Structural diversity of xylans in the cell walls of monocots

Maria J. Peña1 · Ameya R. Kulkarni1,4 · Jason Backe1 · Michael Boyd2 · Malcolm A. O’Neill1 · William S. York1,3

Received: 10 February 2016 / Accepted: 8 April 2016 © Springer-Verlag Berlin Heidelberg 2016

Abstract

Main conclusion  Xylans in the cell walls of monocots are structurally diverse. Arabinofuranose-containing glucuronoxylans are characteristic of commelinids. However, other structural features are not correlated with the major transitions in monocot evolution.

Most studies of xylan structure in monocot cell walls have emphasized members of the Poaceae (grasses). Thus, there is a paucity of information regarding xylan structure in other commelinid and in non-commelinid monocot walls. Here, we describe the major structural features of the xylans produced by plants selected from ten of the twelve monocot orders. Glucuronoxylans comparable to eudicot secondary wall glucuronoxylans are abundant in non-commelinid walls. However, the \( \alpha-D\)-glucuronic acid/4-O-methyl-\( \alpha-D\)-glucuronic acid is often substituted at O-2 by an \( \alpha-L\)-arabinopyranose residue in Alismatales and Asparagales glucuronoxylans. Glucuronoaarabinoxylans were the only xylans detected in the cell walls of five different members of the Poaceae family (grasses). By contrast, both glucuronoxylan and glucuronoaarabinoxylan are formed by the Zingiberales and Commelinales (commelinids). At least one species of each monocot order, including the Poales, forms xylan with the reducing end sequence \(-4)\beta-D-Xylp-(1,3)-\alpha-L-Rhap-(1,2)-\alpha-D-GalpA-(1,4)\alpha-D-Xyl\) first identified in eudicot and gymnosperm glucuronoxylans. This sequence was not discernible in the arabinopyranose-containing glucuronoxylans of the Alismatales and Asparagales or the glucuronoaarabinoxylans of the Poaceae. Rather, our data provide additional evidence that in Poaceae glucuronoaarabinoxylan, the reducing end xylose residue is often substituted at O-2 with 4-O-methyl glucuronic acid or at O-3 with arabinofuranose. The variations in xylan structure and their implications for the evolution and biosynthesis of monocot cell walls are discussed.

Keywords  Monocot · Cell wall · Glucuronoaarabinoxylan · Glucuronoxylan

Abbreviations

AGX  Arabinoglucuronoxylan
AIR  Alcohol-insoluble residue
Araf  Arabinofuranose
AraP  Arabinopyranose
2AB  2-Aminobenzamide
GlcP  Glucuronic acid
MeGlcP  4-O-Methyl glucuronic acid
\(^{13}\)C NMR  Carbon nuclear magnetic resonance spectroscopy
GAX  Glucuronoaarabinoxylan
GX  Glucuronoxylan

M. J. Peña and A. R. Kulkarni contributed equally to this study.

Electronic supplementary material  The online version of this article (doi:10.1007/s00425-016-2527-1) contains supplementary material, which is available to authorized users.

William S. York will@ccrc.uga.edu

1  Complex Carbohydrate Research Center and US Department of Energy Bioenergy Science Center, University of Georgia, Athens, GA 30602, USA

2  Department of Plant Biology, University of Georgia, Athens, GA 30602, USA

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

4  Present Address: Incyte Corporation, Wilmington, DE 19803, USA
**Introduction**

Plant cells are enclosed in a polysaccharide-rich macrostructure referred to as the cell wall. The composition of these walls varies considerably depending on the type and developmental stage of the cell that they surrounded (Fangel et al. 2012). Cells in growing tissues have non-lignified primary walls with a polysaccharide composition that differs from the lignified secondary walls of vascular and woody tissues. Additional diversity in cell wall composition also exists between different taxonomic groups, most notably in the monocots. This diverse group of land plants, which includes species of economic and ecological importance, is comprised of approximately 65,000 species that are distributed in 12 orders (Bremer et al. 2009; Givnish et al. 2010). Studies by Harris and Hartley showed that the monocots could be divided into two groups based on the presence or absence of ester-linked ferulic and coumaric acid in their primary walls (Harris and Hartley 1980). Monocots with primary walls that contain ferulate/coumarate comprise the commelinid clade (Chase et al. 1993). This clade consists of the Poales, which includes the Poaceae (grasses), Cyperaceae (sedges), and Bromeliaceae (bromeliads), together with the Zingiberales, Commelinales, and the basal Arecales (palms) (Givnish et al. 2010). The Dasypogonaceae are also considered to be commelinids but their relationship to other members of the clade has not been resolved (Chase et al. 2006). Ferulic acid esters are absent in the primary cell walls of the non-commelinid monocots, which are a paraphyletic group consisting of the Asparagales, Liliales, Pandanales, Dioscoreales, Petrosaviales, Alismatales, and Acorales. The Acorales are believed to be sister to all the monocots (Bremer et al. 2009; Givnish et al. 2010).

Commelinid and non-commelinid primary cell walls also differ in the type and abundance of their non-cellulosic polysaccharides. These differences led to the recognition of at least two types of primary wall (Bacic et al. 1988; Carpita and Gibeaut 1993; Carpita 1996; Harris et al. 1997). In the so-called type I primary walls of eudicots and non-commelinid monocots, pectin and xyloglucan account for the bulk of the non-cellulosic polysaccharides. Small amounts of xylan have also been reported to be present in these walls (Darvill et al. 1980; Zablackis et al. 1995; Hervé et al. 2009). By contrast, glucuronoxarabinoyxlan (GAX) is the predominant non-cellulosic polysaccharide in the type II primary walls of the Poaceae (commelinids), which contain relatively small amounts of pectin and xyloglucan (Smith and Harris 1999; Gibeaut et al. 2005). The walls of the Poaceae are further distinguished by the presence of variable amounts of (1-3,1-4)-linked β-glucans (Carpita and Gibeaut 1993; Smith and Harris 1999). The limited data available suggest that the primary cell walls of other commelinids have characteristics of both type I and type II walls.

Xylans are the major hemicellulosic polysaccharide present in the secondary (lignified) cell walls of flowering plants. These xylans are classified according to the type and abundance of the substituents on the 1,4-linked β-D-Xylp residues of the polysaccharide backbone (Fig. 1a). Glucuronoxylans (GX), which are major components of the secondary walls of woody and herbaceous eudicots, have α-D-glucuronic acid (GlcA) or 4-O-methyl α-D-glucuronic acid (MeGlcA) substituents at O-2 (Ebringerova et al. 2005). The secondary wall GAX of the Poaceae and other commelinids contains α-L-Araf residues at O-3, which account for the bulk of the backbone substituents, along with a small number of GlcP or MeGlcP substituents at O-2. The α-L-Araf residues may be further substituted at O-2 with a α-L-Araf or a β-D-Xylp residues (Saulnier et al. 1995). The GAX in the primary walls of Poaceae typically contains more Araf substituents. The Araf residues of GAX in Poaceae primary and secondary cell walls are often esterified with ferulic or coumaric acids (Buanañina 2009).

The GX and AGX in the secondary walls of eudicots and gymnosperms have a well-defined glycosyl sequence -4)-β-D-Xylp-(1,3)-α-L-Rhap-(1,2)-α-D-GalpA-(1,4)-β-D-Xylp (sequence I in Fig. 1b) at their reducing end (Johansson and Samuelson 1977; Peña et al. 2007). This sequence is required for normal GX synthesis during secondary cell wall formation and has been proposed to have a role in regulating xylan chain length (Peña et al. 2007; York and O’Neill 2008). However, no sequence I has been detected at reducing end of Poaceae GAX (Kulkarni et al. 2012a). Rather, in wheat endosperm GAX, the reducing end xylose has been reported to be often substituted at O-3 with an α-L-Araf residue or at O-2 with a α-D-GlcP residue (Ratnayake et al. 2014). It is not known whether sequence I is conserved in other commelinid and non-commelinid monocot heteroxylans.

Most studies of the composition and structure of monocot cell walls have emphasized model plants or plants of economic importance rather than phylogenetic diversity. Here, we describe the major structural features of xylans isolated from the cell walls of plants representing 10 of the 12 monocot orders. Cell walls were isolated from the entire actively growing aerial portions of each plant with a major
focus on species-dependent structural variations in polysaccharide structure. In selected cases, xylans were also isolated from specific organs to examine organ-specific differences. Our data support the view that the commelininids are the only monocots to synthesize xylan with Araf substituents. These GAXs predominate in most commelinid walls, with the exception of some members of the basal Arecales, which contain mainly GX. The walls of most of the non-commelinid contain GX. Our data also demonstrate that sequence 1 is more widespread in seed plants than anticipated as it is discernible in xylans from at least one species in each monocot order. Nevertheless, sequence 1 was not detected in any of the Poaceae GAXs analyzed. Rather, the xylosyl residues at the reducing end of these GAXs were often substituted with 4-O-Me-GlcpA or with Araf. Xylans lacking sequence 1 were present in some members of the Asparagales and the Alismatales. Many of these xylans contain GlcpA and/or MeGlcpA that is substituted with an α-L-Arap residue and may originate from the primary wall.

Materials and methods

Plant material and growth conditions

Actively growing stems and leaves from Acorus americanus (American sweet flag, Acorales), Elodea canadensis (Canadian waterweed, Alismatales), Dioscorea alata (purple yam, Dioscoreales), Pandanus utilis (common screwpine, Pandanales), Agapanthus africanus (African lily, Asparagales), Asparagus officinalis (asparagus, Asparagales), Clivia miniata (natal lily, Asparagales), Crinum americanum (southern swamp lily, Asparagales), Iris germanica (German iris, Asparagales), Ludisia discolor (jewel orchid, Asparagales), Allium cepa (onion, Asparagales), Agave americana (agave, Asparagales), Caryota mitis (fishtail palm, Arecales), Howea forsteriana (kentia palm, Arecales), Rhapis excelsa (broadleaf lady palm, Arecales), Sabal etonia (scrub palmetto, Arecales), Cocos nucifera (coconut, Arecales), Tradescantia virginiana (spiderwort, Commelininales), Amomum costatum (cardamom, Zingiberales), Zingiber officinale (ginger, Zingiberales), Hedychium coronarium (white ginger lily, Zingiberales), Strelitzia alba (white bird of paradise, Zingiberales), Calathea lancifolia (rattlesnake plant, Zingiberales), Ctenanthe oppenheimeriana (never–never plant, Zingiberales), Musa sp. (banana, Zingiberales), Tillandsia usneoides (Spanish moss, Bromeliaceae), and Cyperus alternifolius (umbrella papyrus, Cyperaceae) were obtained from the Plant Biology Greenhouse at the University of Georgia (Athens, GA). *Lemma minor* (CPCC 490, Landolt clone 8434, Alismatales) was obtained from the Canadian Phycological Culture Centre (University of Waterloo, Canada). *Spir- odea polyrhiza* (ID7498) was obtained from Joachim
Messing (Waksman Institute of Microbiology, Rutgers University, New Jersey, United States of America). Duckweeds were grown on Schenck and Hildebrandt basal salts (1.6 g L\(^{-1}\)), pH 5.8, containing sucrose (10 g L\(^{-1}\)), morpholinoethanesulfonic acid (0.5 g L\(^{-1}\)), and agar (8 g L\(^{-1}\)) in a controlled environmental growth chamber (Adaptis A1000, Conviron, Winnipeg, Canada) at 19 and 15 °C with a 14-h light and 10-h dark cycle, respectively, with a light intensity of 120 μmol quanta m\(^{-2}\) s\(^{-1}\), and 70 % relative humidity.

Stems and leaves of mature Miscanthus x giganteus (giant miscanthus), Brachypodium distachyon, and Setaria italica (Poaceae) were obtained from the Plant Biology Greenhouse at the University of Georgia (Athens, GA). Oryza sativa (rice, Poaceae) straw was obtained from Wayne Parrot at the University of Georgia. The air-dried stems and leaves were Wiley-milled (−20/+80 mesh) and kept at room temperature. Panicum virgatum (switchgrass, Poaceae) was obtained from the Samuel Roberts Noble Foundation (Ardmore, Oklahoma, USA). The biomass was milled using a Hammer mill with a 1-mm screen and then ground in a Wiley mill using a 1-mm screen. The milled material was sieved to −120 μmol quanta m\(^{-2}\) s\(^{-1}\), and 70 % relative humidity.

Preparation of cell walls

Cell walls were prepared as their alcohol-insoluble residues (AIRs) by homogenizing suspensions of the plant material in aq. 80 % (v/v) EtOH with a Polytron tissue disruptor (Kinematica, Switzerland). To facilitate labeling of grass cell walls with 2-aminobenzamide (2AB), the homogenized material was ball-milled for 16 h at 4 °C and 90–100 rpm in aq. 80 % EtOH (v/v) using ¼ in. zirconium grinding media (U.S. Stoneware, East Palestine, OH). This procedure reduced the AIR particle size to <50 μm. The AIRs were then washed with absolute EtOH, with MeOH:CHCl\(_3\) (1: 1 v/v), and with acetone and then vacuum-dried at room temperature.

2-Aminobenzamide labeling of the reducing sugars present in AIR

The ball-milled AIR (100–200 mg) from the five grasses was suspended in DMSO: H\(_2\)OAc (70:30 v/v) containing sodium cyanoborohydride (1 M) and 2AB (0.5 M) and kept for 16 h at 65 °C (Bigge et al. 1995). The suspensions were then dialyzed against deionized water (3500 molecular weight cut-off tubing) and lyophilized.

Sequential extraction of AIR

The AIR and the 2AB-labeled AIR (0.5 g) were sequentially extracted with 50 mM ammonium oxalate (35 mL) and then with 1 M KOH containing 1 % (w/v) NaBH\(_4\). The 1 M KOH-soluble extracts were neutralized with HOAc, dialyzed (3500 MWCO tubing) against deionized water, and then lyophilized.

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

The 1 M KOH-soluble materials (~20 mg) were suspended in 50 mM ammonium formate, pH 5, and treated for 24 h at 37 °C with endoxyylanase (3.5 unit, Trichoderma viride M1, Megazyme, Wicklow, Ireland). 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 purified by size exclusion chromatography (SEC) using a Superdex-75 HR10/30 column (GE Healthcare) and a Dionex Ultimate 3000 HPLC equipped with a Shodex RI-101 refractive index detector and a fluorescence detector. The column was eluted with 50 mM ammonium formate, pH 5, and the fraction containing xylo-oligosaccharides were collected and freeze-dried. The xylo-oligosaccharides were analyzed by MALDI-TOF MS spectrometry and NMR spectroscopy.

The 2AB-labeled xylo-oligosaccharides generated by enzymatic digestion of the 1 M KOH-soluble materials from the 2AB-labeled AIR were partially purified by SEC as described above. Acidic and neutral 2AB-labeled xylo-oligosaccharides were separated by solid-phase extraction (SPE) using graphitized carbon cartridges (Packer et al. 1998). The oligosaccharides were eluted with a series of increasing concentrations of acetonitrile (10 % (v/v) acetonitrile to 80 % (v/v) acetonitrile). Most of the acidic 2AB-labeled xylo-oligosaccharides were eluted with 40 % (v/v) acetonitrile.

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS)

Positive ion MALDI-TOF mass spectra were recorded using a Bruker LT MALDI-TOF mass spectrometer interfaced to a Bruker Biospectrometry workstation (Bruker...

Electrospray ionization mass spectrometry (ESI MS)

Materials enriched in acidic or neutral 2AB-labeled xylo-oligosaccharides (~1 mg) were dissolved in dry DMSO (0.2 mL) and per-O-methylated and isolated as described (Mazumder and York 2010). The per-O-methylated xylo-oligosaccharides were then analyzed by positive ion ESI MS using a ThermoScientific LTQ XL mass spectrometer (Thermo Scientific, http://www.thermoscientific.com) as described (Mazumder and York 2010).

NMR spectroscopy

Native and 2AB-labeled xylo-oligosaccharides were dissolved in D$_2$O (3–5 mg in 0.3 mL, 99.9 %; Cambridge Isotope Laboratories). One- and two-dimensional NMR spectra were recorded at 298 K with a Varian Inova NMR spectrometer operating at 600 MHz and equipped with a 5-mm NMR cold probe (Agilent, Santa Clara, CA). Homonuclear (gCOSY, TOCSY, and ROESY) and the heteronuclear (gHSQC and gHMBC) experiments were recorded using standard Varian pulse programs. Chemical shifts were measured relative to internal acetone at $\delta_H$ 2.225 and $\delta_C$ 30.89. Data were processed using MestReNova software (Universidad de Santiago de Compostela, Spain).

Results

The cell walls of commelinid and non-commelinid monocots contain structurally distinct heteroxylans

To determine the species-dependent differences in xylan structure, cell walls were isolated from the entire actively growing aerial portions of plants representing 10 (Poales, Zingiberales, Commelinales, Arecales, Asparagales, Liliales, Pandanales, Dioscoreales, Alismatales, and Acorales) of the 12 orders of monocots. Material enriched in xyloglucan was then solubilized by treating the depectinated AIR with 1 M KOH. The xyloans were fragmented with endoxylanase and the resulting xylo-oligosaccharides isolated by SEC and structurally characterized using mass spectrometry and NMR spectroscopy. MALDI-TOF MS provides information regarding the number and identity of the glycosyl residues (pentose, hexose, or hexuronic acid) in each oligosaccharide. 1D and 2D $^1$H NMR spectroscopy of the oligosaccharides provides information about the glycosyl composition, linkage, and sequence of each oligosaccharide. Assignment of NMR resonances was facilitated by comparison with previously published chemical shift data for the $^1$H and $^{13}$C resonances of the residues in oligosaccharides prepared from structurally diverse heteroxylans (Peña et al. 2007; Kulkarni et al. 2012a).

The commelinid clade is resolved into two groups on the basis of their MALDI-TOF MS xylo-oligosaccharide profiles. In the first group, which includes Panicum virgatum, Brachypodium distachyon, Oryza sativa, Miscanthus × giganteus, and Setaria italica (Poaceae), Calathea lancifolia and Ctenanthe opffenheimiana (Marantaceae), and Amomum costatum and Zingiber officinalis (Zingiberaceae) (see Fig. 2, Supplementary Fig. S1), the spectra are dominated by [M + Na]$^+$ ions (m/z 701, 833, 965, and 1097), corresponding to oligosaccharides containing from 5 to 8 pentosyl residues. Ions originating from oligosaccharides containing a MeGlcP A residue (e.g., P$_4$M at m/z 759 and P$_5$M at m/z 891) were present at much lower abundance (Fig. 2, Supplementary Fig. S1). These types and abundance of neutral and acidic oligosaccharides are the expected products of xylanase fragmentation of GAX that contain only a few MeGlcP residues. Such GAXs are the predominant hemicellulose in the cell walls of the Poaceae (Poales) (Ebringerova et al. 2005). The MALDI-TOF MS spectra of the xylo-oligosaccharides from the second group of commelinids (Fig. 2, Supplementary Fig. S1), which includes Cyperus alternifolius (Cyperaceae), Ananas Comosus and Tillandsia usneoides (Bromeliaceae), Musa sp. (Musaceae), Tradescantia virginiana (Commelinaceae), and Sabal etonia, Cocos Nucifera, Caryota mitis, and Rhapis excelsa (Arecaceae), were dominated by [M + Na]$^+$ ions (m/z 759, 891, 1023, and 1155) and [M – H + 2Na]$^+$ ions (m/z 781, 913, 1045, 1177), which correspond to oligosaccharides composed of a single MeGlcP residue and from 4 to 7 pentosyl residues (P$_4$M–P$_7$M).

The MALDI-TOF MS xylo-oligosaccharide profiles of all the non-commelinids contained ions that correspond to acidic oligosaccharides composed of 4-6 xylosyl residues and one uronic acid (Fig. 2, Supplementary Fig. S1). Typically, most of the GlcpA residues were 4-O-methylated. However, xylo-oligosaccharides from Crinum americanum (Amaryllidaceae), Ludisia discolor (Orchidaceae), Iris germanica (Iridaceae), Orotium aquaticum and Spirodela polyrhiza (Araceae) contained substantial amounts of GlcpA lacking a methyl group (Fig. 2, Supplementary Fig. S1).

The ion distribution in most of the MALDI-TOF mass spectra is consistent with the products expected to be formed when a GX is fragmented with a family GH11 xylanase. However, the spectra of the xylo-oligosaccharides of several members of the Asparagales including...
**Ludisia discolor** (Orchidaceae), **Crinum americanum** and **Clivia miniata** (Amaryllidaceae) did not conform to this distribution as they were enriched in ions that correspond to oligosaccharides composed of six pentoses and a uronic acid (Fig. 2, Supplementary Fig. S1). The spectrum of the **Clivia miniata** (Amaryllidaceae) xylo-oligosaccharides also indicates that this plant produces GX with a higher Me-GlcAp to GlcpA ratio than those produced by other Asparagales. The xylo-oligosaccharides from two Araceae (**Orontium aquaticum** (Fig. 2) and **Elodea canadensis** (Supplementary Fig. S1), also contained elevated amounts of oligosaccharides containing six pentoses and a uronic acid. Such oligosaccharides are typically generated by the family GH11 endoxylanase when acting...
on xylans with side chains that contain pentosyl residues. To extend the results of these MS analyses, we used 1D and 2D NMR spectroscopy to gain further insight into the structures of the side chain substituents and diverse pentose-containing side chain structures generated from the commelinid and non-commelinid xylans.

NMR spectroscopy provides further insight into the structural features characteristic of commelinid and non-commelinid xylans

The 1D and 2D NMR spectroscopic analyses of the xylo-oligosaccharides allowed us to obtain information on the structures and relative abundances of the side chain sub-

oligosaccharides allowed us to obtain information on the structures and relative abundances of the side chain sub-

stituents in the xylans from each monocot species examined.

Resonances characteristic of GlcP A substituents were observed in the 1H-NMR spectra of the xylo-oligosaccharides generated from commelinid and non-commelinid monocot heteroxylans (Fig. 3, Supplementary Figs. S2–5). These GlcP A residues were typically 4-O-methylated. The extent of GlcP A methylation in xylans is known to differ depending on the species, the tissue, and the developmental stage. For example, virtually all of the GlcP A is 4-O-methylated in woody dicot GX, whereas in Arabidopsis GX the degree of 4-O-methylation varies from 40 to 70 % depending on the plant’s developmental stage. None of the GlcP A is methylated in the GX of the moss Physcomitrella patens (Kulkarni et al. 2012b).

The GlcA residues in all the monocot xylans analyzed in this study were 4-O-methylated to some extent, with degrees of methylation that varied from 40 to 100 % (Table 1). GlcP A- and MeGlcP A-containing side chains were the only backbone substituents detected in GX from the non-commelinid monocots (Table 1). The xylans from all the commelinids analyzed also contained Araf-substituents at O-3 of many of the backbone xylosyl residues. These Araf residues are the most abundant backbone substituent of the GAX in the Poaceae (Table 1; Fig. 3, Supplementary Fig. S2). We also detected resonances corresponding to Araf bearing a substituent at O-2 (i.e., 2-substituted Araf) in the spectra of all the Poales analyzed (Fig. 1a; Table 1; Supplementary Fig. S2), but not in the spectra of the other commelinids (Supplementary Fig. S3). Thus, 2-substituted Araf may be a unique characteristic of Poales GAX.

Our NMR and MALDI-TOF MS data are consistent with the notion that Araf substituents are one of the structural motifs that distinguishes commelinid xylans from their non-commelinid counterparts (Table 1). The Arecales, which are the basal commelinids, are comprised of only one family, the Arecales. The xylans of the Arecales (Sabal etonia, Cocos nucifera, and Howea forsteriana) contain much more MeGlcP A than Araf (Fig. 3, Supplementary Fig. S3; Table 1). Thus, the Arecales produce xylans that have structural homology with non-commelinid GX including the presence of sequence 1 at the reducing end. We also detected 1 in all the Zingiberales and Commelinales xylans analyzed (Table 1, Supplementary Fig. S3). Previously, it was reported that 1 is absent in the GAX of the Poales (Kulkarni et al. 2012a; Ratnayake et al. 2014). Our analysis indicates that 1 is also absent in the GAX of other Poales including Cyperus alternifolius (Cyperaceae) and Tillandsia usneoides (Bromeliaceae) (Table 1). Weak, but clearly discernible resonances diagnostic for the GalpA and Rhap residues of 1 were present in the 1H-NMR spectrum of the xylo-oligosaccharides of Ananas comosus (pineapple, Bromeliaceae) (Table 1; Fig. 3). Thus, 1 was detected in at least one species of each order in the commelinid clade (Table 1). These results indicate that 1 is far more widespread among the monocots than we anticipated. Nevertheless, 1 was not detected in all monocot xylans. For example, no resonances characteristic of 1 were discernible in the NMR spectra of the xylo-oligosaccharides from several non-commelinid monocots (Table 1; Fig. 3, Supplementary Figs. S4 and S5), including Spirodela polyrhiza and Lemma minor (Araceae, Alismatales) and Agapanthus africanus and Allium cepa (Amaryllidaceae, Asparagales).

Structural characterization of the α-1-Arap-(1 → 2)-α-D-GlcP A-(1 → side chain of Alismatales and Asparagales Xylan

The 1D 1H-NMR spectra of xylo-oligosaccharides generated from the xylan of several Asparagales and Alismatales species (Fig. 3, Supplementary Figs. S4 and S5) contain strong signals at δ 5.41. These signals are diagnostic for GlcP A or MeGlcP A residues bearing a substituent at O-2 (Shatalov et al. 1999; Togashi et al. 2009; Zhong et al. 2014). Signals at δ 5.41 may also originate from starch, which are often a contaminant of isolated cell walls. Nevertheless, our 2D NMR analysis established that these signals did not originate from starch-derived oligosaccharides.

To unambiguously identify the substituted GlcpA/ MeGlcpA residues, we recorded high-resolution gCOSY spectra of the purified xylo-oligosaccharides from Spirodel a polyrhiza (Alismatales, Fig. 4a) and Crinum americanum (Asparagales, Fig. 4b). Analysis of these gCOSY spectra allowed us to identify and assign the resonances of isolated spin systems with chemical shifts and coupling patterns characteristic of 2-substituted GlcpA for Spirodel a polyrhiza (Fig. 4a) or 2-substituted GlcpA and MeGlcpA for Crinum americanum (Fig. 4b). The remaining 1H and 13C resonances of the xylo-oligosaccharide residues were
assigned using homonuclear ROESY and heteronuclear (gHSQC and gHMBC) 2D NMR experiments (Table 2; Figs. 4, 5). These analyses established that the substituted GlcAp is attached to O-2 of a 4-linked Xylp. For example, in the Crinum americanum gHMBC spectrum, the cross-peak between C-1 (δ 97.3) of the 2-linked GlcAp and H-2 (δ 3.56) of the 2,4-linked Xylp confirmed that the GlcAp residue is linked to O-2 of the Xylp residue (Table 2; Fig. 5b).

The results of several studies have indicated that GlcAp is substituted at O-2 with a hexose (α-Galp) or with a partially characterized pentose in some dicots GXS (Shatalov et al. 1999; Togashi et al. 2009; Zhong et al. 2014; Chong et al. 2015; Mortimer et al. 2015). Our MALDI-TOF mass spectra (Fig. 2) are consistent with the presence of xylo-oligosaccharides containing an additional pentose as they contain a high abundance of ions corresponding to oligosaccharides with six pentoses and one uronic acid. No
substantial ions corresponding to xylo-oligosaccharides containing a hexosyl residue were detected, suggesting that oligosaccharides containing galactose linked to GlcA are of low abundance or completely absent in non-commelinid monocot GX.

\[ \beta-D-Galp \text{ and } \alpha-L-Arap \text{ residues are structural homologs that are difficult to distinguish by } { }^1H \text{ NMR spectroscopy but can be differentiated by } { }^{13}C \text{ NMR spectroscopy. Indeed, the } { }^{13}C \text{ NMR spectra of the } Spirodela polyrhiza \text{ and } Crinum americanum \text{ xylo-oligosaccharides contain signals with chemical shifts for C-1 (} \delta 105.2 \text{) and C-5 (} \delta 67.1 \text{) that are diagnostic of Arap residues (Table 2; Figs. 4, 5a) (Glushka et al. 2003; Peña et al. 2008). The gHMBC spectrum (Fig. 5b) contains a cross-peak between C-1 (} \delta 105.2 \text{) of the Arap residue and H-2 (} \delta 3.77 \text{) of the 2-linked GlcA residue that shows that the Arap is attached to O-2 of the GlcA. The presence of this } \alpha-L-Arap-(1 \rightarrow 2)-\beta-D-GlcA-(1 \rightarrow \text{ side chain is consistent with our MALDI analysis of the xylan from selected Asparagales and Alismatales. Further NMR analysis established that } Crinum americanum \text{ xylan contains the structurally related side chain } \alpha-

Table 1  The major structural motifs of monocot xylans

<table>
<thead>
<tr>
<th>Order (family)</th>
<th>Species</th>
<th>Seq. 1</th>
<th>Side chain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Terminal MeGlcA</td>
</tr>
<tr>
<td>Commelinids</td>
<td></td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>Poales (Poaceae)</td>
<td>Brachypodium distachyon</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Panicum virgatum</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Oryza sativa</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Miscanthus × giganteus</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>Poales (Cyperaceae)</td>
<td>Cyperus alternifolius</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>Poales (Bromeliaceae)</td>
<td>Ananas comosus</td>
<td></td>
<td>tr</td>
</tr>
<tr>
<td></td>
<td>Tillandsia usneoides</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>Zingiberales (Zingiberae)</td>
<td>Amomum costatum</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Hedychium coronarium</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Zingiberales (Strelitziaceae)</td>
<td>Strelitzia alba</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Commelinales</td>
<td>Tradescantia virginiana</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>(Commelinacea)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arecales (Arecales)</td>
<td>Sabal etonia</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Cocos nucifera</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Howea forsteriana</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Non-Commelinids</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asparagales (Amaryllidaceae)</td>
<td>Crinum americanum</td>
<td></td>
<td>tr</td>
</tr>
<tr>
<td></td>
<td>Agapanthus africanus</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Allium cepa</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>Asparagales (Asparagaceae)</td>
<td>Asparagus officinalis (Tip)</td>
<td></td>
<td>tr</td>
</tr>
<tr>
<td></td>
<td>Asparagus officinalis (Stem)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Agave americana</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Liliales (Alstroemeriaceae)</td>
<td>Alstroemeria sp.</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Liales (Liliaeae)</td>
<td>Tulipa sp.</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Pandales (Pandanaceae)</td>
<td>Pandanus utilis</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Dioscoreales (Dioscoreaceae)</td>
<td>Dioscorea alata</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Alismatales (Araceae)</td>
<td>Oronium aquaticum</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Spirodela polyrhiza</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Lemna minor</td>
<td></td>
<td>ND</td>
</tr>
<tr>
<td>Acorales (Acoraceae)</td>
<td>Acorus americanus</td>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

ND not detected, tr trace amount (<1 %)
**Table 2** $^1$H and $^13$C NMR assignments for the side chains containing GlcA/MeGlcA substituted with $\alpha-L$-Ara residues present in Asparagales, Alismatales, and *Eucalyptus grandis* GX

<table>
<thead>
<tr>
<th>Residue</th>
<th>$^1$H/C1</th>
<th>$^1$H/C2</th>
<th>$^1$H/C3</th>
<th>$^1$H/C4</th>
<th>$^1$H/C5</th>
<th>$^1$H/C6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$-L-Ara</td>
<td>4.60/105.4</td>
<td>3.62/71.6</td>
<td>3.68/72</td>
<td>3.94/69</td>
<td>3.93–3.67/67.2</td>
<td></td>
</tr>
<tr>
<td>2-$\alpha$-GlcP A</td>
<td>5.42/97.3</td>
<td>3.76/79.3</td>
<td>3.94/72.4</td>
<td>3.53/72.4</td>
<td>4.38/72.5</td>
<td></td>
</tr>
<tr>
<td>2-$\alpha$-Glc</td>
<td>4.71/101.2</td>
<td>3.55/76.2</td>
<td>3.69/72.4</td>
<td>3.81/77.0</td>
<td>4.13–3.44/63.1</td>
<td></td>
</tr>
<tr>
<td>$\alpha$-L-Ara</td>
<td>4.61/105.3</td>
<td>3.61/71.7</td>
<td>3.68/73</td>
<td>3.94/69</td>
<td>3.93–3.67/67.2</td>
<td></td>
</tr>
<tr>
<td>2-$\alpha$-GlcP A</td>
<td>5.41/97.1</td>
<td>3.79/79.3</td>
<td>3.97/72.2</td>
<td>3.29/82.9</td>
<td>4.35/72</td>
<td></td>
</tr>
<tr>
<td>2-$\alpha$-Xyl</td>
<td>4.69/101.3</td>
<td>3.51/76.1</td>
<td>3.66/72.6</td>
<td>3.81/77.0</td>
<td>4.10–3.43/63</td>
<td></td>
</tr>
<tr>
<td>$\beta$-GlcP A</td>
<td>4.70/104.7</td>
<td>3.58/76.7</td>
<td>3.65/72.6</td>
<td>3.80/77</td>
<td>4.10–3.40/63.3</td>
<td></td>
</tr>
<tr>
<td>2-$\beta$-Xyl</td>
<td>4.67/101.6</td>
<td>3.51/76</td>
<td>3.66/72</td>
<td>3.80/77</td>
<td>4.10–3.40/63</td>
<td></td>
</tr>
</tbody>
</table>

The corresponding assignments for the $\beta$-Galp-(1 → 2)-MeGlcP A side chain in *Eucalyptus* GX are also shown.

* Chemical shifts are reported in ppm relative to internal acetone, $\delta_H$ 2.225 and $\delta_C$ 30.89

$\alpha$-L-Ara-(1 → 2)-4-O- Me-$\alpha$-d-GlcP A-(1 → 2) - (Table 2; Figs. 4, 5). This disaccharide is the most abundant side chain in the xylan from several Asparagales species, including *Crinum americanum*, *Agapanthus africanus*, *Clivia miniata*, *Allium cepa* (Amaryllidaceae), and *Ludisia discolor* (Orchidaceae) and in the Alismatales, *Spirodela polyrhiza*, *Orontium aquaticum*, and *Elodea canadensis* (Table 1; Fig. 2, Supplementary Figs. S4 and S5). Weak signals that are diagnostic for an $\alpha$-L-Ara-containing acidic side chain were also detected in the $^1$H-NMR spectra of xylo-oligosaccharides of numerous other monocots (see Table 1).

The xylo-oligosaccharides generated from *Eucalyptus globulus* wood GX have been shown to contain a $\beta$-Galp-.
Thus, we next determined using NMR if Eucalyptus xylan also has Arap linked to GlcpA. Resonances diagnostic for both the β-Galp (1 → 2)-4-O-methyl-α-GlcpA-(1 → 2)-side chain are present in the gCOSY spectra of Eucalyptus grandis xylo-oligosaccharides (Supplementary Fig. S6). Together, these data show that GX containing the Arap-GlcA side chain is a common feature of non-commelinid monocots and provide further evidence that it is present in dicot GX (Chong et al. 2015; Mortimer et al. 2015).

**Xylan structure shows organ-specific variations in selected monocots genera**

The cell wall xylans of commelinids, other than the Poales, have structural features that are comparable to the GAX of Poaceae walls (single Araf side chains) and to the GX of eudicot secondary walls (high abundance of MeGlcA and GlcpA and presence of sequence 1) (see Table 1). To determine whether these different structural features are predominantly associated with primary or with secondary wall xylans, the xylans isolated from leaves and stems of Tradescantia virginiata (family Commelinaceae, Commelininales) were structurally characterized. Acidic xylo-oligosaccharides that are characteristic of GX in eudicot secondary walls were predominant in the MALDI-TOF MS of the stem xylo-oligosaccharides (Fig. 6a). By contrast, neutral xylo-oligosaccharides were more abundant in leaves, which are rich in primary walls. Thus, the Tradescantia virginiata leaf xylan is comparable with the Poaceae GAX where neutral Araf-containing substituents predominate.

In a second series of experiments, we found that the xylans isolated from Asparagus officinalis organs enriched in either primary or secondary walls are also structurally distinct. For example, the GX from growing tips of shoots contains the Arap-(1,2)-GlcpA-(1, side chain but only trace amounts of 1 were discernible. The stem and “leaf” GXs had features characteristic of eudicot secondary wall GX, including the presence of sequence 1 at the reducing end (Fig. 6b). Although this result at first sight was surprising, it is consistent with the fact that in Asparagus what appear to be the leaves are in fact modified stems (cladodes or phylloclades); the true leaves are scale-like structures on the stem.

**The reducing end xylose of Poaceae secondary wall xylan is often substituted at O-2 with MeGlcA or at O-3 with Araf**

Our results and previous studies have shown that GAX isolated from secondary cell walls of Poaceae lacks...
discernible amounts of 1 at their reducing end (Kulkarni et al. 2012a). Ratnayake and colleagues reported that the reducing end xylose of *Triticum aestivum* (wheat) endosperm GAX is often substituted at O-3 with an α-L-Ara residue or at O-2 with an α-D-GlcP A residue (Ratnayake et al. 2014). To extend these studies and to identify the glycosyl sequence at the reducing end of the GAX in the secondary cell walls of Poaceae, we used the following approach. The AIRs of the five Poaceae were treated with 2AB in the presence of sodium cyanoborohydride. This reductive amination procedure converts any exposed reducing end glycose to its corresponding 2AB-labeled derivative. The GAX was then solubilized from the AIR and fragmented with endoxylanase. Xylo-oligosaccharides containing xylitol-2AB correspond to the sequence of glycoses that were located at the reducing end of the intact polysaccharide. The 2AB-labeled xylo-oligosaccharides were purified and separated into neutral and acidic fractions using a combination of SEC and solid-phase extraction with graphitized carbon and then structurally characterized using MALDI-TOF MS, ESI MS^n, and 1D and 2D NMR spectroscopy.

Complete chemical shift assignments of the glycosyl residues comprising the most abundant components of the mixture of 2AB-labeled acidic xylo-oligosaccharides of *Setaria italica* (foxtail millet) were obtained using 1D and 2D (TOCSY, HSQC, ROESY, HMBC) NMR experiments (Fig. 7; Table 3). The predominant 2AB-labeled oligoglycosyl alditol contained one MeGlcP A residue, two xylosyl residues, and a 2AB-derivatized xylitol. The chemical shift values for 2AB-labeled xylitol are, with the exception of the C-2 resonance, consistent with previously published
2AB-xylitol showing that a MeGlcA residue is linked to O-2 of the xylitol (Fig. 7; Table 3). The gHMBC spectrum also contained a cross-peak correlating H-1 of an internal 4-linked β-D-Xylp residue with C-4 of the xylitol-2AB residue, indicating that this β-D-Xylp residue is linked to O-4 of the xylitol AB. Finally, a cross-peak correlated H-1 of a terminal β-D-Xylp residue with C-4 of the internal 4-linked β-D-Xylp residue, which itself is linked to O-4 of the xylitol-2AB residue (Fig. 7; Table 3). These results together show that the predominant oligosaccharide in the mixture is β-D-Xylp-(1→4)-β-D-Xylp-(1→4)-[MeGlcA-(1→2)]-xylitol-2AB (Fig. 1b). Other weaker resonances observed in the 2D spectra indicate that small amounts of Xyl-(Ara)-Xyl-(MeGlcA)-Xyl-2AB and Xyl-Xyl-(Ara)-Xyl-(MeGlcA)-Xyl-2AB were also present in the Setaria italica GAX. Thus, the reducing end xylene residue of this GAX is frequently substituted at O-2 with a MeGlcA residue.

Our NMR analysis of the 2AB-labeled neutral xylo-oligosaccharides from Setaria italica indicated a high abundance of Araf side chains, but the mixture was too complex to be fully characterized by this technique. Thus, the neutral and acidic 2AB-labeled xylo-oligosaccharides were per-O-methylated and analyzed by ESI MS. The ESI MS spectra of these oligosaccharides are dominated by Y and B ions, which provide extensive sequence information (Domon and Costello 1988; Mazumder and York 2010). Y ions, which correspond to fragments generated by loss of residues from the non-reducing end, were abundant due to the presence of the 2AB-labeled xylitol at the former reducing end of the oligosaccharides. B ions are fragments generated by loss of residues from the reducing terminus and were much less abundant in the ESI spectra.

The full-scan ESI mass spectrum of the per-O-methylated acidic oligosaccharides from Setaria italica contained a series of [M + Na]+ ions at m/z 769, 929, 1089, 1249, which correspond to 2AB-labeled oligosaccharides containing a 2AB-labeled xylitol, one to four pentosyl residues, and a single uronic acid (Supplementary Fig. S7a). Based on our NMR analysis, the uronic acid is almost always MeGlcA. The parallel series of ions at 14 Da lower mass likely corresponds to oligosaccharides lacking one O-methyl substituent and may result from incomplete methylation of the 2AB moiety. The fully methylated precursor ions at m/z 929 and 1089 were selected for MS fragmentation studies.

The quasimolecular ion at m/z 929 in this spectrum corresponds to a methylated 2AB-labeled oligosaccharide composed of a 2AB-labeled xylitol, two pentoses, and one uronic acid. The ESI MS2 and MS3 spectra (Fig. S8) contained a fragment ions series 929-755-595/523 that, together with NMR analysis and the mode of action of the family GH11 xylanase used to fragment the xylan, establish that the structure of this oligosaccharide is Xyl-Xyl-(MeGlcA)-Xyl-2AB.

The quasimolecular ion at m/z 1089 in the ESI spectrum of the per-O-methylated 2AB-labeled, acidic, xylo-oligosaccharides from Setaria italica corresponds to an oligosaccharide composed of a 2AB-labeled xylitol, three pentosyl residues, and one GlcA, which is 4-O-methylated in the native structure. ESI MS here indicated the existence of two or more distinct oligosaccharides (Supplementary Fig. S9). For example, the series of fragment ions at m/z 1089-915-755-595/523 is consistent with the sequence Xyl-Xyl-(GlcA)-Xyl-2AB, whereas the series of fragment ions at m/z 1089-915-741-595 provides evidence that sequence Xyl-(Pentosyl)-Xyl-(GlcA)-Xyl-2AB is present, where the pentosyl residue is most likely Araf based on our NMR data.

The ESI mass spectrum of the fraction enriched in the methylated neutral 2AB-labeled oligosaccharides (obtained by elution of the SPE cartridge with 25 % acetonitrile)
gave a series of [M + Na]+ ions at m/z 711, 871, 1031, and 1191 (Supplementary Fig. S7b), corresponding to oligosaccharides containing between 2 and 5 xyloses and a 2AB-xylitol. ESI MS® of the ion at m/z 1031 indicated that this ion can arise from both a linear and a branched sequence (Fig. 8). For example, the fragmentation pathways m/z 1031-857-683-537-363 and 1031-857-697-537-363 establish the presence of a pentosyl residue linked to the 2AB-xylitol residue Fig. 8). However, the fragmentation pathway at m/z 1031-857-697-377 corresponded to a linear oligosaccharide, i.e., Xyl-Xyl-Xyl-Xyl-2AB, thus indicating the presence of oligosaccharides in which the reducing end xylose is not branched.

Our combined ESI MS and NMR data provide substantial evidence that most but not all of the xylose at the reducing end of the GAX isolated from Setaria italica secondary cell walls is branched. Comparable results were obtained with the 2AB-labeled xylo-oligosaccharides generated from the 2AB-treated AIR of Miscanthus x giganteus, Panicum virgatum, Oryza sativa, and Brachyspiadium distachyon. Thus, in all of the Poaceae GAXs examined, the predominant reducing end structure is a xylose substituted either at O-2 with 4-O-MeGlcPA or at O-3 with Araf residue (Figs. 7, 8; Supplementary Figs. S7–S9). The presence of a 4-O-MeGlcPA substituent rather than a GlcPA substituent on the reducing end xylose of secondary wall GXs is one structural feature that distinguishes them from their wheat endosperm counterparts (Ratnayake et al. 2014).

Discussion

Monocot xylans are more structurally diverse than was previously reported

Our structural characterization of cell wall xylans from genera of 10 of the 12 monocot orders provides a new perspective on the diversity, function, and evolution of monocot cell walls. For example, the xylose from many Alismatales and Asparagales lack sequence I and their most abundant backbone substituents are disaccharides consisting of MeGlcPA or GlcPA with an Araf substituent at O-2 (Tables 1, 2; Figs. 1, 4, 5). These disaccharides are widespread among non-commelinid xylans but are present only in trace amounts in Arecales xylan and were not detected in the GAXs of the more derived commelinids (Table 1).

Side chains containing a substituted MeGlcPA were first identified in Eucalyptus GX (Shatalov et al. 1999; Reis et al. 2005). Detailed analyses of the xylo-oligosaccharides generated from Eucalyptus globulus wood GX established that one of these side chains is β-Galp-(1 → 2)-4-O-methyl-α-GlcPA-(1 → 2)- (Togashi et al. 2009). GlcPA substituted with an incompletely characterized pentosyl residue has been detected in the GX isolated from Arabidopsis thaliana inflorescence stems (Chong et al. 2015). Mortimer et al. (2015) have tentatively identified the pentose as Arap and suggested that the disaccharide is a component of a xylan present in the Arabidopsis thaliana primary wall. Our data unambiguously demonstrate that, in monocots and also in Eucalyptus, the pentose in this disaccharide side chain is α-L-Arap. To date, only the GX isolated from the secondary cell walls of Eucalyptus grandis has both β-Galp-(1 → 2)-MeGlcPA and α-L-Arap-(1 → 2)-MeGlcPA side chains, which are structurally homologous. The role of these disaccharides in xylan function(s) is not known. Further studies are also required to determine the functional consequences of this structural diversity.

Our data suggest that the Arap-containing GX we isolated from non-commelinid monocots is structurally similar to the primary wall GX described by Mortimer et al. (2015). Our results with Asparagus officinalis (Fig. 6b) also show that the Arap-containing GX is most abundant in actively growing organs, which are rich in cells surrounded by a primary wall. Nevertheless, the presence of this side chain is not a fundamental feature of non-commelinid monocot primary wall GX as it was not detected in the GX isolated from Lemna minor cell walls (Table 1). Lemna minor, which is an aquatic plant, is composed predominantly of growing cells surrounded by primary walls and contains a very limited vascular system (Landolt 1986). Additional studies are required to determine whether the distribution of Arap-containing GX is cell or tissue specific and to determine the functions of primary wall xylan.

Changes in xylan structure during the evolution and diversification of monocots

The combined results of numerous studies have shown that the primary walls of the graminids (clade in the Poales, commelinid monocots) are xylan rich and thus distinct from the pectin- and xyloglucan-rich primary walls of eudicots, non-commelinid monocots, and the Arecales (a basal non-graminid commelinid) (Harris and Hartley 1980; Bacic et al. 1988; Carpita and Gibbaut 1993; Carpita 1996; Harris et al. 1997). By contrast, the commelinids Zingiberales and Commelinales, which are more derived than the Arecales, have primary walls that contain similar proportions of pectin and hemicellulose (Harris et al. 1997; Harris 2006; Harris and Trethewey 2010). Our study extends the notion that wall composition varies according to commelinid phylogeny by showing that evolution and diversification within this clade were accompanied by pronounced changes in xylan structure, most notably by the appearance of Arap-containing side chains and subsequent increase in
the abundance of Araf-containing xylan in their primary and secondary cell walls (see Table 1).

The Poales (Poaceae, Cyperaceae, and Bromeliaceae) are the most derived commelinids and produce GAX that contain side chains with terminal and 2-linked Araf residues and little if any sequence 1. However, Poaceae GAX contains relatively low levels of GlcpA and MeGlcpA and is thus distinct from Cyperaceae and Bromeliaceae GAX. No sequence 1 was detected in the Poaceae GAX; nevertheless, we cannot preclude the possibility that a small number of cells in these plants produce a xylan containing this sequence as trace amounts of 1 were detected in the xylan from Ananas cosmosus (Bromeliaceae) walls.

Araf-containing side chains predominate in Poaceae GAX but account for only half of the backbone glycosyl substituents in the Cyperaceae and Bromeliaceae. Xylans in all the Zingiberales and Commelinales analyzed also contain similar proportions of Araf- and GlcpA-MeGlcpA-containing side chains, but were shown unambiguously to contain sequence 1 and lack side chains with 2-linked Araf. The abundance of Araf side chains is reduced to its lowest level in the GAX of the Arecales and results in a xylan with structural features that are more in common with the GXs of the non-commelinid monocots. We suggest that, among extant plants, the Arecales most closely resemble the ancestral monocots that first synthesized GAX. Such an observation is consistent with phylogenetic studies showing that the Arecales are the least derived commelinid (Carnachan and Harris 2000; Chase et al. 2006; Smith and Donoghue 2008).

Glycosyl sequence 1 is located at the reducing end of many but not all of the xylans produced by monocots (Table 1). The presence or absence of this sequence occurs both within and between orders and is thus unlikely to be directly correlated with the appearance of new orders during monocot evolution. Sequence 1 was readily identified in the GX isolated from Asparagus officinalis stems, yet little if any of this sequence was detected in the GX isolated from growing tips of the shoots (Fig. 6; Table 1). Similarly, this sequence was not detected in the GXs isolated from the cell walls of two different duckweeds (Araceae), Spirodela polyrhiza and Lemma minor (see Table 1). These aquatic plants are composed predominantly of growing cells and contain very little vascular tissue (Landolt 1986). Thus, the presence of sequence 1 is likely to be a characteristic of the secondary cell wall xylans of many monocots. Notable exceptions are the GAXs in the secondary cell walls of Poales (Poaceae, Cyperaceae, and Bromeliaceae). These GAXs lack sequence 1 but often contain substituted xylosyl residues at their reducing end (Figs. 7, 8, Supplementary Figs. S7, S8, and S9). This structural feature was also identified in the primary wall GAX of Triticum aestivum endosperm (Ratnayake et al. 2014). It is not known whether the substituted reducing end xyloses are also present in the GXs of the Alismatales and Asparagales that also lack sequence 1. Additional studies are required to identify the genes and enzymes involved in the biosynthesis of these reducing end sequences and to determine the relationship between the xylan’s reducing end structure and its function in the cell wall.

Do non-graminid commelinids synthesize distinct GAXs and GXs?

We have shown that the Zingiberales and Commelinales analyzed in this study produce xylans with structural features characteristic of both graminid GAX and non-commelinid/dicot GX (Table 1). Most studies, including ours, have determined the structural characteristics of xylans by characterizing the oligosaccharides generated by endoxylanase fragmentation. This procedure cannot readily distinguish between the presence of two distinct polysaccharides that contain different structural motifs or a single polysaccharide that contains all the structural motifs. Nevertheless, we believe our results are more consistent with the presence of both GAX and GX as separate polysaccharides. For example, Araf substituents are abundant in the xylan from the leaves of Tradescantia virginiana (Commelinacae), whereas GlcpA/MeGlcpA substituents predominate in the stem xylan (Fig. 6). Such results are consistent with the predominance of GAX in leaf primary walls and GX in stem secondary walls. The presence of these two types of xylans in different organs of the plant may result from cell- and/or tissue-specific differences in the polysaccharide synthesis machinery as proposed for primary and secondary wall synthesis in Arabidopsis (Mortimer et al. 2015).

The cell walls of many Commelinales and Zingiberales contain xyloglucan that has characteristics of both eudicots
(XXXG type) and grasses (XXGn type) (Hsieh and Harris 2009). XXXG- and XXGn-type xyloglucans are also present in comparable proportions in the walls of the Bromeliaceae (Poales). It is not known whether one xyloglucan contains both structural motifs or whether distinct XXXG- and XXGn-type xyloglucans are formed by specific cells and/or tissues in the plant organs. This latter notion is not unprecedented as several examples of cell and tissue-specific differences in hemicellulose structure have been described. For example, the root hair walls of numerous plants including Arabidopsis thaliana and Oryza sativa (Peña et al. 2012; Liu et al. 2015), and Nicotiana alata pollen tube walls (Lampugnani et al. 2013) contain a xyloglucan, which is structurally distinct from the xyloglucan present in other tissues of the same plant. GX in the interfascicular fibers of Arabidopsis thaliana stems contains GlcA and MeGlcA substituents, whereas the GlcA is fully methylated in the GX in stem and root vessels (Urbanowicz et al. 2012).

Our results suggest that organ-, tissue-, and cell-specific structural differences of the cell wall polysaccharides may be more substantial than the differences between taxonomic groups. This notion adds an additional layer of complexity to the study of cell wall diversity and evolution and is especially relevant in studies that infer the activities of gene products by the inability to detect a particular polysaccharide structure. Structurally distinct polysaccharide may indeed be synthesized by a limited number of specific cell types.

Origins of the graminid GAXs

The evolution and diversification of the commelinids were accompanied by substantial changes in the polysaccharide composition of the cell wall together with changes in the fine structures of the wall polysaccharides. Our results suggest that some of the changes in xylan structure are not necessarily species specific, since we found structural features characteristic of different types of xylan in the same species (Fig. 6; Table 1). We propose that one pivotal event involved the addition of Araf residues to GX to form a prototypical GAX in a basal commelinid. This may have first occurred in cells that only produce primary walls but was then attained by cells producing secondary walls. This hypothesis is consistent with our results showing that the amount of Araf-containing xylans increases through the commelinids and that these GAXs are the most abundant xylan in the Poaceae, which are the most derived monocots. The addition of Araf residues may have endowed the newly formed GAX with physical properties that allowed it to perform some of the functions previously performed by xyloglucan and/or pectin, as previously suggested (Carpita 1996). Subsequently, the transition to primary cell walls with decreased amounts of pectin and xyloglucan may have been facilitated by the ability to synthesize xylan with Araf side chains.

In summary, we have shown that several structural motifs of xylan have changed during monocot evolution. Such data will be of value in studies to determine whether this structural diversity is reflected in changes in the molecular and cellular mechanisms that control the synthesis of these polysaccharides. Knowledge of differences in xylan structure will also facilitate the use of biomass from diverse monocots for conversion to biofuel and thereby contribute to sustainable biomass production.

Author contribution statement WSY, MJP, and MAO conceived and designed research. AK, JB, MAO, WSY, and MJP carried out experiments and analyzed the data. M. B. identified and provided plant materials. MJP, MAO, AK, and WSY wrote the manuscript. All authors read and approved the manuscript.

Acknowledgments This research was funded by the BioEnergy Science Center (BESC). BESC is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. We also acknowledge the U.S. Department of Energy-funded Center for Plant and Microbial Complex Carbohydrates (Grant DE-FG02-93ER20097) for equipment support.

Conflict of interest The authors declare that they have no conflict of interest

References

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 3:1–130


Johansson M, Samuelson O (1977) Reducing end groups in birch and liverwort xyloglucans containing galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. Glycobiology 18:891–899


Landolt E (1986) Biosystematic investigation in the family of duckweeds (‘‘Lemnaceae’’). The family of ‘‘Lemnaceae’’: a monographic study, vol 2. Veröffentlichungen des Geobotanischen Institutes der ETH, Zürich


