

The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes.

Ameya Kulkarni², Maria J. Peña², Utku Avci², Koushik Mazumder², Breeanna Urbanowicz², Sivakumar Pattathil², Yanbin Yin³, Malcolm A. O'Neill², Alison W. Roberts⁴, Michael G. Hahn^{2,5}, Ying Xu³, Alan G. Darvill^{2,6}, and William S. York^{1,2,6}

²Complex Carbohydrate Research Center and US Department of Energy Bioenergy Science Center, The University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA,

³Computational System Biology Laboratory, Department of Biochemistry and Molecular Biology and DOE Bioenergy Science Center, The University of Georgia, Athens, GA 30602, USA,

⁴Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881, USA,

⁵Department of Plant Biology, The University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA,

⁶Department of Biochemistry and Molecular Biology, The University of Georgia, Athens, GA 30602, USA

¹To whom correspondence should be addressed: Tel:1-706-542-4628; Fax: 706-542-4412;

e-mail: will@ccrc.uga.edu

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 (GlcA) sidechain at O-2. By contrast, the lycopodiophyte *Selaginella kraussiana* synthesizes a glucuronoxylan substituted with 4-O-Me- α -D-GlcA sidechains, as do many hardwood species. The monilophyte *Equisetum hyemale* produces a glucuronoxylan with both 4-O-Me- α -D-GlcA and α -D-GlcA 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.

Keywords: Glucuronoxylan, plant cell wall, Physcomitrella, Selaginella, Equisetum, land plant evolution

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). Although branched 1,4-linked β -D-xylans are found in both primary and secondary cell walls of vascular plants, they are typically a minor component of primary cell walls, except in grasses, where most cell walls contain a considerable amount of xylan. 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 2011, Popper and Tuohy 2010, Sorensen et al. 2010).

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 (Xylp) residues but differ in the type, location, and number of glycosyl residues attached to this backbone. For example, many eudicots synthesize glucuronoxylans that have α -D-glucosyluronic acid (α -D-GlcpA) and/or a 4-O-methyl

α -D-glucosyluronic acid (4-O-Me-GlcpA) sidechains at O-2 of the backbone residues. Gymnosperms synthesize glucuronoarabinoxylans in which backbone residues are substituted at O-2 with 4-O-Me-GlcpA and at O-3 with α -L-arabinofuranosyl (Araf) 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-Xylp-(1,4)- β -D-Xylp-(1,3)- α -L-Rhap-(1,2)- α -D-GalpA-(1,4)-D-Xylp at their reducing ends (Johansson and Samuelson 1977, Peña et al. 2007, Shimizu et al. 1976). The Poaceae (grasses) typically produce arabinoxylans and glucuronoarabinoxylans substituted predominantly with α -L-Araf residues at *O*-2 and/or *O*-3 and less frequently with GlcpA and/or 4-*O*-Me-GlcpA 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 glucuronoxyllans that are substituted at *O*-2 with 4-*O*-Me-GlcpA (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 glucuronoxyllan that is structurally homologous to glucuronoxyllans 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 glucuronoxyllans 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 1a, 1c and

S1a) and specific polysaccharide epitopes were localized by immunolabeling (Figures 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 1b and 1d, arrows) identified as axillary hair cells (Hiwatashi et al. 2001, Ligrone 1986), which are located between leaves and the leafy shoot. High-resolution transmission electron microscopy (Figure S2) 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 S1c – 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 1f and 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 1f and 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 S1f). Leafy shoot cells were not labeled by any of the xylan-directed monoclonal antibodies used in this study (Figures 1b, 1d, 1f, 1h, and S1c – f). Furthermore, immunolabeling of cross sections prepared from protonema indicated that the abundance of xylan epitopes in this tissue is very low (data not shown).

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 1e and 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 S1g), 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 S1h).

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 S3) 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. In contrast, analysis of alkali extracts prepared from protonema AIR indicated that this material contained little if any xylan. Glycosyl-linkage composition analysis also revealed that derivatives of 1,4-linked and 1,2,4-linked Xylp residues are abundant in the 4M KOH extract (Figure S4). 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 ^1H NMR spectroscopy, and electrospray ionization multiple mass spectrometry (ESI-MSⁿ).

Virtually all of the detected oligosaccharides generated by xylanase treatment of the 4M KOH soluble extract were shown by MALDI-TOF-MS (Figure S5) 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, obscured by matrix ions.

The purified, xylanase-generated *P. patens* oligosaccharides were further characterized by ^1H NMR spectroscopy. The 2D gCOSY spectrum (Figure 2a) provided chemical shift and scalar coupling information that, in combination with previously published data (Peña et al. 2007, Verbruggen et al. 1998), 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-GlcpA residues linked to O2 of the xylosyl backbone (residues A-G, Table I, Figure 2A). The downfield shift of H-2 of β -Xylp residue F (δ

3.482) relative to H-2 of unbranched β -Xylp residues B-F (δ 3.25-3.29) confirmed that the xylan backbone is substituted at O-2 by α -D-GlcpA (Peña et al. 2007). No resonance that could be assigned to 4-*O*-Me GlcpA, Araf sidechains attached to Xylp residues or Xylp residues bearing Araf 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 ^1H -NMR spectra of the crude endoxylanase-treated 4 M KOH extract, but these were not observed in the ^1H -NMR spectra of the purified glucuronoxylan oligomers. (See Figure S6). Thus, the Araf 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 S3). These data suggest that Araf sidechains, if present at all, are a minor component of the *P. patens* glucuronoxylan.

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-MSⁿ. 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 (P₆G₂) is an oligosaccharide with two hexuronosyl sidechains separated by a single xylosyl residue (Figure 3). MSⁿ of the low-abundance quasimolecular ion at m/z 1407 (P₇G, 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, $^1\text{H-NMR}$ analysis indicates that the purified *P. patens* endoxylanase-generated oligosaccharides contain few, if any, arabinofuranosyl sidechains.

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 endoxylanase to generate oligosaccharides, which were partially purified by SEC and analyzed by $^1\text{H-NMR}$ spectroscopy. The 2D gCOSY NMR spectrum of the *S. kraussiana* xylo-oligosaccharides (Figure 2B, Table II) revealed resonances with chemical shifts and scalar coupling patterns (Peña et al. 2007) diagnostic for the presence of 1,4-linked $\beta\text{-D-Xylp}$ residues, some of which are substituted at O-2 with 4-*O*-Me- $\alpha\text{-D-GlcpA}$. No resonances indicating the presence of (unmethylated) $\alpha\text{-D-GlcpA}$ sidechains were observed in this spectrum. The 2D gCOSY spectrum of *E. hyemale* xylo-oligosaccharides (Figure 2C, Table III) contained resonances diagnostic for the presence of both 4-*O*-Me- $\alpha\text{-D-GlcpA}$ and $\alpha\text{-D-GlcpA}$ sidechains at O-2 of 1,2,4-linked $\beta\text{-D-Xylp}$ 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- $\alpha\text{-D-GlcpA}$ and $\alpha\text{-D-GlcpA}$ sidechains (Peña et al. 2007)

Two gymnosperms (spruce and birch) and the eudicot Arabidopsis have been shown to synthesize xylans with the glycosyl sequence 4- $\beta\text{-D-Xylp}$ -(1,4)- $\beta\text{-D-Xylp}$ -(1,3)- $\alpha\text{-L-Rhap}$ -(1,2)-

α -D-GalpA-(1,4)-D-Xylp at their reducing end (Johansson and Samuelson 1977, Peña et al. 2007, Shimizu et al. 1976). 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 (Brown et al. 2009, Zhong et al. 2005, Zhong and Ye 2003, Zhou et al. 2006). *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 (Kong et al. 2009, Lee et al. 2009, Peña et al. 2007). 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 (Lee et al. 2010, Peña et al. 2007, 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. 2010).

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).

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 S1). 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 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 4, Table S2). However, it seems unlikely that the GT43 clade B3 genes of *S. moellendorffii* 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 glycosyltransferases (Figure S8). 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 5, Table S2).

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 (Mishler and Kelch 2009, Qiu et al. 2006). Subsequent evolutionary innovations led to the appearance of vascular plants (tracheophytes) which diverged from the bryophytes around 420 million years ago (Graham et al. 2000, Niklas 1997, 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 (Carafa et al. 2005, Matsunaga et al. 2004, Peña et al. 2008, Popper 2011, Popper 2008, Popper and Fry 2003). 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 GlcpA, 4-*O*-Me-GlcpA, or Araf residues. For example, the cell walls of *Chaetangium fastigiatum* and *Scinaia hatei* have been reported to contain linear 1,4-linked xylans (Mandal et al. 2009, Matulewicz and Cerezo 1987), whereas *Rhodymenia palmata* synthesizes a linear xylan composed of 1,3- and 1,4-linked Xylp (Percival and Chanda 1950). 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-GlcpA 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-GlcpA 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 GlcpA sidechains in the moss glucuronoxylan is also unusual, with pairs of GlcpA sidechains separated by a single xylosyl residue. This branching pattern leads to the generation of oligosaccharide fragments bearing two GlcpA residues upon treatment with endoxylanase. Less densely substituted regions give rise to oligosaccharides bearing zero or one GlcpA 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 moellendorffii* 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-GlcA 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 gymnosperms and eudicots (Figure 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 glucuronoxyylan 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 1e). In contrast, the xylan-directed antibodies LM11, CCRC-M137 (Figure 1), CCRC-M147, CCRC-M154 and CCRC-M160 (Figure S1) 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 glucuronoxyylan primarily 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 glucuronoxyylan 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; Medina et al. 2011). 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 tracheophytes. 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 gymnosperms and eudicots in that it lacks the distinctive oligosaccharide structure present at the reducing end and none of its GlcA 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 GlcA sidechains and the presence of the distinctive reducing end sequence, may be associated with a plant's 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 $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ /8 h dark cycle at 23 $^{\circ}$ 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 $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$) at 24 $^{\circ}$ 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 $^{\circ}$ 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 $^{\circ}$ 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 $^{\circ}$ C 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 μ L). The solution was removed and then KPBS (10 μ L) added to the section for 5 min. The KPBS was removed and undiluted hybridoma supernatant (10 μ L) 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-insoluble 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 CHCl₃-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 μ L 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 starch in the residue. Microscopic analysis of the stained residue 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 μ L 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) NaBH₄ 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) NaBH₄ 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 μ L of glacial acetic acid (Ahlgren and Goring 1971, Wise

et al. 1946). 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 supernatant 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 μ L of the diluted cell wall extracts (60 μ g of sugar/ mL) and allowed to dry overnight at 37^o 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 μ L 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).

¹H-NMR spectroscopy

Xylo-oligosaccharide-enriched material (~1 mg) was dissolved in D₂O (0.5-1.0 mL, 99.9%). ¹H-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 (Kato et al. 2005) 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.oxfordjournals.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; *Araf*, 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 siphonous 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 birch 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, Dombrowska 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, Vliegthart 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.

Figure Legends

Figure 1. Bright field (a and c) and immunofluorescence (b, d – h) 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) Bright-field image of toluidine blue-stained cross section.

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

(c) Magnified portion of bright-field image in 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, which 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.

Figure 2: Partial 600-MHz gCOSY NMR spectrum of purified xylo-oligosaccharides generated by β -endoxyylanase 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.

Figure 3: ESI-MSⁿ 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.

Figure 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. 2005) 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 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).

Table I. ^1H 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 -Xylp (reducing)	5.186	3.546	3.787	---	---	---
B	β -Xylp (reducing)	4.584	3.251	3.546	3.7829	3.377	4.058
C	β -1,4-Xylp (internal) (major)	4.474	3.283	3.578	3.795	3.439	4.155
D	β -1,4-Xylp (internal) (minor)	4.47	3.29	3.56	3.793	3.377	4.107
E	β -Xylp (terminal)	4.457	3.254	3.426	3.624	3.296	3.964
F	β -1,2,4-Xylp (α -Glc pA)	4.642	3.482	3.635	3.812	3.391	4.11
G	α -Glc pA	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 pA) is a β -linked xylosyl residue that bears a Glc pA 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.

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

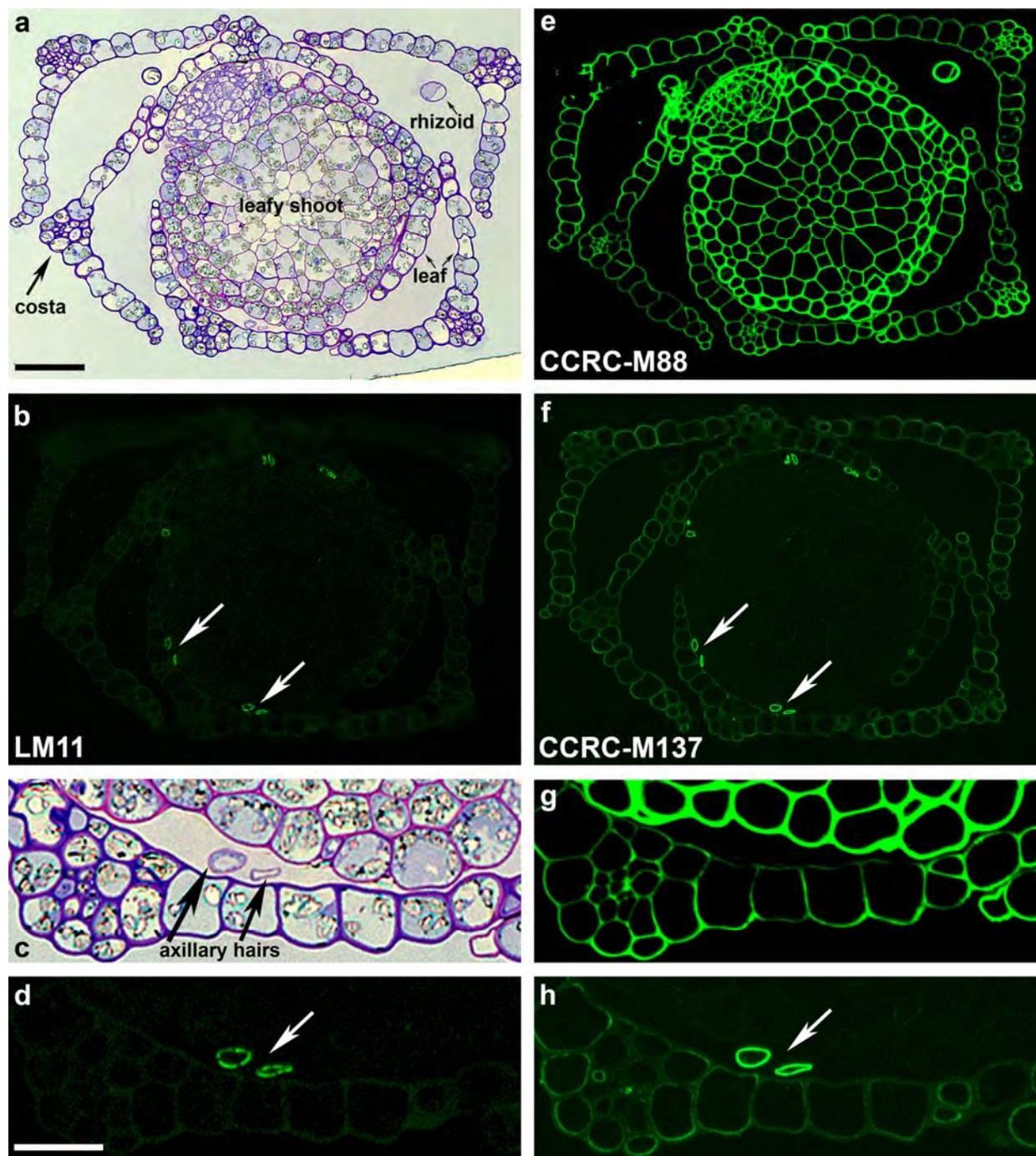
Key	Residue	H1	H2	H3	H4	H5 _{ax}	H5 _{eq}
A	α -Xylp (reducing)	5.184	3.545	3.783			
B	β -Xylp (reducing)	4.584	3.251	3.546	3.778	3.374	4.056
C	β -1,4-Xylp (internal) (major)	4.469	3.276	3.570	3.791	3.435	4.146
D	β -1,4-Xylp (internal) (minor)	4.477	3.288	3.55	3.79	3.37	4.103
E	β -Xylp (terminal)	4.457	3.254	3.427	3.623	3.300	3.975
H	β -1,2,4-Xylp (α -4Me- Glc pA)	4.625	3.435	3.620	3.803	3.382	4.102
I	α -4Me-Glc pA	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 III. ^1H 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 -Xylp (reducing)	5.183	3.543	3.77			
B	β -Xylp (reducing)	4.583	3.251	3.546	3.778	3.376	4.055
C	β -1,4-Xylp (internal) (major)	4.473	3.279	3.575	3.797	3.439	4.149
D	β -1,4-Xylp (internal) (minor)	4.470	3.282	3.56	3.79	3.37	4.10
E	β -Xylp (terminal)	4.457	3.257	3.427	3.622	3.301	3.970
F	β -1,2,4-Xylp (α -Glc pA)	4.643	3.480	3.635	3.810	3.390	4.109
G	α -Glc pA	5.305	3.552	3.738	3.470	4.363	---
H	β -1,2,4-Xylp (α -4Me- Glc pA)	4.627	3.436	3.623	3.803	3.383	4.102
I	α -4Me-Glc pA	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.

**Fig. 1**

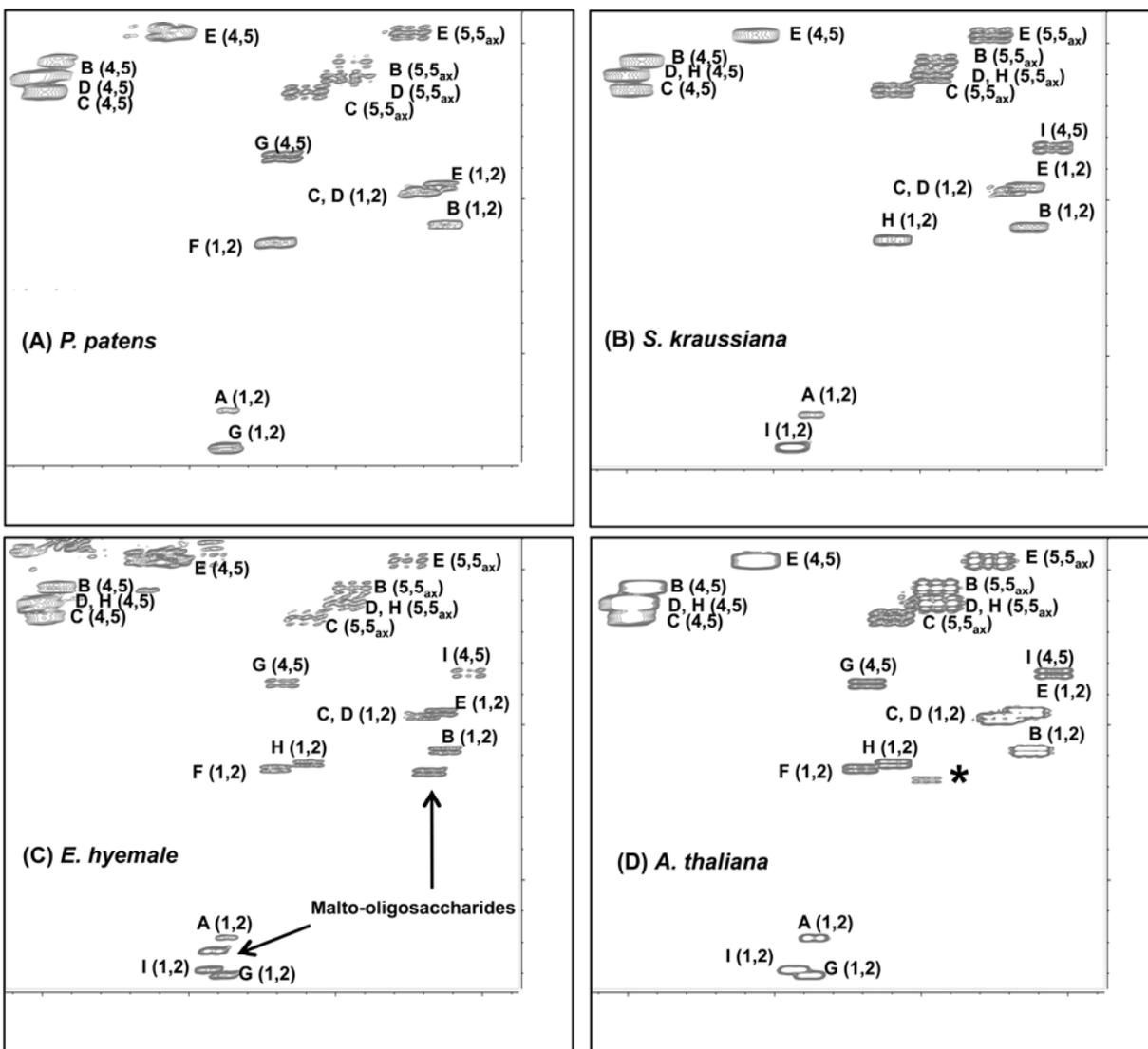


Fig. 2

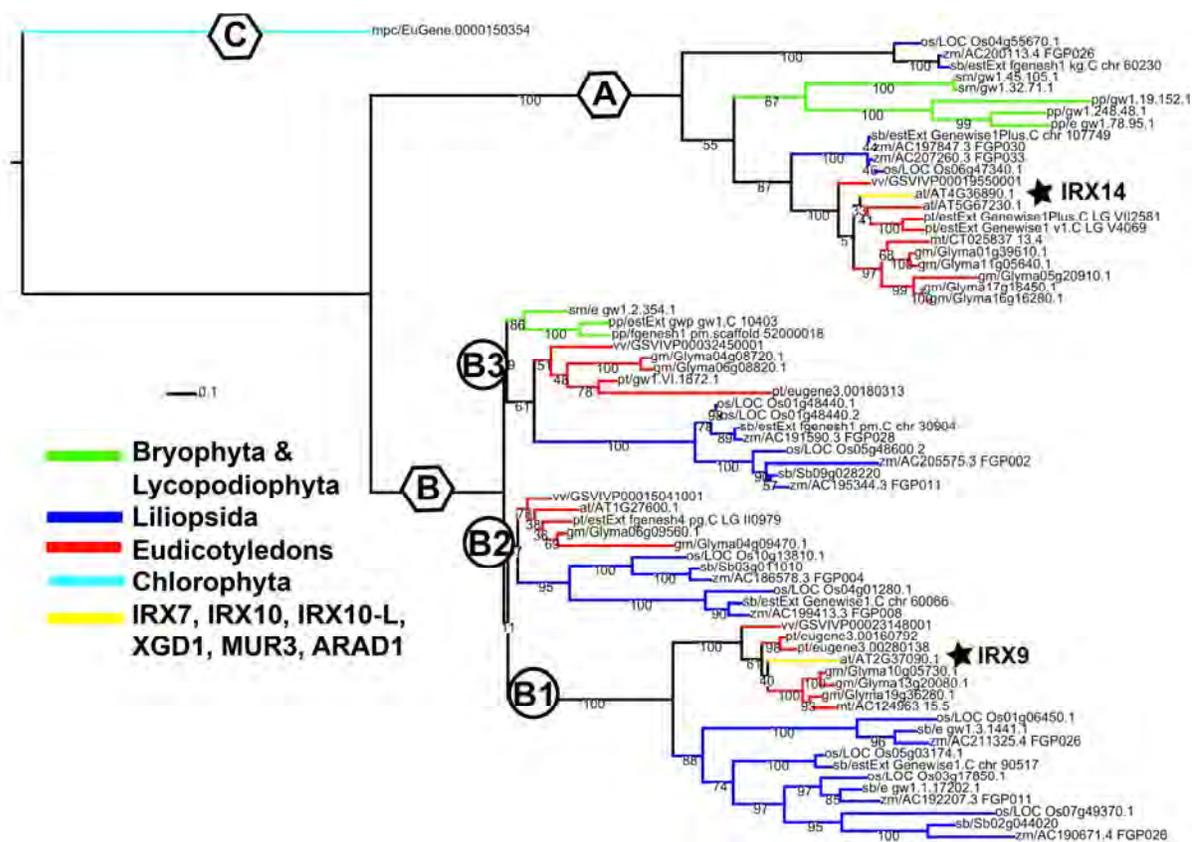


Fig. 4

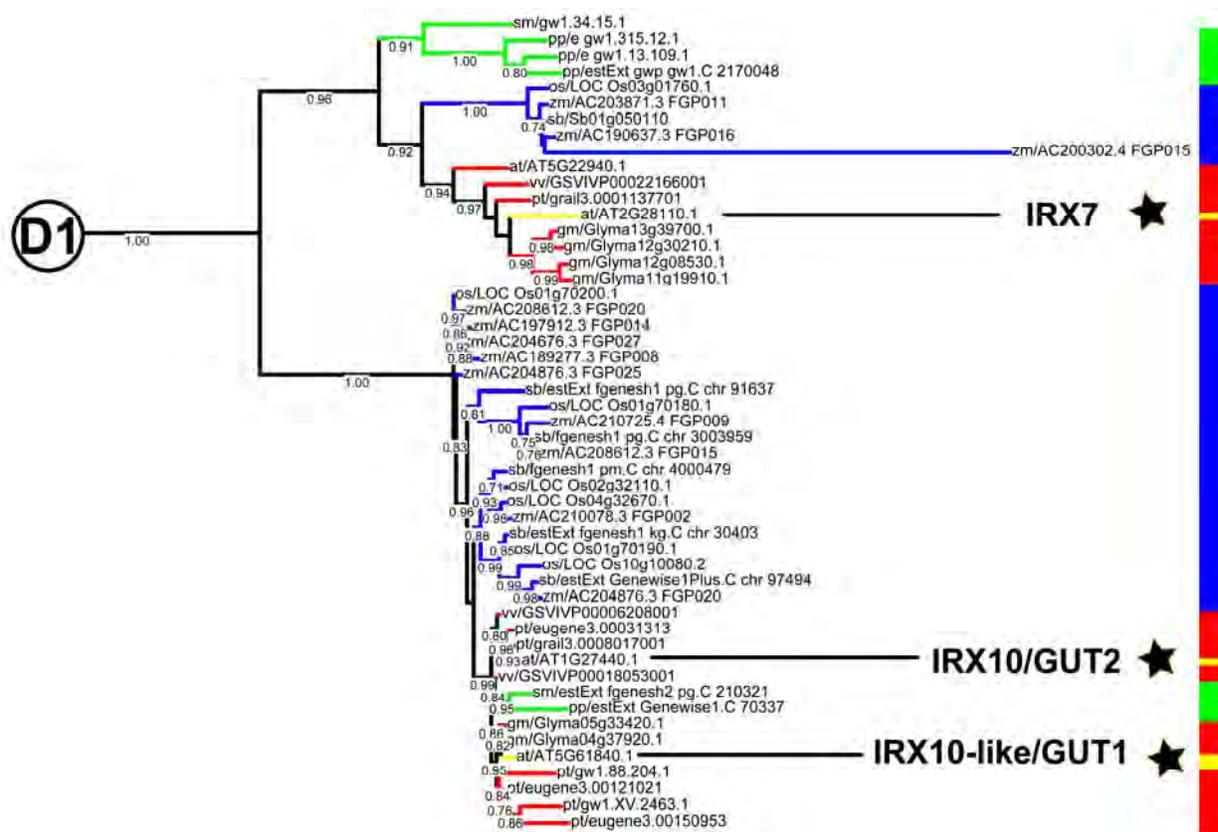


Fig. 5