soluble material lyophilized. The xylo-oligosaccharides were isolated by size exclusion chromatography using Superdex 75 column and further emriched in 2-AB labeled xylo-oligosaccharides by Solid Phase Extraction (SPE) using graphitized carbon. 2-AB labeled xylo-oligosaccharides were eluted with a series of different concentrations of acetonitrile. Most of the 2AB labeled xylo-oligosaccharides were eluted with 40% (v/v) acetonitrile.

MALDI-TOF Mass Spectrometry

Positive ion MALDI-TOF mass spectra were recorded using a Bruker LT MALDI-TOF mass spectrometer interfaced to a Bruker biospectrometry workstation (Bruker Daltonics, http://www.bdal.com) as described in (Mazumder and York 2010, Kulkarni et al. 2012a).

^{1}H-NMR Spectroscopy

^{1}H NMR spectra of 2-AB labeled xylo-oligosaccharides from the five grasses and the xylo-oligosaccharides from monocot species (~1 mg in ~0.5 mL 99.9% D_2O) were recorded at 600

MHz using a Varian Inova NMR spectrometer (Agilent, Santa Clara, CA). 2D spectra were recorded using standard Varian pulse programs. Chemical shifts were measured relative to internal acetone at δ_H 2.225 and δ_C 30.89.

Per-O-Methylation of the Xylo-Oligosaccharides

Xylo-oligosaccharide-enriched materials (~1 mg) were dissolved in dry DMSO (0.2 mL) and per-*O*-methylated as described (Mazumder and York 2010).

ESI Mass Spectrometry

The per-*O*-methylated xylo-oligosaccharides were analyzed by positive ion ESI mass spectrometry using a Thermo Scientific LTQ XL mass spectrometer (Thermo Scientific, http://www.thermoscientific.com) as described (Mazumder and York 2010).

Acknowledgement

This research was funded by the Bio-Energy Science Center (BESC) administered by Oak Ridge National Laboratory and funded by a grant (DE-AC05-00OR22725) from the Office of Biological and Environmental Research, Office of Science, U.S. Department of Energy.

References

Ahlgren PA, Goring DAI. 1971. Removal of wood components during chlorite delignification of black spruce. *Can. J. Chem* 49:1272-1275.

Ahola V, Aittokallio T, Vihinen M, Uusipaikka E. 2006. A statistical score for assessing the quality of multiple sequence alignments. *BMC Bioinformatics.* 7:484.

Anders N, Wilkinson MD, Lovegrove A, Freeman J, Tryfona T, Pellny TK, Weimar T, Mortimer JC, Stott K, Baker JM*, et al.* 2012. Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proc. Natl. Acad. Sci. U S A* 109:989-993.

Aspeborg H, Schrader J, Coutinho PM, Stam M, Kallas A, Djerbi S, Nilsson P, Denman S, Amini B, Sterky F*, et al.* 2005. Carbohydrate-active enzymes involved in the secondary cell wall biogenesis in hybrid aspen. *Plant Physiol.* 137:983-997.

Atkins E, Parker K, Preston R. 1969. The Helical structure of the -1, 3-linked xylan in some siphoneous green algae. *Proc. Roy. Soc. B.* 173:209-221.

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

Bendahou A, Dufresne A, Kaddami H, Habibi Y. 2007. Isolation and structural characterization of hemicelluloses from palm of *Phoenix dactylifera L. Carbohydr. Polym.* 68:601-608.

Bigge JC, Patel TP, Bruce JA, Goulding PN, Charles SM, Parekh RB. 1995. Nonselective and Efficient Fluorescent Labeling of Glycans Using 2-Amino Benzamide and Anthranilic Acid. *Anal. Biochem.* 230:229-238.

Bremer B, Bremer K, Chase MW, Fay MF, Reveal JL, Soltis DE, Soltis PS, Stevens PF, Anderberg AA, Moore MJ*, et al.* 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* 161:105-121.

Bremer K. 2000. Early Cretaceous lineages of monocot flowering plants. *Proc. Natl. Acad. Sci. U S A* 97:4707-4711.

Bremner I, Wilkie K. 1966. The hemicelluloses of bracken:: Part I. An acidic xylan. *Carbohydr Res.* 2:24-34.

Bromley JR, Busse-Wicher M, Tryfona T, Mortimer JC, Zhang Z, Brown DM, Dupree P. 2013. GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *Plant J* 74:423-434.

Brown D, Wightman R, Zhang Z, Gomez LD, Atanassov I, Bukowski JP, Tryfona T, McQueen-Mason SJ, Dupree P, 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 J.* 66:401-413.

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

Brown DM, Zhang Z, Stephens E, Dupree P, Turner SR. 2009. Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *Plant J.* 57:732-746.

Buchala A, Meier H. 1972. Hemicelluloses from the stalk of *Cyperus papyrus*. *Phytochemistry* 11:3275-3278.

Carafa A, Duckett JG, Knox JP, Ligrone R. 2005. Distribution of cell-wall xylans in bryophytes and tracheophytes: new insights into basal interrelationships of land plants. *New Phytol.* 168:231-240.

Carnachan SM, Harris PJ. 2000. Polysaccharide compositions of primary cell walls of the palms *Phoenix canariensis* and *Rhopalostylis sapida*. *Plant Physiol. Bioch.* 38:699-708.

Carpita NC. 1996. Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47:445-476.

Carpita NC, Gibeaut DM. 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3:1-30.

Carpita NC, Kanabus J. 1987. Extraction of starch by dimethyl sulfoxide and quantitation by enzymatic assay. *Anal. Biochem.* 161:132-139.

Cartmell A, McKee LS, Peña MJ, Larsbrink J, Brumer H, Kaneko S, Ichinose H, Lewis RJ, Vikso-Nielsen A, Gilbert HJ*, et al.* 2011. The structure and function of an arabinan-specific alpha-1,2-arabinofuranosidase identified from screening the activities of bacterial GH43 glycoside hydrolases. *J. Biol. Chem.* 286:15483-15495.

Chase MW, Fay MF, Devey DS, Maurin O, R√∏nsted N, Davies TJ, Pillon Y, Petersen G, Seberg O, Tamura MN. 2006. Multigene analyses of monocot relationships: a summary. *Aliso* 22:63-75.

Chiniquy D, Sharma V, Schultink A, Baidoo EE, Rautengarten C, Cheng K, Carroll A, Ulvskov P, Harholt J, Keasling JD*, et al.* 2012. XAX1 from glycosyltransferase family 61 mediates xylosyltransfer to rice xylan. *Proc. Natl. Acad. Sci. U S A* 109:17117-17122.

Chiniquy D, Varanasi P, Oh T, Harholt J, Katnelson J, Singh S, Auer M, Simmons B, Adams PD, Scheller HV*, et al.* 2013. Three Novel Rice Genes Closely Related to the Arabidopsis IRX9, IRX9L, and IRX14 Genes and Their Roles in Xylan Biosynthesis. *Front Plant Sci.* 4:83.

Ciucanu I, Kerek F. 1984. A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res.* 131:209-217.

Correia MAS, Mazumder K, Bras JLA, Firbank SJ, Zhu Y, Lewis RJ, York WS, Fontes CMGA, Gilbert HJ. 2011. The structure and function of an arabinoxylan-specific xylanase. *J. Biol. Chem.* 286:22510-22520.

Dalessandro G, Northcote DH. 1977. Changes in enzymic activities of nucleoside diphosphate sugar interconversions during differentiation of cambium to xylem in pine and fir. *Biochem. J.* 162:281-288.

de O. Buanafina MM. 2009. Feruloylation in Grasses: Current and Future Perspectives. *Mol. Plant* 2:861-872.

Dhugga KS, Tiwari SC, Ray PM. 1997. A reversibly glycosylated polypeptide (RGP1) possibly involved in plant cell wall synthesis: purification, gene cloning, and trans-Golgi localization. *Proc Natl Acad Sci U S A* 94:7679-7684.

Doering A, Lathe R, Persson S. 2012. An update on xylan synthesis. *Mol Plant* 5:769-771.

Domon B, Costello CE. 1988. A Systematic Nomeclature for Carbohydrate Fragmentations in FAB-MS/MS Spectra of Glycoconjugates. *Glycoconj. J.* 5:397-409.

Domozych DS, Sorensen I, Willats WG. 2009. The distribution of cell wall polymers during antheridium development and spermatogenesis in the Charophycean green alga, Chara corallina. *Ann. Bot.* 104:1045-1056.

Doust AN, Kellogg EA, Devos KM, Bennetzen JL. 2009. Foxtail millet: a sequence-driven grass model system. *Plant Physiol.* 149:137-141.

Ebringerová A, Hromadkova Z, Heinze T. 2005. Hemicellulose. *Adv. Polym. Sci.* 186:1-67.

Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics.* 14:755-763.

Evert RF. 2006. Esau's plant anatomy: meristems, cells, and tissues of the plant body: their structure, function, and development Wiley-Liss.

Faik A. 2010. Xylan biosynthesis: News from the grass. *Plant Physiol.* 153:396-402.

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R*, et al.* 2006. Pfam: clans, web tools and services. *Nucleic Acids Res.* 34:D247-251.

Fu H, Yadav MP, Nothnagel EA. 2007. *Physcomitrella patens* arabinogalactan proteins contain abundant terminal 3-O-methyl-L: -rhamnosyl residues not found in angiosperms. *Planta.* 226:1511-1524.

Gibeaut DM, Pauly M, Bacic A, Fincher GB. 2005. Changes in cell wall polysaccharides in developing barley (*Hordeum vulgare*) coleoptiles. *Planta* 221:729-738.

Givnish T, Sytsma K. 2006. Phylogeny of the monocotyledons based on the highly informative plastid gene ndhF: Evidence for widespread concerted convergence. *Monocots: Comparative Biology and Evolution (Excluding Poales). Rancho Santa Ana Botanic Garden, Claremont, California*:28-51.

Givnish TJ, Ames M, McNeal JR, McKain MR, Steele PR, dePamphilis CW, Graham SW, Pires JC, Stevenson DW, Zomlefer WB. 2010. Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of Poales 1. *Annals of the Missouri Botanical Garden* 97:584-616.

Graham L, Cook M, Busse J. 2000. The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proc. Natl. Acad. Sci. USA* 97:4535.

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52:696-704.

Heaton EA, Dohleman FG, Long SP. 2008. Meeting US biofuel goals with less land: the potential of Miscanthus. *Global Change Biol.* 14:2000-2014.

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315:804-807.

Hiwatashi Y, Nishiyama T, Fujita T, Hasebe M. 2001. Establishment of gene-trap and enhancer-trap systems in the moss *Physcomitrella patens*. *Plant J.* 28:105-116.

Hörnblad E, Ulfstedt M, Ronne H, Marchant A. 2013. Partial functional conservation of IRX10 homologs in *physcomitrella patens* and *Arabidopsis thaliana* indicates an evolutionary step contributing to vascular formation in land plants. *BMC Plant Biol.* 13:3.

Hsieh YS, Harris PJ. 2009. Xyloglucans of monocotyledons have diverse structures. *Molecular Plant* 2:943-965.

Ishii T, Konishi T, Ono H, Ohnishi-Kameyama M, Togashi H, Shimizu K. 2008. Assignment of the ^{1}H and ^{13}C NMR spectra of 2-aminobenzamide-labeled xylo-oligosaccharides. *Carbohydr. Polym.* 74:579-589.

Izydorczyk M, Biliaderis C. 1995. Cereal arabinoxylans: advances in structure and physicochemical properties. *Carbohydr. Polym.* 28:33-48.

Jacobs A, Palm M, Zacchi G, Dahlman O. 2003a. Isolation and characterization of water-soluble hemicelluloses from flax shive. *Carbohydr. Res.* 338:1869-1876.

Jacobs AK, Lipka V, Burton RA, Panstruga R, Strizhov N, Schulze-Lefert P, Fincher GB. 2003b. An Arabidopsis callose synthase, GSL5, is required for wound and papillary callose formation. *The Plant Cell Online* 15:2503-2513.

Jarvis MC, Forsyth W, Duncan HJ. 1988. A survey of the pectic content of nonlignified monocot cell walls. *Plant Physiol.* 88:309-314.

Jensen JK, Kim H, Cocuron JC, Orler R, Ralph J, Wilkerson CG. 2011. The DUF579 domain containing proteins IRX15 and IRX15-L affect xylan synthesis in Arabidopsis. *Plant J.* 66:387-400.

Jia Z, Qin Q, Darvill AG, York WS. 2003. Structure of the xyloglucan produced by suspension-cultured tomato cells. *Carbohydr. Res.* 338:1197-1208.

Johansson M, Samuelson O. 1977. Reducing end groups in brich xylan and their alkaline degradation. *Wood Sci. Technol.* 11:251-263.

Katoh K, Kuma K, Toh H, Miyata T. 2005a. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33:511-518.

Katoh K, Kuma K, Toh H, Miyata T. 2005b. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic acids research* 33:511-518.

Keppler BD, Showalter AM. 2010. IRX14 and IRX14-LIKE, two glycosyl transferases involved in glucuronoxylan biosynthesis and drought tolerance in Arabidopsis. *Mol. Plant* 3:834-841.

Kong Y, Zhou G, Avci U, Gu X, Jones C, Yin Y, Xu Y, Hahn MG. 2009. Two poplar glycosyltransferase genes, PdGATL1.1 and PdGATL1.2, are functional orthologs to PARVUS/AtGATL1 in Arabidopsis. *Mol. Plant* 2:1040-1050.

Kremer C, Pettolino F, Bacic A, Drinnan A. 2004. Distribution of cell wall components in Sphagnum hyaline cells and in liverwort and hornwort elaters. *Planta* 219:1023-1035.

Kulkarni AR, Pattathil S, Hahn MG, York WS, O'Neill MA. 2012a. Comparison of arabinoxylan structure in bioenergy and model grasses. *Ind. Biotech.* 8:222-229.

Kulkarni AR, Pena MJ, Avci U, Mazumder K, Urbanowicz BR, Pattathil S, Yin Y, O'Neill MA, Roberts AW, Hahn MG*, et al.* 2012b. The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes. *Glycobiology* 22:439-451.

Lee C, O'Neill MA, Tsumuraya Y, Darvill AG, Ye ZH. 2007a. The irregular xylem9 mutant is deficient in xylan xylosyltransferase activity. *Plant Cell Physiol.* 48:1624-1634.

Lee C, Teng Q, Huang W, Zhong R, Ye Z. 2009a. The poplar GT8E and GT8F glycosyltransferases are functional orthologs of Arabidopsis PARVUS involved in glucuronoxylan biosynthesis. *Plant Cell Physiol.* 50:1982.

Lee C, Teng Q, Huang W, Zhong R, Ye Z. 2010a. The arabidopsis family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. *Plant Physiol.* 153:526.

Lee C, Teng Q, Huang W, Zhong R, Ye ZH. 2009b. The F8H glycosyltransferase is a functional paralog of FRA8 involved in glucuronoxylan biosynthesis in Arabidopsis. *Plant Cell Physiol.* 50:812-827.

Lee C, Teng Q, Huang W, Zhong R, Ye ZH. 2009c. The poplar GT8E and GT8F glycosyltransferases are functional orthologs of Arabidopsis PARVUS involved in glucuronoxylan biosynthesis. *Plant Cell Physiol.* 50:1982-1987.

Lee C, Teng Q, Huang W, Zhong R, Ye ZH. 2010b. The Arabidopsis family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. *Plant Physiol.* 153:526-541.

Lee C, Teng Q, Zhong R, Ye ZH. 2011a. The four Arabidopsis reduced wall acetylation genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. *Plant. Cell. Physiol.* 52:1289-1301.

Lee C, Teng Q, Zhong R, Ye ZH. 2011b. Molecular dissection of xylan biosynthesis during wood formation in poplar. *Mol. Plant* 4:730-747.

Lee C, Zhong R, Richardson EA, Himmelsbach DS, McPhail BT, Ye ZH. 2007b. The PARVUS gene is expressed in cells undergoing secondary wall thickening and is essential for glucuronoxylan biosynthesis. *Plant Cell Physiol.* 48:1659-1672.

Lee C, Zhong R, Ye ZH. 2012a. Arabidopsis family GT43 members are xylan xylosyltransferases required for the elongation of the xylan backbone. *Plant Cell Physiol.* 53:135-143.

Lee C, Zhong R, Ye ZH. 2012b. Biochemical characterization of xylan xylosyltransferases involved in wood formation in poplar. *Plant Signal Behav* 7:332-337.

Lewandowski I, Scurlock JMO, Lindvall E, Christou M. 2003. The development and current status of perennial rhizomatous grasses as energy crops in the US and Europe. *Biomass Bioenergy* 25:335-361.

Liepman AH, Wightman R, Geshi N, Turner SR, Scheller HV. 2010. Arabidopsis - a powerful model system for plant cell wall research. *Plant J.* 61:1107-1121.

Ligrone R. 1986. Structure, development and cytochemistry of mucilage-secreting hairs in the moss Timmiella-barbuloides (brid) moenk. *Ann. Bot.* 58:859-868.

Manabe Y, Nafisi M, Verhertbruggen Y, Orfila C, Gille S, Rautengarten C, Cherk C, Marcus SE, Somerville S, Pauly M*, et al.* 2011. Loss-of-function mutation of REDUCED WALL ACETYLATION2 in Arabidopsis leads to reduced cell wall acetylation and increased resistance to Botrytis cinerea. *Plant Physiol.* 155:1068-1078.

Manabe Y, Verhertbruggen Y, Gille S, Harholt J, Chong S-L, Pawar PM, Mellerowicz E, Tenkanen M, Cheng K, Pauly M. 2013. RWA Proteins Play Vital and Distinct Roles in Cell Wall O-Acetylation in Arabidopsis thaliana. *Plant Physiol.*:pp. 113.225193.

Mandal P, Pujol C, Damonte E, Ghosh T, Ray B. 2009a. Xylans from Scinaia hatei: Structural features, sulfation and anti-HSV activity. *Int. J. Biol. Macromol.* 46:173-178.

Mandal P, Pujol C, Damonte E, Ghosh T, Ray B. 2009b. Xylans from Scinaia hatei: structural features, sulfation and anti-HSV activity. *Int. J. Biol. Macromol.*

Masuko T, Minami A, Iwasaki N, Majima T, Nishimura S, Lee YC. 2005. Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. *Anal. Biochem.* 339:69-72.

Matsunaga T, Ishii T, Matsumoto S, Higuchi M, Darvill A, Albersheim P, O'Neill M. 2004. Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. *Plant Physiol.* 134:339-351.

Matulewicz M, Cerezo A. 1987. Alkali-soluble polysaccharides from *Chaetangium fastigiatum*: Structure of a xylan. *Phytochemistry* 26:1033-1035.

Mazumder K, York WS. 2010. Structural analysis of arabinoxylans isolated from ball milled switchgrass biomass. *Carbohydr. Res.* 345:2183-2193.

McCartney L, Marcus SE, Knox JP. 2005. Monoclonal antibodies to plant cell wall xylans and arabinoxylans. *J. Histochem. Cytochem.* 53:543-546.

Meikle PJ, Bonig I, Hoogenraad NJ, Clarke AE, Stone BA. 1991. The location of (1-->3)-β-glucans in the walls of pollen tubes of *Nicotiana-alata* using a (1-->3)-β-glucan-specific monoclonal-antibody. *Planta* 185:1-8.

Meikle PJ, Hoogenraad NJ, Bonig I, Clarke AE, Stone BA. 1994. A (1-->3,1-->4)-β-glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1-->3,1-->4)-β-glucans. *Plant J.* 5:1-9.

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

Mishler B, Kelch D. 2009. Phylogenomics and early land plant evolution. *Bryophyte Biology*:173–197.

Mohnen D, Bar-Peled M, Somerville C. 2008. Cell wall polysaccharide synthesis. In: Himmel ME editor. Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. : Blackwell Publishing, Singapore. p. 94-159.

Moller I, Sorensen I, Bernal AJ, Blaukopf C, Lee K, Obro J, Pettolino F, Roberts A, Mikkelsen JD, Knox JP*, et al.* 2007. High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *Plant J.* 50:1118-1128.

Mortimer JC, Miles GP, Brown DM, Zhang Z, Segura MP, Weimar T, Yu X, Seffen KA, Stephens E, Turner SR*, et al.* 2010. Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. *Proc. Natl. Acad. Sci. U S A* 107:17409-17414.

Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M. 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.* 96:673-686.

Mur LA, Allainguillaume J, Catalan P, Hasterok R, Jenkins G, Lesniewska K, Thomas I, Vogel J. 2011. Exploiting the Brachypodium Tool Box in cereal and grass research. *New Phytol.* 191:334-347.

Niklas K. 1997. The evolutionary biology of plants University of Chicago Press.

Nuin PA, Wang Z, Tillier ER. 2006. The accuracy of several multiple sequence alignment programs for proteins. *BMC Bioinformatics* 7:471.

Obel N, Porchia AC, Scheller HV. 2002. Dynamic changes in cell wall polysaccharides during wheat seedling development. *Phytochemistry* 60:603-610.

Oikawa A, Joshi HJ, Rennie EA, Ebert B, Manisseri C, Heazlewood JL, Scheller HV. 2010. An integrative approach to the identification of Arabidopsis and rice genes involved in xylan and secondary wall development. *PLoS One* 5:e15481.

Pattathil S, Avci U, Baldwin D, Swennes AG, McGill JA, Popper Z, Bootten T, Albert A, Davis RH, Chennareddy C*, et al.* 2010. A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol.* 153:514-525.

Pattathil S, Avci U, Miller JS, Hahn MG. 2012. Immunological approaches to plant cell wall and biomass characterization: glycome profiling. In: Himmel ME editor. Biomass Conversion: Methods and Protocols: Humana Press.

Pauly M, Albersheim P, Darvill A, York WS. 1999. Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. *Plant J* 20:629-639.

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

Pauly M, Keegstra K. 2008. Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J.* 54:559-568.

Pauly M, Keegstra K. 2010. Plant cell wall polymers as precursors for biofuels. *Curr. Opin. Plant Biol.* 13:304-311.

Peña MJ, Darvill AG, Eberhard S, York WS, O'Neill MA. 2008. Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology* 18:891-904.

Peña MJ, Zhong R, Zhou GK, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye ZH. 2007. Arabidopsis irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19:549-563.

Percival E, Chanda S. 1950a. The xylan of *Rhodymenia palmata*.

Percival EGV, Chanda SK. 1950b. The Xylan of Rhodymenia palmata. *Nature* 166:787-788.

Persson S, Wei H, Milne J, Page GP, Somerville CR. 2005. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Natl Acad Sci U S A* 102:8633-8638.

Popper Z. 2011. Evolution and diversity of plant cell walls: from algae to flowering plants. *Ann. Rev. Plant Biol.* 62:567-590.

Popper ZA. 2008. Evolution and diversity of green plant cell walls. *Curr. Opin. Plant Biol.* 11:286-292.

Popper ZA, Fry SC. 2003. Primary cell wall composition of bryophytes and charophytes. *Ann. Bot.* 91:1-12.

Popper ZA, Tuohy MG. 2010. Beyond the green: understanding the evolutionary puzzle of plant and algal cell walls. *Plant Physiol.* 153:373-383.

Qiu Y, Li L, Wang B, Chen Z, Knoop V, Groth-Malonek M, Dombrovska O, Lee J, Kent L, Rest J. 2006. The deepest divergences in land plants inferred from phylogenomic evidence. *Proc. Natl. Acad. Sci. U S A* 103:15511-15516.

Reis A, Pinto P, Evtuguin D, Neto CP, Domingues P, Ferrer-Correia A, Domingues MRM. 2005. Electrospray tandem mass spectrometry of underivatised acetylated xylo-oligosaccharides. *Rapid Commun. Mass Sp.* 19:3589-3599.

Rennie EA, Hansen SF, Baidoo EE, Hadi MZ, Keasling JD, Scheller HV. 2012. Three members of the Arabidopsis glycosyltransferase family 8 are xylan glucuronosyltransferases. *Plant Physiol.* 159:1408-1417.

Sandhu AP, Randhawa GS, Dhugga KS. 2009. Plant cell wall matrix polysaccharide biosynthesis. *Mol. Plant* 2:840-850.

Scheller HV, Ulvskov P. 2010. Hemicelluloses. *Ann. Rev. Plant Biol.* 61:263-289.

Selvendran RR, Davies AMC, Tidder E. 1975. Cell wall glycoproteins and polysaccharides of mature runner beans. *Phytochemistry* 14:2169-2174.

Selvendran RR, O'Neill MA. 1987. Isolation and analysis of cell walls from plant material. *Methods Biochem. Anal.* 32:25-152.

Shatalov AA, Evtuguin DV, Pascoal Neto C. 1999. (2-O-α-D-galactopyranosyl-4-O-methyl-α-D-glucurono)-D-xylan from *Eucalyptus globulus* Labill. *Carbohydr. Res.* 320:93-99.

Shimizu K, Hashi M, Sakurai K. 1978. Isolation from a softwood xylan of oligosaccharides containing two 4-O-methyl-D-glucuronic acid residues. *Carbohydr. Res.* 62:117-126.

Shimizu K, Ishihara M, Ishihara T. 1976. Hemicellulases of brown rotting fungus Tyromyces palustris. II. The oligosaccharides from the hydrolysate of a hardwood xylan by the intracellular xylanase. *Mokuzai Gakkaishi*:618-625.

Simas-Tosin FF, de Souza LM, Wagner R, Pereira GC, Barraza RR, Wendel CF, Sassaki GL, Iacomini M, Gorin PA. 2013. Structural characterization of a glucuronoarabinoxylan from pineapple (*Ananas comosus* (L.) Merrill) gum exudate. *Carbohydr. Polym.* 94:704-711.

Smith B, Harris P. 1999. The polysaccharide composition of Poales cell walls:: Poaceae cell walls are not unique. *Biochem. Syst. Ecol.* 27:33-53.

Smith BG, Harris PJ. 1995. Polysaccharide composition of unlignified cell walls of pineapple [*Ananas comosus* (L.) Merr.] fruit. *Plant Physiol.* 107:1399-1409.

Sorensen I, Domozych D, Willats WG. 2010. How have plant cell walls evolved? *Plant Physiol.* 153:366-372.

Sterling JD, Atmodjo MA, Inwood SE, Kumar Kolli VS, Quigley HF, Hahn MG, Mohnen D. 2006. Functional identification of an Arabidopsis pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proc Natl Acad Sci U S A* 103:5236-5241.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24:1596.

Taylor T, Taylor E, Krings M. 2009. Paleobotany: The biology and evolution of fossil plants Academic Pr.

Toole GA, Le Gall G, Colquhoun IJ, Nemeth C, Saulnier L, Lovegrove A, Pellny T, Wilkinson MD, Freeman J, Mitchell RA*, et al.* 2010. Temporal and spatial changes in cell wall composition in developing grains of wheat cv. Hereward. *Planta* 232:677-689.

Turvey JR, Williams EL. 1970. Structures of some xylans from red algae. *Phytochemistry* 9:2383-&.

Urbanowicz BR, Pe√±a MJ, Ratnaparkhe S, Avci U, Backe J, Steet HF, Foston M, Li H, O,ÄôNeill MA, Ragauskas AJ. 2012. 4-O-methylation of glucuronic acid in Arabidopsis glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. *Proc. Natl. Acad. Sci. U S A* 109:14253-14258.

Verbruggen MA, Spronk BA, Schols HA, Beldman G, Voragen AG, Thomas JR, Kamerling JP, Vliegenthart JF. 1998. Structures of enzymically derived oligosaccharides from sorghum glucuronoarabinoxylan. *Carbohydr. Res.* 306:265-274.

Vogel J. 2008. Unique aspects of the grass cell wall. *Curr. Opin. Plant Biol.* 11:301-307.

Willis LM, Whitfield C. 2013. Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways. *Carbohydr. Res.* 378:35-44.

Wise LE, Murphy M, Daddieco AA. 1946. Chlorite Holocellulose, Its Fractionation and Bearing on Summative Wood Analysis and on Studies on the Hemicelluloses. *Tech Assoc Pap* 29:210-218.

Wu AM, Hornblad E, Voxeur A, Gerber L, Rihouey C, Lerouge P, 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 Physiol.* 153:542-554.

Wu AM, Rihouey C, Seveno M, Hornblad E, Singh SK, Matsunaga T, Ishii T, Lerouge P, Marchant A. 2009. The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J.* 57:718-731.

Xiong G, Cheng K, Pauly M. 2013. Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the Arabidopsis mutant tbl29. *Mol Plant* 6:1373-1375.

Yamagaki T, Maeda M, Kanazawa K, Ishizuka Y, Nakanishi H. 1997. Structural clarification of Caulerpa cell wall -1, 3-xylan by NMR spectroscopy. *BioSci., BioTechnol., and Biochem* 61:1077-1080.

Yin Y, Chen H, Hahn MG, Mohnen D, Xu Y. 2010. Evolution and function of the plant cell wall synthesis-related Glycosyltransferase Family 8. *Plant Physiol.*

Yoon H, Hackett J, Ciniglia C, Pinto G, Bhattacharya D. 2004. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 21:809.

York WS, O'Neill MA. 2008. Biochemical control of xylan biosynthesis - which end is up? *Curr. Opin. Plant Biol.* 11:258-265.

Zeng W, Jiang N, Nadella R, Killen TL, Nadella V, Faik A. 2010. A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. *Plant Physiol.* 154:78-97.

Zhong R, Peña MJ, Zhou GK, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WH, 3rd, Darvill AG, York WS, Ye ZH. 2005. Arabidopsis fragile fiber8, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. *Plant Cell* 17:3390-3408.

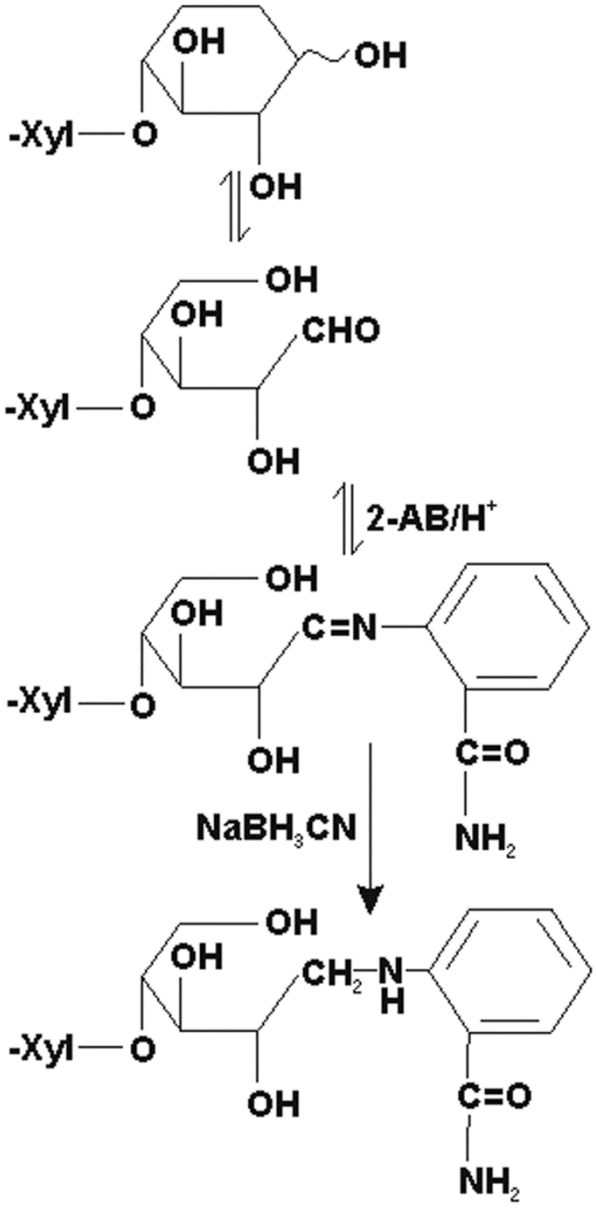
Zhong R, Ye ZH. 2003. Unraveling the functions of glycosyltransferase family 47 in plants. *Trends Plant Sci.* 8:565-568.

Zhong R, Ye ZH. 2007. Regulation of cell wall biosynthesis. *Curr. Opin. Plant Biol.* 10:564-572.

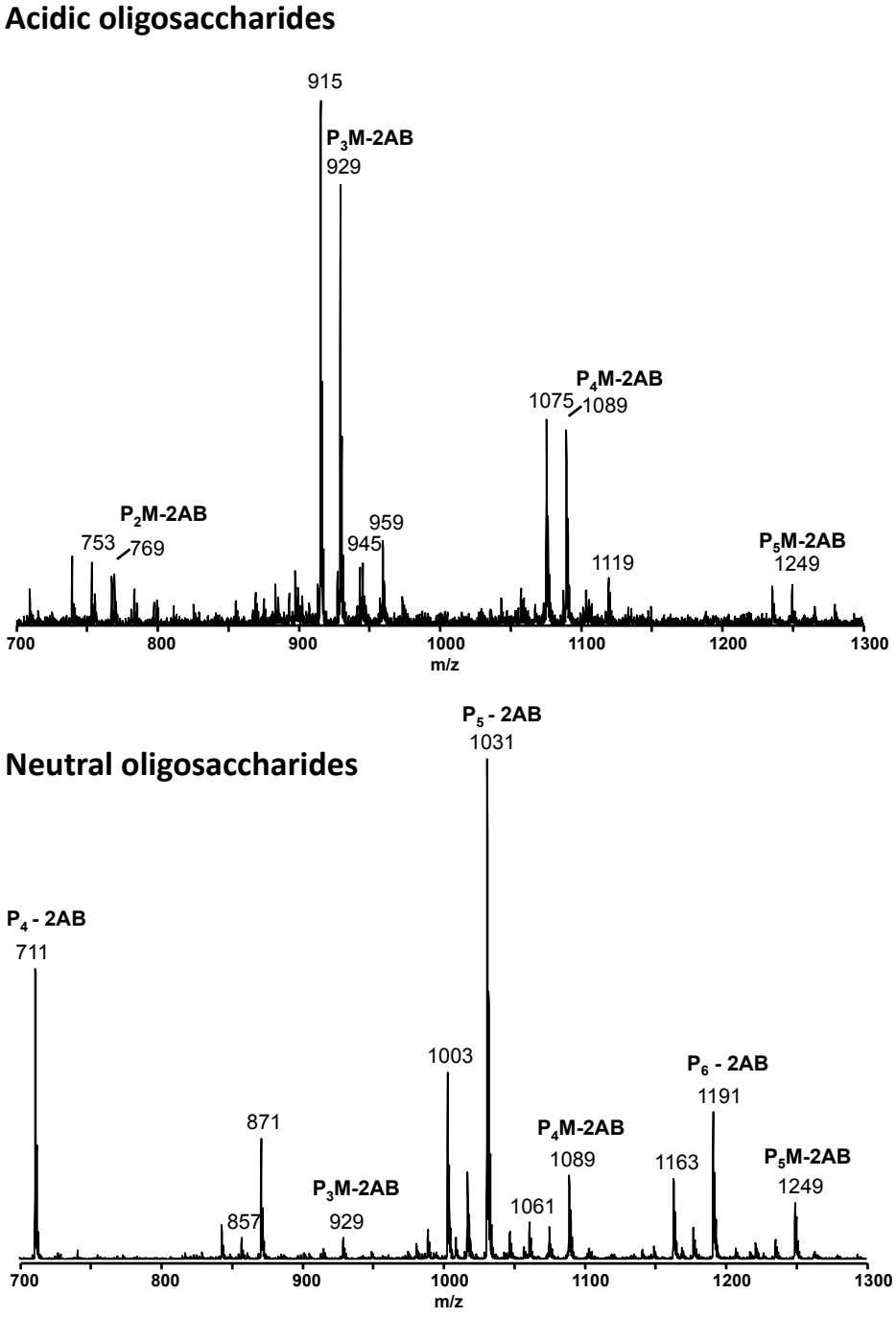
Zhou GK, Zhong R, Himmelsbach DS, McPhail BT, Ye ZH. 2007. Molecular characterization of PoGT8D and PoGT43B, two secondary wall-associated glycosyltransferases in poplar. *Plant Cell Physiol.* 48:689-699.

Zhou GK, Zhong R, Richardson EA, Morrison WH, 3rd, Nairn CJ, Wood-Jones A, Ye ZH. 2006. The poplar glycosyltransferase GT47C is functionally conserved with Arabidopsis Fragile fiber8. *Plant Cell Physiol.* 47:1229-1240.

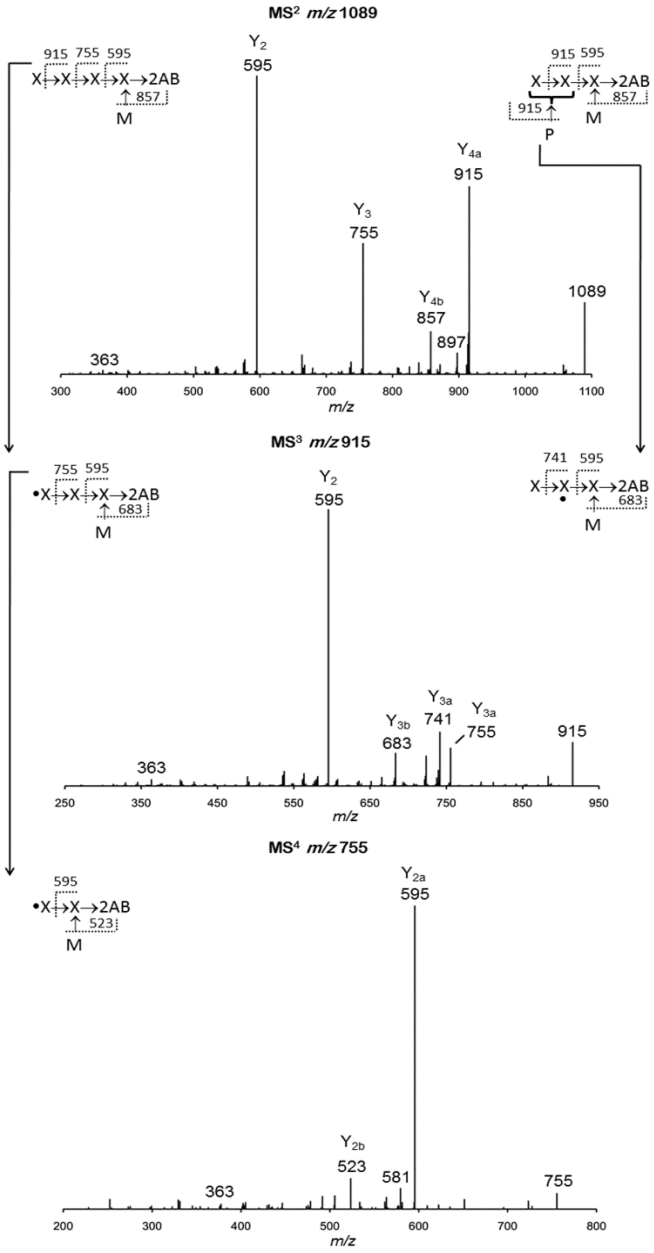
Supplemental Information



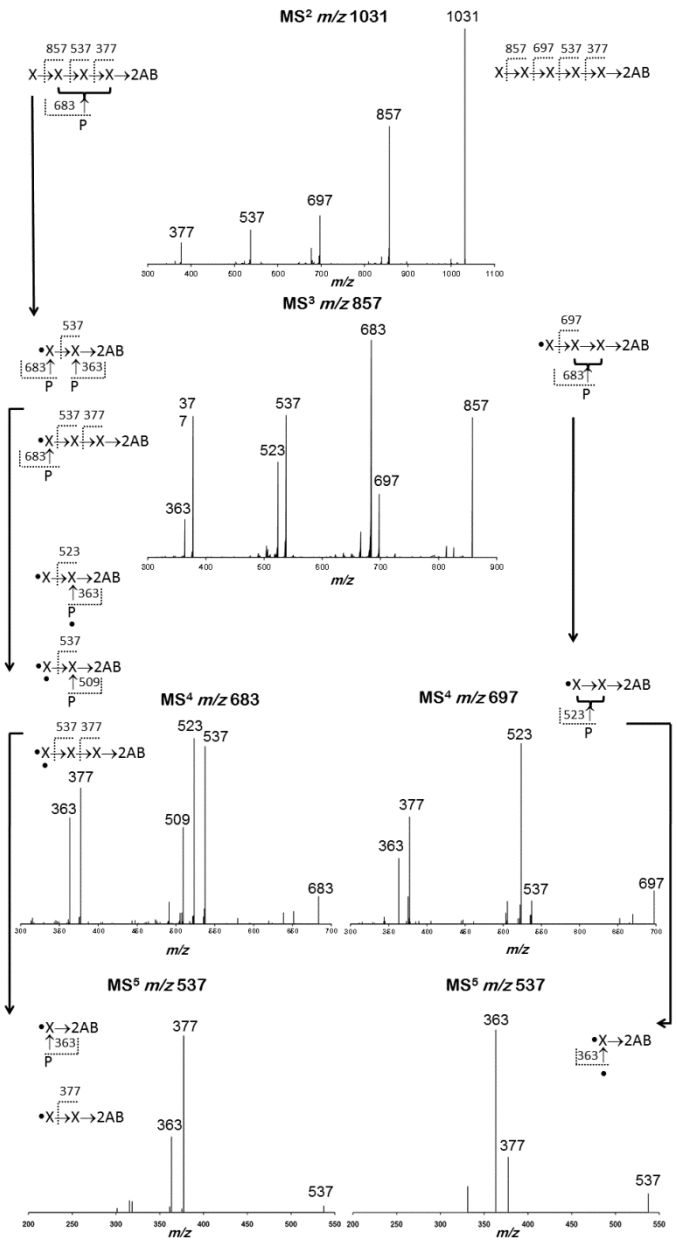
Supplemental Figure S4.1: 2-AB labeling of the free reducing ends of xylan. The amino group of the 2-AB couples with the open ring form of the glycan to form a Schiff's base. The Schiff's base is reduced to give a stable 2-AB labeled conjugate.



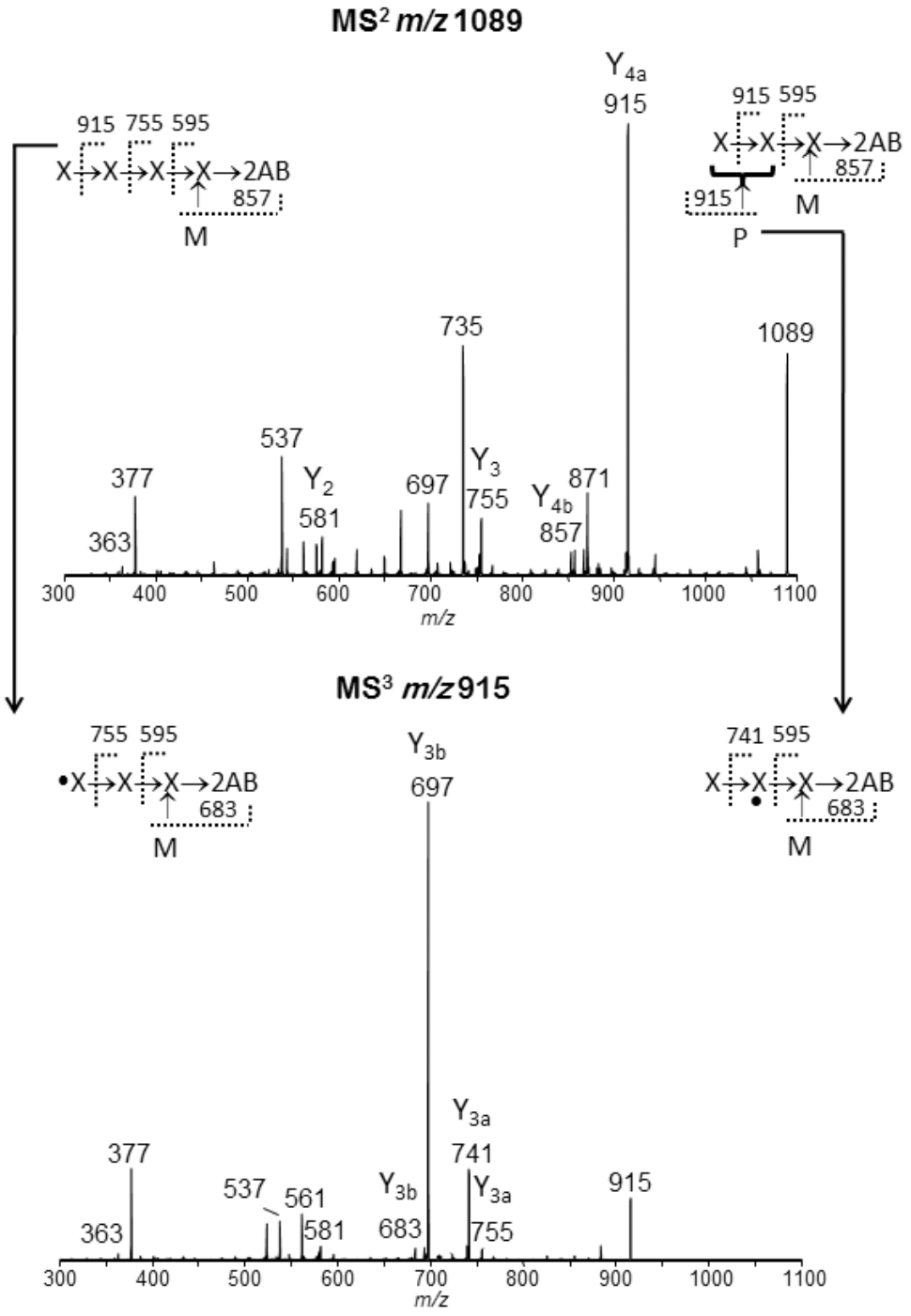
Supplemental Figure S4.2: MS spectrum of per-*O*-methylated 2-AB labeled xylo-oligosaccharides from S. italica indicating the presence of [M + Na]+ ions at *m/z* 769, 929, 1089, and 1249 composed of P_2M, P_3M, P_4M, and P_5M respectively.



Supplemental Figure S4.3: ESI-MS^n of precursor ion at *m/z* 1089 of xylo-oligosaccharide generated by per-*O*-methylation of endo-xylanase treated 1 M KOH soluble extract of 2-AB labeled AIR from *S. italica*. Similar structures were found in *Miscanthus*, *P. virgatum*, *Oryza*, and *Brachypodium*. Fragmentation events leading to the generation of Y ions are indicated in each spectrum.



Supplemental Figure S4.4: ESI-MS^n of precursor ion at *m/z* 1031 of xylo-oligosaccharide generated by per-*O*-methylation of endo-xylanase treated 1 M KOH soluble extract of 2-AB labeled AIR from *S. italica*. Similar structures were found in *Miscanthus*, *P. virgatum*, *Oryza*, and *Brachypodium*. Fragmentation events leading to the generation of Y ions are indicated in each spectrum.



Supplemental Figure S4.5: ESI-MS^n spectra of per-*O*-methylated 2-AB labeled xylo-oligosaccharides (precursor ion at *m/z* 1089) from *Crinum* establishing the presence of substituted xylose residue at the reducing end, consistent with that in grasses.

Chapter 5

Conclusion

The current interest in the use of woody plants and grasses for the production of biofuels and other value-added chemicals has led to increased efforts to understand the molecular and biochemical factors that control xylan synthesis and structure in these plants. Such knowledge is required to facilitate the development of technologies to modify xylan structure and its interaction with other wall components to improve the economic value of plant biomass.

In my dissertation research, I have described the structural characterization of the predominant xylans present in the cell walls of members of the major groups of land plants. My study illustrated the diversity of xylan structures (See Chapter 2, 3, and 4) and the conserved nature of the glycosyl sequences at the reducing ends of monocots and other land plants. My results indicate that the xylans from eudicots, together with all monocots except the Asparagles and Poales have the unique glycosyl sequence 4-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,3)-α-L-Rha*p*-(1,2)-α-D-Gal*p*A-(1,4)-D-Xyl*p* (Sequence **1**) at their reducing ends. Sequence 1 is also present at the reducing end of gymnosperm xylans. This knowledge together with the results of studies correlating that mutations in the synthesis of Sequence **1** with the production of xylan having a considerably low DP suggest this glycosyl sequence is involved in controlling xylan chain length and possibly its ability to interact with other wall components. Indeed, several authors have hypothesized that sequence **1** is either a primer that initiates xylan chain formation or a sequence that is used to terminate xylan synthesis (Peña et al. 2007, York and O'Neill 2008). In my thesis I

have extended these models for the mechanism by which sequence **1** may control xylan synthesis.

The presence of xylans in the cell walls of the mosses (a group of avascular land plants) has been was a subject of debate. I resolved this debate by isolating and structurally characterizing a glucuronoxylan from the cell walls of the moss *Physocmitrella patens*. I also structurally characterized the xylans present in the cell walls of two vascular non-flowering plants, the lycophyte *Selaginella krussiana,* and the vascular monilophyte *Equisetum hyemale* (Chapter 2). My study established that mosses synthesize glucuronoxylan, albeit in small amounts, that is structurally homologous to the glucuronoxylans present in the secondary cell walls of lycopodiophytes, monilophytes, and many seed-bearing plants. However, *P. patens*, *S. kraussiana,* and *E. hyemale* glucuronoxylans lack the distinct reducing end sequence characteristic of gymnosperm and dicot GX. I also performed phylogenetic analysis of glycosyltransferase family 43, and 47 to demonstrate that several of the glycosyl transferases required for glucuronoxylan synthesis likely evolved before the evolution of vascular plants. The presence of these common glucuronoxylan structures and glycosyl transferase genes in both *P. patens* and vascular plants led us to hypothesize that they have a common ancestry and that changes in cell wall polysaccharides provide insight into some of the events that occurred during the evolution of land plants.

It has been previously suggested that Sequence **1** is absent in grass heteroxylans (York and O'Neill 2008, Fincher 2009, Scheller and Ulvskov 2010, Pauly et al. 2013). To further extend our understanding of xylan structures in grasses, I structurally characterized xylans from the cell walls of two bioenergy grasses (switchgrass and miscanthus) and three model grasses (rice, foxtail millet, and brachypodium). I have demonstrated that the xylans of these grasses

lacked discernable amounts of Sequence **1** at their reducing ends suggesting that the xylan chain length in Poaceae is controlled by a mechanism that does not involve Sequence **1** (Chapter 3).

There is no published literature on the nature of the reducing ends in grasses and other non-grass monocots. Thus, the possibility could not be eliminated that monocot xylans possess a unique glycosyl sequences at their reducing ends. I therefore isolated xylans from the walls of plants representing 10 of the 12 orders of monocots and determined their major structural features including the nature of their reducing ends (Chapter 4). These analyses provided insight in the transition of glucuronoxylan to arabinoxylan during monocot evolution. The data supported the notion that glucuronoarabinoxylan is predominant in commelinids walls, whereas glucuronoxylan is predominant in the walls of the non-commelinids. My characterization of monocot xylans also showed different trends in the cell wall structures, thus providing further correlation between changes in cell wall structures and taxonomy of land plants. My data clearly demonstrate that heteroxylans synthesized by members of all of the orders of monocots with the exception of Poales and Asparagales have Sequence **1** at their reducing ends. Thus, Sequence **1** was not eliminated when monocots and dicots diverged but was lost during the evolution of some monocot orders. However, the absence of Sequence **1** in the xylans from the Poales and Asparagales suggest that this glycosyl sequence is not required for xylan synthesis in these plants.

The absence of Sequence **1** in xylans in the walls of the Poales and Asparagales led me to determine the nature of the glycosyl sequences at the reducing ends of these xylans. Such information is required to develop an understanding of the mechanisms involved in arabinoxylan/glucuronoarabinoxylan synthesis. I showed that in these plants the reducing end xylose is substituted at *O*-2 with a MeGlc*p*A (ARK-1 and ARK-2 sequences) or at *O*-3 with a

Ara*f* residue (ARK-3 and ARK-4 sequences). The presence of these sequences in the Poales and Asparagales, suggest that new mechanisms for xylan synthesis using ARK sequences have evolved in these plants. My data also suggest that these distinctive ARK sequences are conserved within these monocots orders.

The conserved ARK sequences present at the reducing ends of Poales and Asparagales xylans and Sequence **1** in most other seed-bearing plants suggest that these two structures may have similar roles in xylan synthesis. That is, ARK sequences may prime or terminate xylan synthesis. However, the lack of mutants encoding glycosyltransferase involved in the formation of the ARK sequences limits our ability to examine the function of these sequences in xylan synthesis. The possibility that dicots synthesize xylan using a block wise assembly process has been discussed (**Chapter 1, Figure 1.5**).

The results of my research provide important information regarding changes in xylan structure during the evolutionary of land plants. This information also provides further evidence to support our hypothesis on correlation between changes in cell wall polysaccharide structures and land plant evolution. Through my study I was also able to establish that ARK sequences are present at the reducing end of the xylans synthesized by the Poales and Asparagales. This information provides the basis for developing new insight into the mechanism of synthesis of xylan in grasses and allowed me to make predictive models for this process. My studies also have a direct bearing on approaches for modifying xylan structure to improve the value of plant biomass for commercial biofuel production and for the production of other added values chemicals. The results presented in this thesis are relevant for the emerging bioprocessing industries as it is likely that such technologies will be designed to use enzymes that saccharify biomass that contains either glucuronoxylan or arabinoxylan. My studies are also likely to be of

relevance in the development of bio-based processing technologies that generate value-added chemicals from diverse plant waste products rather than from dedicated energy crops.

References

Fincher GB. 2009. Revolutionary Times in Our Understanding of Cell Wall Biosynthesis and Remodeling in the Grasses. *Plant Physiol.* 149:27-37.

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

Peña MJ, Zhong R, Zhou GK, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye ZH. 2007. Arabidopsis irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19:549-563.

Scheller HV, Ulvskov P. 2010. Hemicelluloses. *Ann. Rev. Plant Biol.* 61:263-289.

York WS, O'Neill MA. 2008. Biochemical control of xylan biosynthesis - which end is up? *Curr. Opin. Plant Biol.* 11:258-265.