Identification of an arabinopyranosyltransferase from *Physcomitrella patens* involved in the synthesis of the hemicellulose xyloglucan

Lei Zhu[1], Murali Dama[2], and Markus Pauly[2]

[1] Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA, 94720, USA

[2] Institute of Plant Cell and Biotechnology, University of Dusseldorf, Dusseldorf, Germany

To whom correspondence should be addressed. Markus Pauly

Institute of Cell Biology and Biotechnology,

Heinrich-Heine University Duesseldorf,

Düsseldorf, Germany

E-mail: m.pauly@hhu.de

Tel: +49 211 81-15680

ABSTRACT

The hemicellulose xyloglucan consists of a backbone of a β-1,4 glucan substituted with xylosyl moieties and many other, diverse sidechains that are important for its proper function. Many, but not all glycosyltransferases involved in the biosynthesis of xyloglucan have been identified. Here, we report the identification of an hitherto elusive xyloglucan:arabinopyranosyltransferase. This glycosyltransferase was isolated from the moss *Physcomitrella patens*, where it acts as a **X**yloglucan "**D"**-side-chain **T**ransferase (**XDT**). Heterologous expression of *XDT* in the *Arabidopsis thaliana* double mutant *mur3.1 xlt2*, where xyloglucan consists of a xylosylated glucan without further glycosyl substituents, results in the production of the arabinopyranose-containing "D" side chain as characterized by oligosaccharide mass profiling, glycosidic linkage analysis, and NMR analysis. In addition, expression of a related *Physcomitrella* glycosyltransferase hortholog of *XLT2* leads to the production of the galactose-containing "L" side chain. The presence of the "D" and "L" xyloglucan side chains in *PpXDT mur3.1 xlt2* and *PpXLT2 mur3.1 xlt2* transgenic plants, respectively, rescue the dwarfed phenotype of untransformed *mur3.1 xlt2* mutants to nearly wild-type height. Expression of *PpXDT* and *PpXLT2* in the Arabidopsis *mur3.1 xlt2* mutant also enhanced root growth.

Keywords

Hemicellulose – Xyloglucan – glycosyltransferase – arabinopyranosyltransferase - *Physcomitrella patens*

INTRODUCTION

The plant cell wall is a complex extracellular matrix composed of polysaccharides such as cellulose, hemicellulose, and various pectic polysaccharides, glycoproteins and the polyphenol lignin. The major hemicellulose xyloglucan (XyG) is found in all land plants and is especially abundant in the primary cell wall of dicots (1). XyG in the primary cell wall attaches to cellulose microfibrils non-covalently via H-bonds and its metabolism in the wall is thought to play a role in cell elongation (2, 3, 4). However, the precise molecular role of XyG in plant growth and development is not clear (5, 6, 7) as mutant plants lacking XyG do not exhibit an obvious growth phenotype (8). Initially it was thought that a particular XyG structure is plant species specific, but recently tissue specific structures within a plant species have emerged (9, 10, 11). XyG has not only been found in higher plants, but also in non-vascular plants such as liverworts and mosses (12).

XyG consists of a backbone of β-1,4 glucan substituted with xylosyl residues that are often further decorated with other sugar residues and/or acetyl-residues, leading to the discovery of more than 20 structurally different XyG side chains to date (13, 14, 15). Due to the structural diversity, a one-letter code has been established describing XyG side chains (16). According to this code G refers to an unsubstituted glucosyl backbone residue, while X depicts a xylosylated glucosyl residue as in α-D-xylose-6-β-D-glucose. X can be further extended on the xylosyl unit at *O-2* with galactosyl-, arabinopyranosyl-, galacturonsyl-, xylosyl- or arabinofuranosyl residues resulting in L, D, Y, U, and S side chains, respectively (12, 17, 18, 19, 20).

XyG assembly requires various glycosyltransferases (GTs) that add specific sugars to the extending polymer. Many GTs involved in XyG synthesis have been identified that belong to various Carbohydrate-Active Enzymes (CaZy) families based on gene-sequence homology (1, 21). One of the CaZy families involved in XyG sidechain biosynthesis is the GT47 family, including MUR3, XLT2, and XST (Fig. 1). MUR3 represents a XyG:galactosyltransferase, which adds a β-galactosyl-residue at the *O*-2 position to a xylosyl-residue resulting in the sidechain L (22). MUR3 transfers the galactosyl-moiety to a specific xylosyl-residue on the XyG chain leading to the occurrence of an XXLG oligosaccharide motive in XyG. In contrast, a related GT47 protein, XLT2, adds the galactosyl-residue to another xylosyl-residue leading to a XLXG motive indicating that these GTs exhibit

78 regioselectivity (23). GT47 family members can also transfer galacturonic acid (XUT1) or arabinofuranosyl-
79 moieties (XST) (17, 24) to the xylosyl-residue.

80 The moss *Physcomitrella patens* was found to contain XyG (12) with branched side chains containing
81 galacturonosyl and arabinopyranosyl residues at the *O*-2 position of their xylosyl residues (12). The
82 arabinopyranosyl residue is unique as it has also been found in the XyG of lower plants such as the Lycophytes
83 including Selaginella, Equisetales, Polypodiales and Cycadales (12, 18), but not in any gymnosperm or
84 angiosperm plant to date.

85 To gain insights into the function of the XyG:arabinopyranosyl residue on XyG side chains, we describe here the
86 identification of the responsible *Physcomitrella* arabinopyranosyltransferase present in the GT47 family.
87 Because the simplest XyG side chain containing an arabinopyranosyl residue has been abbreviated with the one
88 letter code D (1, 16), we named the responsible protein XDT (**X**yG **D** side chain **T**ransferase).

89

90 **RESULTS**

91 **Identification of XyG-related GT47 Family Members in the Moss *Physcomitrella patens***

92 The amino acid sequence of the Arabidopsis XyG-related GT47 family member AtMUR3 was used as a bait to
93 identify related GT candidates of *Physcomitrella* present in the Joint Genome Institute database Phytozome
94 (phytozome.jgi.doe.gov). Based on amino acid sequence homology 13 *Physcomitrella* proteins were identified,
95 which were also homologous to other, known GT47 XyG related genes from various species (AtXLT2, AtXUT,
96 OsMUR3, OsXLT2, and SlXST; Fig. 2). Of the 13 Physcomitrella proteins, 6 members grouped closely in a
97 MUR3 subclade. The other 7 Physcomitrella proteins fell into the XLT2 subclade that also included XST and
98 XUT. Based on the location in the protein phylogenetic tree 9 non-redundant proteins were chosen for further,
99 functional investigation (Pp1918, Pp42620, Pp201625, Pp2661, Pp21725, Pp173836, Pp156311, Pp110748, and
00 Pp13057).

01

02 **Functional Complementation in Arabidopsis and Characterization of XDT**

03 To assign GT functions to the 9 selected Physcomitrella GT47 family members heterologous expression of

04 individual proteins in the Arabidopsis double mutant *mur3.1 xlt2* was pursued. XyG derived from the various
05 complemented Arabidopsis plants was analyzed by oligosaccharide mass profiling (OLIMP) (25), whereby XyG
06 was solubilized from wall materials using a xyloglucan specific endoglucanase and the resulting XyG
07 oligosaccharide mixture was analyzed by MALDI-TOF mass spectrometry (Fig. 3). The OLIMP profile of
08 untransformed *mur3.1 xlt2* mutant plants shows the occurrence of a single oligosaccharide motive with a m/z of
09 1,085 representing the XyG oligosaccharide XXXG consisting only of the glucan backbone with xylosyl moieties
10 but no further substitutions. This OLIMP profile was retained when seven of the Physcomitrella genes were
11 constitutively expressed in Arabidopsis *mur3.1 xlt2* indicating that in Arabidopsis these genes are not involved in
12 XyG biosynthesis.
13
14 However, expression of *Pp201625 (PpXDT)* in *mur3.1 xlt2* resulted in a XyG that contained an oligosaccharide
15 of m/z 1,217 indicating that this GT affects XyG biosynthesis and is responsible for adding an additional pentosyl
16 residue to XXXG in *mur3.1 xlt2* (Fig. 3). Moreover, 5 additional XyG oligosaccharides were observed when
17 *PpXDT* was expressed in the double mutant *mur3.1 xlt2* (Fig. S1). These ions with an m/z of 1,349, 1,363, 1,481,
18 1,495, and 1,627 correspond to oligosaccharide structures consisting of 4 hexoses and 5 pentoses (H4P5), H4P4
19 with an additional deoxysugar, likely to be fucose, H4P6, H4P5 with an additional fucose, and H4P6 with a fucose,
20 respectively. Mass spectrometry neither gives an indication which kind of pentose was added nor where the
21 pentose would be attached. To determine the fine structure of the dominant novel XyG oligosaccharide (m/z 1,217)
22 the oligosaccharide was isolated/ enriched by subjecting the XyG oligosaccharide mixture obtained from wall
23 material of *Pp201625 mur3.1 xlt2* to High Performance Anion Exchange Chromatography with Pulsed
24 Amperometric Detection (HPAEC-PAD; Fig. 4). Oligosaccharide(s) with a molecular mass of m/z 1,217 eluted
25 at ~13.2 min and were collected for further analysis. Some impurities of the XyG oligosaccharide XXXG were
26 present in the collected fraction due to its adjacent elution. The retention time of the novel oligosaccharide was
27 found to be the same as a well characterized XyG oligosaccharide isolated from *Selaginella kraussiana* (18),
28 termed XXDG, a XyG oligosaccharide containing an arabinopyranosyl residue (Fig. 4). The isolated/enriched
29 oligosaccharide with a m/z of 1217 isolated from *Pp201625 mur3.1 xlt2* was subjected to glycosidic linkage

analysis (Table 1). The results indicate the presence of t-Ara*p* and 2-Xyl*p* in an equal ratio supporting the hypothesis that Ara*p* is attached to Xyl*p* at *O*-2 thus representing the structure of XyG D side chain. No t-Ara*f* was detected. To gain further insights into this structure ^{1}H NMR was performed (Fig. 5, Additional Fig. S1). Based on previously observed chemical resonances (12) the data confirmed the presence of an anomeric signal of an α-linked arabinopyranosyl residue (chemical shift of 4.488 ppm). In addition, the corresponding substituted 2-α-Xyl*p* residue was identified with a chemical shift of 5.133 ppm. The observed chemical shifts are therefore consistent with the presence of the oligosaccharide XXDG (Fig. 5). Hence, Pp201625, when expressed in the Arabidopsis mutant *mur3.1 xlt2*, plays a role in transferring arabinopyranose to XXXG generating a XyG D side chain and was thus termed XyG D-side-chain Transferase (XDT).

Heterologous expression of *PpXLT2*

The Physcomitrella GT47 family also contains Pp42620, a protein that phylogenetically belongs to the AtXLT2 subclade (Fig.2). Expression of Pp42620 in Arabidopsis *mur3.1 xlt2* resulted in the generation of various XyG oligosaccharides, when the transgenic plants were analyzed by oligosaccharide mass profiling. The oligosaccharides with a m/z of 1,247 represents XXXG plus an additional hexose – the minor new oligosaccharide with a m/z of 1,409 represents XXXG plus 2 hexoses (Fig. 3). To determine the fine structure of the dominant XyG oligosaccharide (m/z 1,247) the XyG oligosaccharide mixture generated from wall materials of *Pp42620 mur3.1 xlt2* was analyzed by HPAEC-PAD. Compared to previously published data (26) and using commercially available tamarind XyG oligosaccharides as standards the novel oligosaccharide exhibited the same retention time as tamarind XyG oligosaccharide XLXG (Fig. 4). Also oligosaccharides with the retention time of XXLG and the double galactosylated XLLG oligosaccharide were present suggesting that *Pp42620,* when expressed in Arabidopsis, exhibits mainly XLT2 activity, hence its name PpXLT2, but also to some extent MUR3 galactosyltransferase activity.

Growth Phenotypes of *PpXDT mur3.1 xlt2* and *PpXLT2 mur3.1 xlt2*

The structure of XyG has been shown to affect vegetative growth. For example, the Arabidopsis double mutant *mur3.1 xlt2* containing XyG entirely composed of XXXG units exhibits dwarfism (9, 24) (Fig. 6). When *PpXDT* and *PpXLT2* were expressed in *mur3.1 xlt2* using a constitutive promotor vegetative (stem) growth was restored to nearly normal heights in most of the lines (Fig. 6 A, B). The Arabidopsis double mutant *mur3.1 xlt2* exhibits also shorter roots (17) (Fig. 7 A, B). Expression of Pp*XDT* and *PpXLT2* in the double mutant leads to root growth that is not significantly different than Arabidopsis WT.

DISCUSSION

Identification of a XyG:Arabinopyranosyltransferase (XDT)

The xylosyl residue of XyG is often substituted at the *O*-2 position with a variety of glycosyl residues including galactosyl, galacturonosyl, xylopyranosyl, arabinofuranosyl or arabinopyranosyl moieties (1). Using *in vitro* assays, loss-of-function mutants and functional complementation approaches in Arabidopsis have led to the successful identification of many of the responsible GTs including two galactosyltransferases MUR3 and XLT2 (22, 23), a galacturonsyltransferase XUT1 (17), an arabinofuranosyltransferase XST (24) and as identified and characterized in this study an arabinopyranosyltransferase XDT.

The availability of the *Physcomitrella patens* genome (27) allowed us to identify this XyG:arabinopyranosyltransferase XDT. Expression of the *XDT* gene from Physcomitrella led to synthesis of the XyG D side chain in the *Arabidopsis* double mutant *mur3.1 xlt2* as evidenced by XyG analysis by MALDI-TOF MS, HPAEC-PAD, linkage analysis and NMR. When a MUR3 ortholog from rice is expressed in the Arabidopsis *mur3.1 xlt2* double mutant XyG does not only become galactosylated, it also becomes fucosylated (9) as the galactosylated side chain L is the required acceptor substrate for the XyG:fucosyltransferase (28) resulting in the F sidechain. Here, expression of *XDT* resulted not only in arabinosylated side chains, but also to a much lesser extent in side chains containing an additional deoxyhexose. This data is consistent with the occurrence of a fucosylated D side chain termed E, which has been observed in XyG derived from *Equisetum hyemale* and *Selaginella kraussiana* (12, 18). The Arabidopsis XyG:fucosyltransferase AtFUT1/AtMUR2 is apparently not only able to transfer fucosyl residues to galactosyl but also arabinopyranosyl residues. Similar to previous reports

(12, 18) acetylated versions of the D side chain were not observed, indicating that the Arabidopsis XyG:*O*-acetyltransferase AtAXY4/AtAXY4L (29) specifically adds acetyl substituents to galactosyl residues.

The galactosyltransferases AtMUR3 and AtXLT2 act regiospecific, i.e. they transfer the galactosyl moiety to a specific xylosyl residue leading to the generation of XXLG or XLXG, respectively. The expression of *PpXDT* in Arabidopsis also lead to XyG oligosaccharides that in addition to arabinopyranosyl residue contain additional pentoses. Although the nature and position of these additional pentosyl residues remain to be determined it seems clear that XDT is more promiscuous in nature than MUR3/XLT2. Either the enzyme can transfer other pentoses than arabinopyranoses such as arabinofuranoses or more likely it can add arabinopyranoses to different positions on XyG resulting not only in the XyG oligosaccharide XXDG (m/z 1,217, Table S1) but also XDXG (m/z 1,217), XDDG (m/z 1,349), and even DDDG (m/z 1,481) and their fucosylated versions XXEG (m/z 1,363), XDEG (m/z 1,495), and DDEG (m/z 1,627). Ions of all these oligosaccharides were present when *PpXDT* was expressed in the *mur3.1 xlt2* mutant (Table S1). Thus, XDT does not seem to act regiospecific.

Functional Conservation of XyG-related Genes in the GT47 Family

The GT47 family is a large carbohydrate-active enzyme (CaZy) family involved in cell wall biogenesis containing a subclade represented by the XyG:galactosyltransferase AtMUR3 (22) and includes XLT2, another XyG:galactosyltransferase (23). The identified Physcomitrella XyG:arabinopyranosyltransferase XDT also belongs to this subclade, as it forms the same glycosidic linkage on the same acceptor substrate albeit utilizing a different donor substrate (UDP-L-arabinopyranose). Taken together with the characterized functions of other GT47 members such as XUT1 from *Arabidopsis* and XST from tomato, members of the GT47 MUR3 subclade in land plants have evolved in transferring a glycosyl moiety to the xylosyl residue of XyG at *O*-2.

Within this clade another functional XyG GT was identified in *Physcomitrella patens*, PpXLT2. Analysis of the XyG present in the constitutive expression of *PpXLT2* in *Arabidopsis mur3.1 xlt2* resulted in the occurrence of XLXG indicating that PpXLT2 in *Arabidopsis* can carry out the same function as AtXLT2. However, in addition XXLG was formed including its fucosylated version XXFG. Thus, unlike AtXLT2 from Arabidopsis (23), SlXLT2 from tomato (24), and OsXLT2 from rice (9) the Physcomitrella ortholog PpXLT2 does not act

regiospecific, it also exhibits MUR3 activity. Therefore, while the function of XLT2 is functionally conserved across land plants including bryophytes, its regioselectivity of this GT has apparently evolved later as it has to date only been observed in angiosperm species.

XyG side chains impact aerial and root growth

Arabidopsis mutant *mur3.1* contains a point mutation in the *MUR3* gene and renders it inactive (22). As a result, XyG in this mutant does not contain the XXLG oligosaccharide, but still the galactosylated XLXG motive. Mutant plants show normal plant growth except for minor effects in trichome morphology (22). However, the double mutant *mur3.1 xlt2*, whose XyG does neither contain XXLG nor XLXG oligosaccharides, but consists entirely of XXXG oligosaccharides displays a dwarfed phenotype (30). Complementing this mutant with various XyG GT47 genes results in the rescue of the growth phenotype. This rescue has been observed with the expression of *MUR3* and *XLT2* from a variety of species such as from rice *OsXLT2* (9), tomato *SlXLT2* (24), or as shown here from Physcomitrella *PpXLT2*. Moreover, arabinofuranosylation by expressing XST (24) and as shown here arabinopyranosylation through XDT also restores the phenotype of the double mutant, not only the growth of vegetative tissue, but also root growth. This indicates that galactosylation or the occurrence of the L side chain is not required for normal growth, but that alternative substitutions such as arabinofuranosylation and arabinopyranosylation resulting in the S and D side chains, respectively, suffice for normal plant growth. It is known that XyG that consists only of XXXG self-aggregates and precipitates *in vitro* (31, 32). Such precipitation of non-galactosylated XyG in the *mur3.1 xlt2* mutant might occur already during its biosynthesis in the Golgi-apparatus of impacting the endomembrane system function. Indeed, the Arabidopsis *mur3.3* mutant, an insertional knockout mutant, exhibits severe dwarfism with a concomitant aggregation of endomembranes and intracellular accumulation of polymers (30). However, when the *mur3.3* mutant is crossed with the XyG-lacking *xxt1 xxt2* mutant the resulting mutant plant exhibits not only again a normal growth phenotype but also a normal endomembrane morphology. XyG is still lacking in these plants. Hence a structurally abberant XyG with low or lacking galactosylation is detrimental to plant development, whereas a lack of XyG is not.

METHODS

Plant growth

Seeds of Arabidopsis wild-type Col-0, *mur3.1 xlt2* (24) and the transgenic plants generated here were germinated either in soil pots or on half MS agar plates. Plants were grown in a Percival growth chamber at 21°C under 16/8 hour light/dark cycle with 70% humidity.

Phylogenetic analysis

Pyscomitrella XyG-related GTs were identified by using AtMUR3 as a template for a BlastP search of the *Physcomitrella* phytozome database (version 10.1). Alignment of the XyG-related GT47 proteins was achieved by MUSCLE alignment and construction of a phylogeny tree using PhyML (33, 34).

Gene constructs and plant transformation

Physcomitrella candidate genes were amplified either by PCR from genomic DNA or by RT-PCR from total RNA extracted from 1-week-old protonemal tissue and 1-month-old gametophytes of *Physcomitrella patents*. Primer sequences used for cloning are listed in the additional table S2. The amplified genes were cloned into the expression vector pORE-E4, which was transformed to *Agrobacterium tumefaciens* strain GV3101, and subsequently transformed to Arabidopsis via the floral dip method (35). Three generations of transgenic *PpXDT* and *PpXLT2* plants were selected on half MS agar (0.8%) plates containing 60µg ml^{-1} Kanamycin. Germinated seedlings were then move into soil for continuous growth.

Analysis of xyloglucan

XyG oligosaccharides were extracted from leaf tissue of Arabidopsis Col-0, *Atmur3.1 xlt2*, *PpXDT mur3.1 xlt2, PpXLT2 mur3.1 xlt2*, and *Selaginella Kraussiana* by alcohol insoluble residue (AIR) preparation followed by xyloglucanase digestion (36) and subsequent XyG oligosaccharide profiling by MALDI-TOF MS and HPAEC-PAD as described (23, 24).

58

59 **Purification of xyloglucan oligosaccharide XXDG**

60 Extraction of the XyG oligosaccharide XXDG (m/z 1217) from transgenic plants was performed
61 according to methods described in Schultink et al., 2013. However, the reduction of the oligosaccharides by
62 sodium borohydride was performed after separation of the oligosaccharides by HPAEC-PAD. The reduced
63 oligosaccharides were neutralized, de-salted using a ENVI-CARB reverse phase column (Sigma Aldrich, USA)
64 and freeze dried in a lyophilizer.

65

66 **Glycosyl linkage analysis**

67 Glycosidic linkage analysis of XyG oligosaccharides was performed as described (24).

68

69 **NMR analysis**

70 Reduced oligosaccharides were dissolved in 0.3 ml of D_2O (99.9%, company?), freeze dried and dissolved
71 again in 0.3 ml of D_2O (99.9%) containing 0.05 % of 3-(trimethylsilyl)-propionic-2,2,3,3-d4 acid sodium salt).
72 The ^{1}H NMR spectra were recorded on a Bruker Avance 600 MHz NMR spectrometer equipped with an inverse
73 gradient TXI $^{1}H/^{13}C/^{15}N$ Cryoprobe at 298 K. All chemical shifts were referenced relative to 3-(trimethylsilyl)
74 propionic-2,2,3,3-d4 acid (0.00 ppm for ^{1}H). The NMR data processing and analysis was performed using
75 Bruker's Topspin 3.1 software.

76

77 **Abbreviations**

78 XyG	Xyloglucan
79 XEG	Xyloglucan-specific endoglucanase
80 GT	Glycosyltransferase
81 CaZy	Carbohydrate-active enzyme
82 XDT	Xyloglucan D-side-chain Transferase
83 OLIMP	Oligosaccharide mass profiling

MALDI-TOF	Matrix-Assisted Laser Desorption/ Ionization Time of Flight Mass Spectrometry
HPAEC-PAD	high-performance anion-exchange chromatography withpulsed amperometric detection
GC-MS	Gas chromatography mass spectrometer
AIR	Alcohol insoluble residue
t-Ara*p*	terminal arabinopyranose
2-Xyl*p*	2-xylopyranose
t-Ara*f*	terminal arabinofuranose

Table 1. Glycosidic linkage analysis of XyG oligosaccharide fraction with a m/z 1217 (with contamination of XXXG) from leaf walls of *PpXDT mur3.1 xlt2*.

Sugar moiety	Abundance (%)
t-Ara*p*	2.88
t-Xyl	38.20
2-Xyl	2.88
6-Glc	14.57
4-Glc	13.29
4,6-Glc	28.18

Note: t-Arap, terminal arabinopyranose; t-Xylp, terminal xylopyranose; 2-Xyl, 2-linked xylose; 6-Glc, 6-linked glucose; 4-Glc, 4-linked glucose; 4,6-Glc, 4,6-linked glucose

Figure legends:

Figure 1. Xyloglucan oligosaccharide structures and GT47 glycosyltransferases involved in its synthesis. The xyloglucan one-letter code nomenclature is indicated below the structure. MUR3 and XLT2 - XyG:galactosyltransferases; XST - XyG:arabinofuranosyltransferase; XDT – XyG:arabinopyranosyltransferase

Figure 2. Phylogeny of XyG-related GT47 proteins. Red font - Protein sequences from known XyG:galactosyltransferases AtMUR3, OsMUR3, AtXLT2 and OsXLT2, the galacturonosyltransferase AtXUT, and the arabinofuranosyltransferases SlXST1 and SlXST2. Black font - *Physcomitrella patens* proteins obtained from a BlastP search against Phytozome database. Phylogenetic tree built by PhyML.

Figure 3. XyG oligosaccharide mass profiling by MALDI-TOF MS. XyG oligosaccharides derived from leaf tissue of the Arabidopsis double mutant *mur3.1 xlt2*, transgenic lines expressing *PpXDT* or *PpXLT2* in *mur3.1 xlt2*. Numbers indicate m/z and their potential structure is shown in supplemental table 1. m/z 1085 represents the known XyG oligosaccharide structure XXXG.

Figure 4. XyG oligosaccharide separation by HPAEC-PAD. Peaks were assigned based on retention times from published work (Schultink et al., 2013; Hsieh & Harris, 2012; Megazyme Inc., Ireland) as well as assignment based on mass spectrometry. The fractions containing m/z 1217 (XXDG) was collected and further analyzed by glycosidic linkage analysis (Table 1) and NMR spectroscopy (Fig. 5).

Figure 5. Anomeric region of the ^{1}H NMR spectra of the reduced XyG oligosaccharide XXDGol contaminated with XXXGol. ^{1}H NMR peaks were labelled according to CCRC (ccrc.uga.edu) database and the literature (Pena et al., 2008). The putative structure of the oligosaccharide is shown in the upper left corner indicating the order of the side-chains on top and the one-letter code below.

Figure 6. A) Growth habit of 8-week old Arabidopsis plants. B) Height of Inflorescence stems of 8-week old Arabidopsis plants. n≥3. Small letters – ANOVA analysis.

Figure 7. Root growth of Arabidopsis plants. A) Root habit B) Root length of 8-day-old seedlings, n=10, small letters – ANOVA analysis.

Supplemental Figures:

Table S1. The XyG oligosaccharide composition and relative abundance of ions (m/z) seen in Fig. 3.

Table S2. Primers for PCR amplification of genes and quantitative RT-PCR.

Figure S1: Figure S1. Anomeric region of the ^{1}H NMR spectra of the oligosaccharide XXXG obtained from *mur3.1 xlt2*. ^{1}H NMR peaks were labelled according to CCRC (ccrc.uga.edu) database. The putative structure of the oligosaccharide is shown in the upper right corner indicating the order of the side-chains on top and the one-letter code below.

48 **Declarations**

49 **Ethics approval and consent**

50 The leaf samples used in this study were collected from growth chamber in our lab and from botanical garden at
51 University of California, Berkeley with permission of the curator. The experimental research was undertaken in
52 accordance with local guidelines. For access to the plants, please contact the corresponding author.

53 **Consent for publication**

54 Not applicable

55 **Availability of data and materials**

56 Constructs described in this work and datasets analysed during the current study are available from the
57 corresponding author upon request.

58 **Competing interests**

59 The authors declare no conflict of interest.

60 **Funding**

61 LZ and MD were supported by a grant from the Energy Biosciences Institute, University of California,
62 Berkeley, USA

63 **Acknowledgements**

64 We thank Dr. Tom Kleist and Professor Sheng Luan for providing moss protonemal and gametophyte tissue of
65 Physcomitrella, and the botanical garden of the University of California, Berkeley for *Selaginella kraussiana.*

66 **Authors Contributions**

67 LZ performed the experiments, analyzed the results and wrote the manuscript, MD performed the NMR
68 experiments and wrote the manuscript, MP conceived the work, analyzed the data, and wrote the manuscript.

69 **Author's Affiliations**

70 (1) Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA, 94720, USA
71 (2) Institute of Plant Cell and Biotechnology, University of Dusseldorf, Dusseldorf, Germany

72

73

74 **References**

75 1. Pauly M, Keegstra K: Biosynthesis of the Plant Cell Wall Matrix Polysaccharide Xyloglucan. Annu Rev
76 Plant Biol 2016, 67:235-259.
77 2. Keegstra K, Talmadge K, Bauer W, Albersheim P: Structure of plant cell walls: III. Amodel of the walls
78 of suspension-cultured sycamore cells based on interconnections of macromolecular componenets. Plant
79 Physiol 1973, 51: 188-196.
80 3. Valent BS, Albersheim P: The structure of plant cell walls: v. On the binding of xyloglucan to cellulose
81 fibers. Plant Physiol 1974, 54: 105-108
82 4. Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredez A, Persson S,
83 Raab T et al: Toward a systems approach to understanding plant cell walls. Science 2004,
84 306(5705):2206-2211.
85 5. Talbott LD, Ray PM: Molecular size and separability features of pea cell wall polysaccharides:
86 implications for models of primary wall structure. Plant physiology 1992, 98(1):357-368.
87 6. Thompson DS: How do cell walls regulate plant growth? Journal of experimental botany 2005,
88 56(419):2275-2285.
89 7. Park YB, Cosgrove DJ: Xyloglucan and its interactions with other components of the growing cell wall.
90 Plant Cell Physiol 2015, 56(2):180-194.
91 8. Cavalier DM, Lerouxel O, Neumetzler L, Yamauchi K, Reinecke A, Freshour G, Zabotina OA, Hahn
92 MG, Burgert I, Pauly M et al: Disrupting two Arabidopsis thaliana xylosyltransferase genes results in
93 plants deficient in xyloglucan, a major primary cell wall component. The Plant cell 2008, 20(6):1519-
94 1537.
95 9. Liu L, Paulitz J, Pauly M: The presence of fucogalactoxyloglucan and its synthesis in rice indicates
96 conserved functional importance in plants. Plant physiology 2015, 168(2):549-560.
97 10. Lampugnani ER, Moller IE, Cassin A, Jones DF, Koh PL, Ratnayake S, Beahan CT, Wilson SM, Bacic
98 A, Newbigin E: In vitro grown pollen tubes of Nicotiana alata actively synthesise a fucosylated
99 xyloglucan. PloS one 2013, 8(10): e77140.

11. Dardelle F, Le Mauff F, Lehner A, Loutelier-Bourhis C, Bardor M, Rihouey C, Causse M, Lerouge P, Driouich A, Mollet JC: Pollen tube cell walls of wild and domesticated tomatoes contain arabinosylated and fucosylated xyloglucan. Ann Bot 2015, 115(1):55-66.

12. Pena MJ, Darvill AG, Eberhard S, York WS, O'Neill MA: Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. Glycobiology 2008, 18(11):891-904.

13. Schultink A, Liu L, Zhu L, Pauly M: Structural Diversity and Function of Xyloglucan Sidechain Substituents. Plants (Basel) 2014, 3(4):526-542.

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

15. Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink A, Xiong G: Hemicellulose biosynthesis. Planta 2013, 238(4):627-642.

16. Fry SC, Aldington S, Hetherington PR, Aitken J: Oligosaccharides as signals and substrates in the plant cell wall. Plant physiology 1993, 103(1):1-5.

17. Pena MJ, Kong Y, York WS, O'Neill MA: A galacturonic acid-containing xyloglucan is involved in Arabidopsis root hair tip growth. Plant Cell 2012, 24: 4511-4524

18. Hsieh YS, Harris PJ: Structures of xyloglucans in primary cell walls of gymnosperms, monilophytes (ferns sensu lato) and lycophytes. Phytochemistry 2012, 79:87-101.

19. Hilz H, de Jong LE, Kabel MA, Verhoef R, Schols HA, Voragen AG: Bilberry xyloglucan--novel building blocks containing beta-xylose within a complex structure. Carbohydrate research 2007, 342(2):170-181.

20. Jia Z, Cash M, Darvill AG, York WS: NMR characterization of endogenously O-acetylated oligosaccharides isolated from tomato (Lycopersicon esculentum) xyloglucan. Carbohydrate research 2005, 340(11):1818-1825.

21. Coutinho PM, Deleury E, Davies GJ, Henrissat B: An evolving hierarchical family classification for glycosyltransferases. J. Mol. Biol. 2003, 328:307-317

22. Madson M, Dunand C, Li X, Verma R, Vanzin GF, Caplan J, Shoue DA, Carpita NC, Reiter WD: The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. The Plant cell 2003, 15(7):1662-1670.

23. Jensen JK, Schultink A, Keegstra K, Wilkerson CG, Pauly M: RNA-Seq analysis of developing nasturtium seeds (Tropaeolum majus): identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis. Molecular plant 2012, 5(5):984-992.

24. Schultink A, Cheng K, Park YB, Cosgrove DJ, Pauly M: The identification of two arabinosyltransferases from tomato reveals functional equivalency of xyloglucan side chain substituents. Plant physiology 2013, 163(1):86-94.

25. Gunl M, Kraemer F, Pauly M: Oligosaccharide mass profiling (OLIMP) of cell wall polysaccharides by MALDI-TOF/MS. Methods Mol Biol 2011, 715:43-54.

26. Kooiman P: The constitution of Tamarindus-amyloid. Recl. Trav. Chim. Pays-Bas 1961, 80: 849-65.

27. Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y et al: The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science 2008, 319(5859):64-69.

28. Vanzin GF, Madson M, Carpita NC, Raikhel NV, Keegstra K, Reiter WD: The mur2 mutant of Arabidopsis thaliana lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. Proceedings of the National Academy of Sciences of the United States of America 2002, 99(5):3340-3345.

29. Gille S, Pauly M: O-acetylation of plant cell wall polysaccharides. Frontiers in plant science 2012, 3:12.

30. Kong Y, Pena MJ, Renna L, Avci U, Pattathil S, Tuomivaara ST, Li X, Reiter WD, Brandizzi F, Hahn MG et al: Galactose-depleted xyloglucan is dysfunctional and leads to dwarfism in Arabidopsis. Plant physiology 2015, 167(4):1296-1306.

31. Buckeridge MS: Seed cell wall storage polysaccharides: models to understand cell wall biosynthesis and degradation. Plant physiology 2010, 154(3):1017-1023.

32. Gibeaut DM, Pauly M, Bacic A, Fincher GB: Changes in cell wall polysaccharides in developing barley (Hordeum vulgare) coleoptiles. Planta 2005, 221(5):729-738.

33. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M et al: Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 2008, 36(Web Server issue): W465-469.

34. Dereeper A, Audic S, Claverie JM, Blanc G: BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC evolutionary biology 2010, 10:8.

35. Clough SJ, Bent AF: Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant journal: for cell and molecular biology 1998, 16(6):735-743.

36. Pauly M, Andersen LN, Kauppinen S, Kofod LV, York WS, Albersheim P, Darvill A: A xyloglucan-specific endo-beta-1,4-glucanase from Aspergillus aculeatus: expression cloning in yeast, purification and characterization of the recombinant enzyme. Glycobiology 1999, 9(1):93-100.

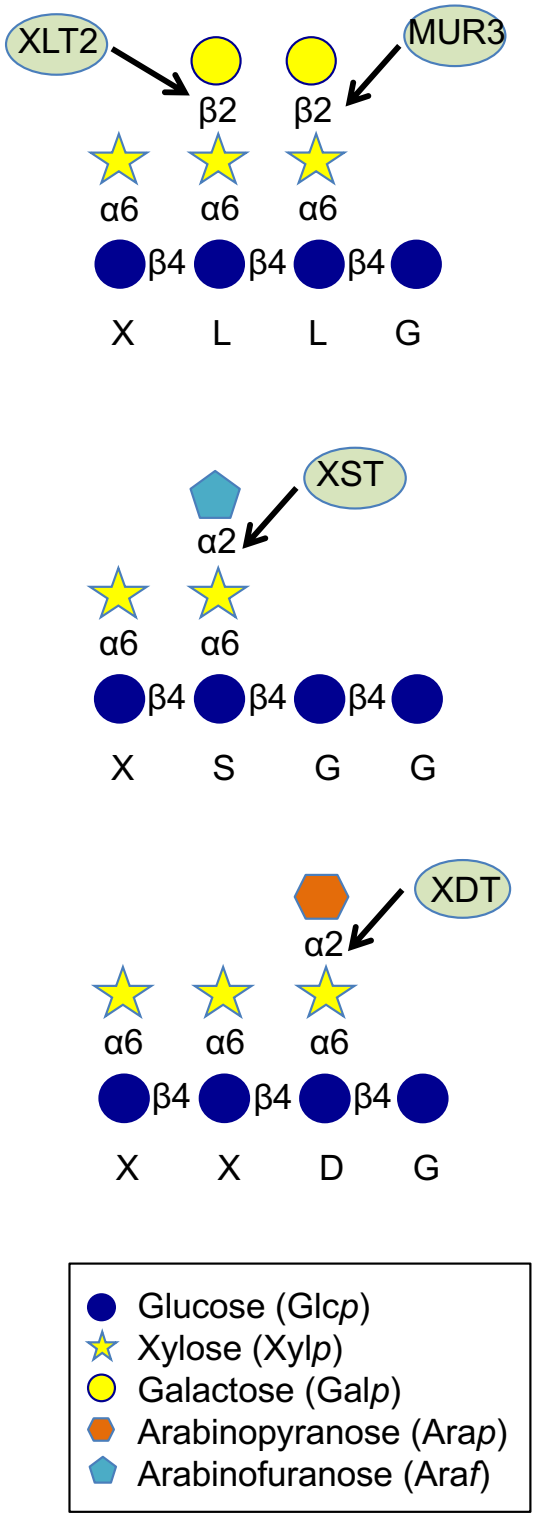


Figure 1. Xyloglucan oligosaccharide structures and GT47 glycosyltransferases involved in its synthesis.

The xyloglucan one-letter code nomenclature is indicated below the structure. MUR3 and XLT2 - XyG:galactosyltransferases; XST - XyG:arabinofuranosyltransferase; XDT – XyG:arabinopyranosyltransferase

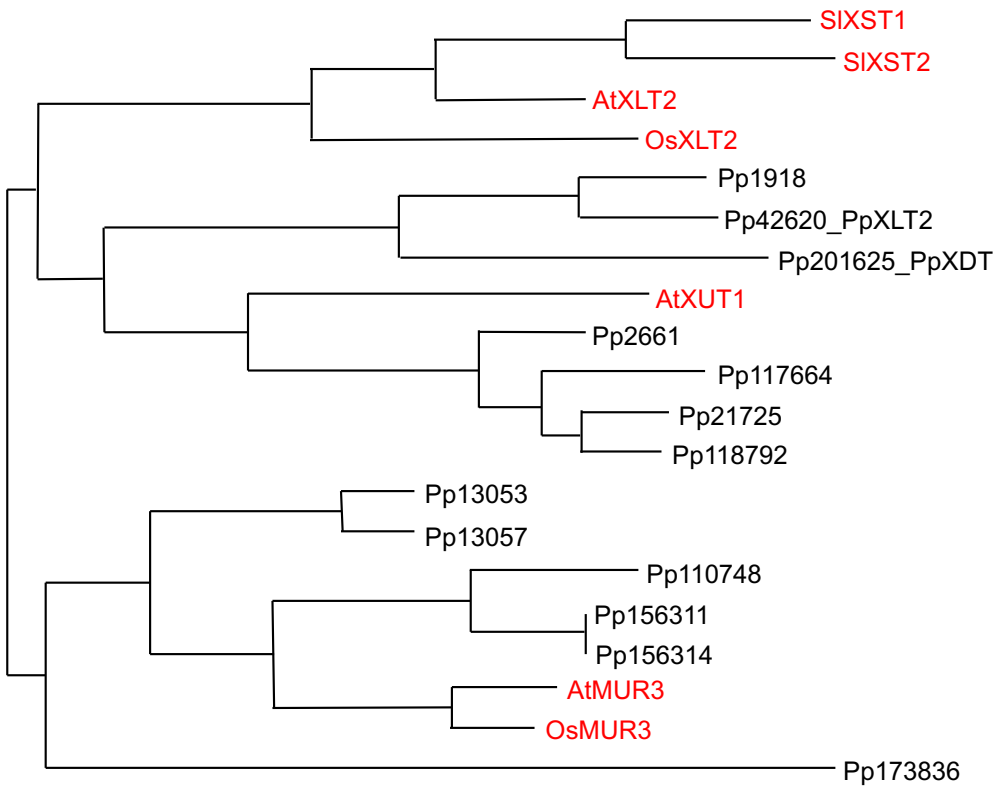


Figure 2. Phylogeny of XyG-related GT47 proteins. Red font - Protein sequences from known XyG:galactosyltransferases AtMUR3, OsMUR3, AtXLT2 and OsXLT2, the galacturonosyltransferase AtXUT, and the arabinofuranosyltransferases SlXST1 and SlXST2. Black font - *Physcomitrella patens* proteins obtained from a BlastP search against Phytozome database. Phylogenetic tree built by PhyML.

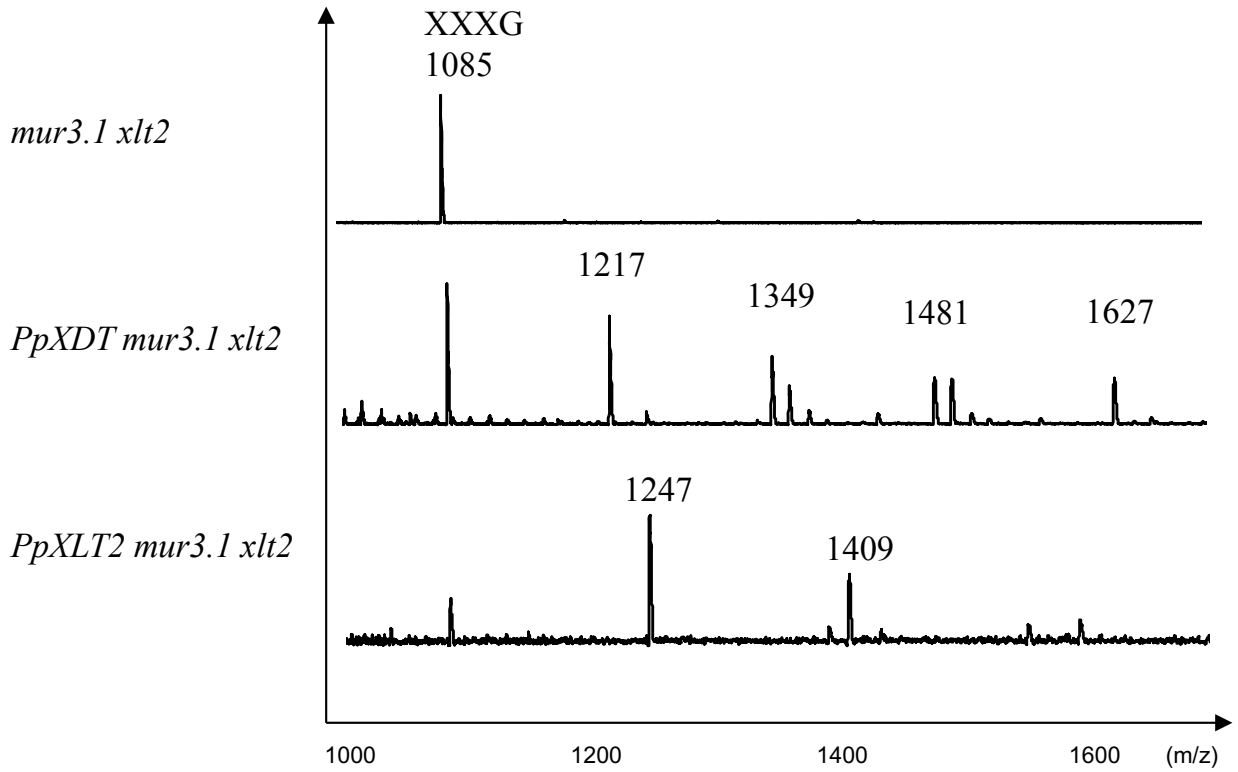


Figure 3. XyG oligosaccharide mass profiling by MALDI-TOF MS. XyG oligosaccharides derived from leaf tissue of the Arabidopsis double mutant *mur3.1 xlt2*, transgenic lines expressing *PpXDT* or *PpXLT2* in *mur3.1 xlt2*. Numbers indicate m/z and their potential structure is shown in supplemental table 1. m/z 1085 represents the known XyG oligosaccharide structure XXXG.

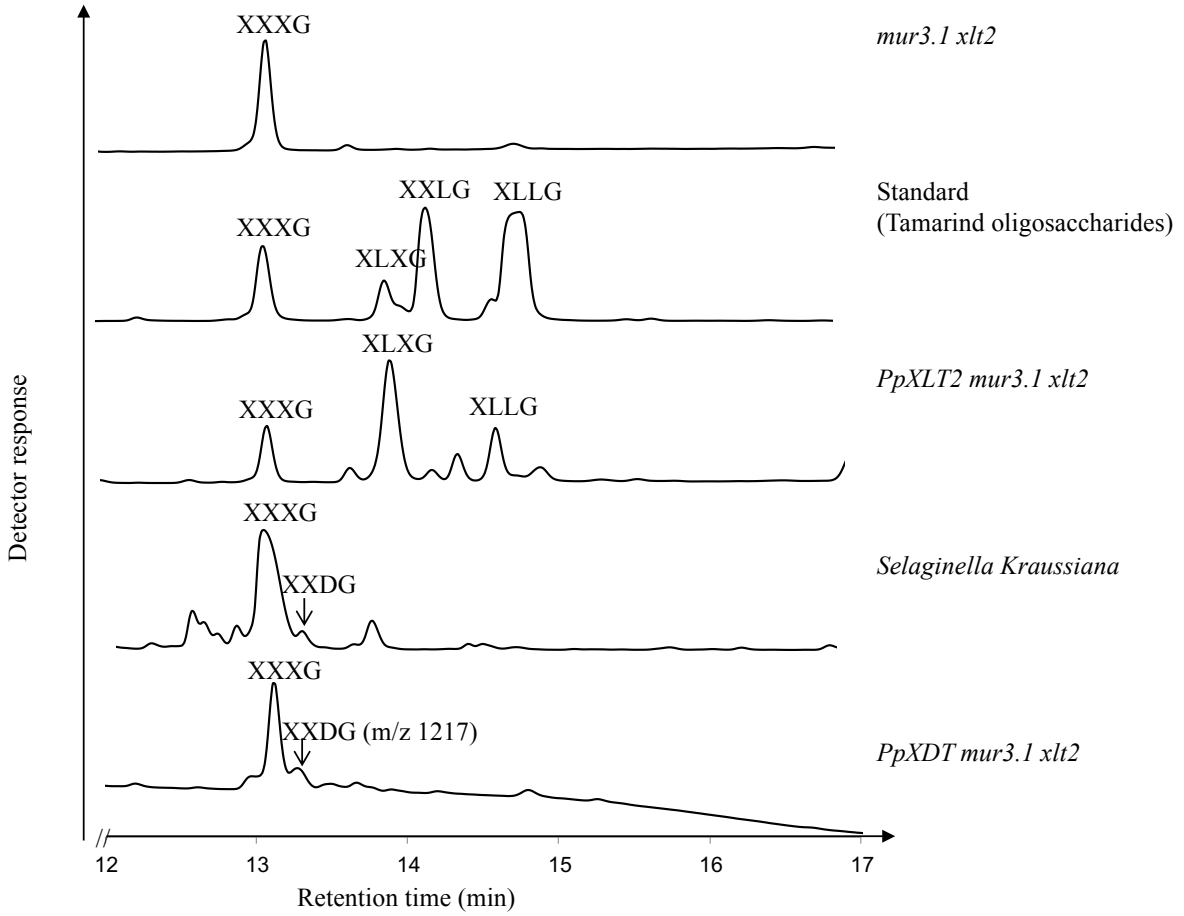


Figure 4. XyG oligosaccharide separation by HPAEC-PAD. Peaks were assigned based on retention times from published work (Schultink et al., 2013; Hsieh & Harris, 2012; Megazyme Inc., Ireland) as well as assignment based on mass spectrometry. The fractions containing m/z 1217 (XXDG) was collected and further analyzed by glycosidic linkage analysis (Table 1) and NMR spectroscopy (Fig. 5).

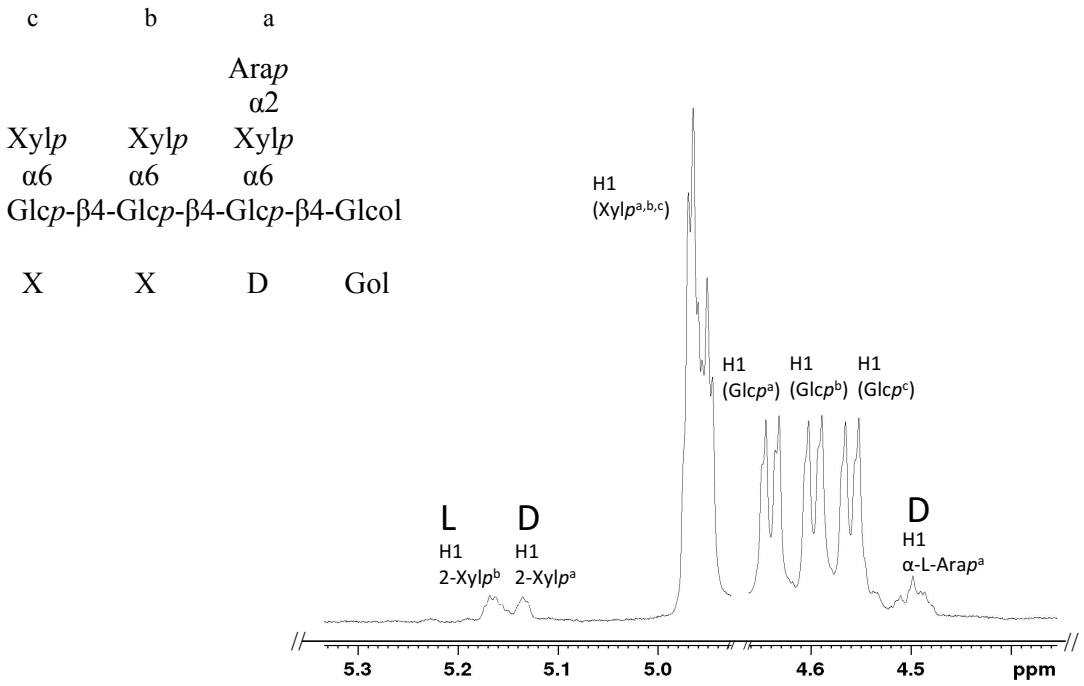


Figure 5. Anomeric region of the ^{1}H NMR spectra of the reduced XyG oligosaccharide XXDGol contaminated with XXXGol. ^{1}H NMR peaks were labelled according to CCRC (ccrc.uga.edu) database and the literature (Pena et al., 2008). The putative structure of the oligosaccharide is shown in the upper left corner indicating the order of the side-chains on top and the one-letter code below.

A)

Wild type | *mur3.1 xlt2* | *PpXLT2 mur3.1 xlt2* | *PpXDT mur3.1 xlt2*

B)

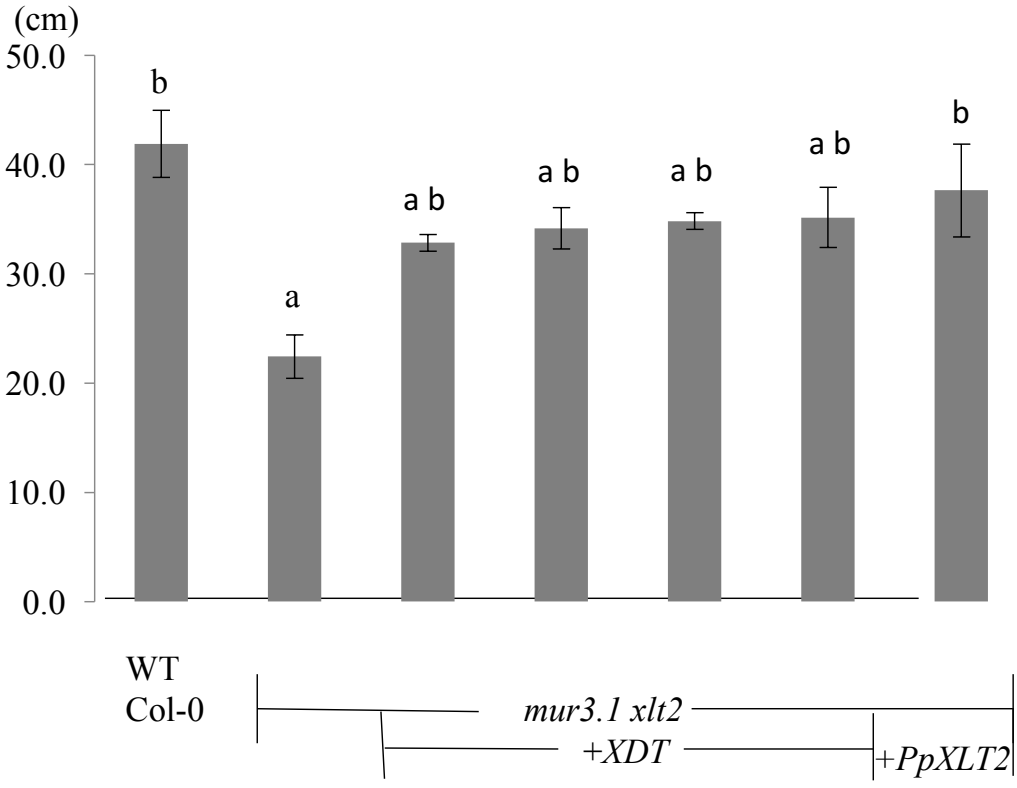
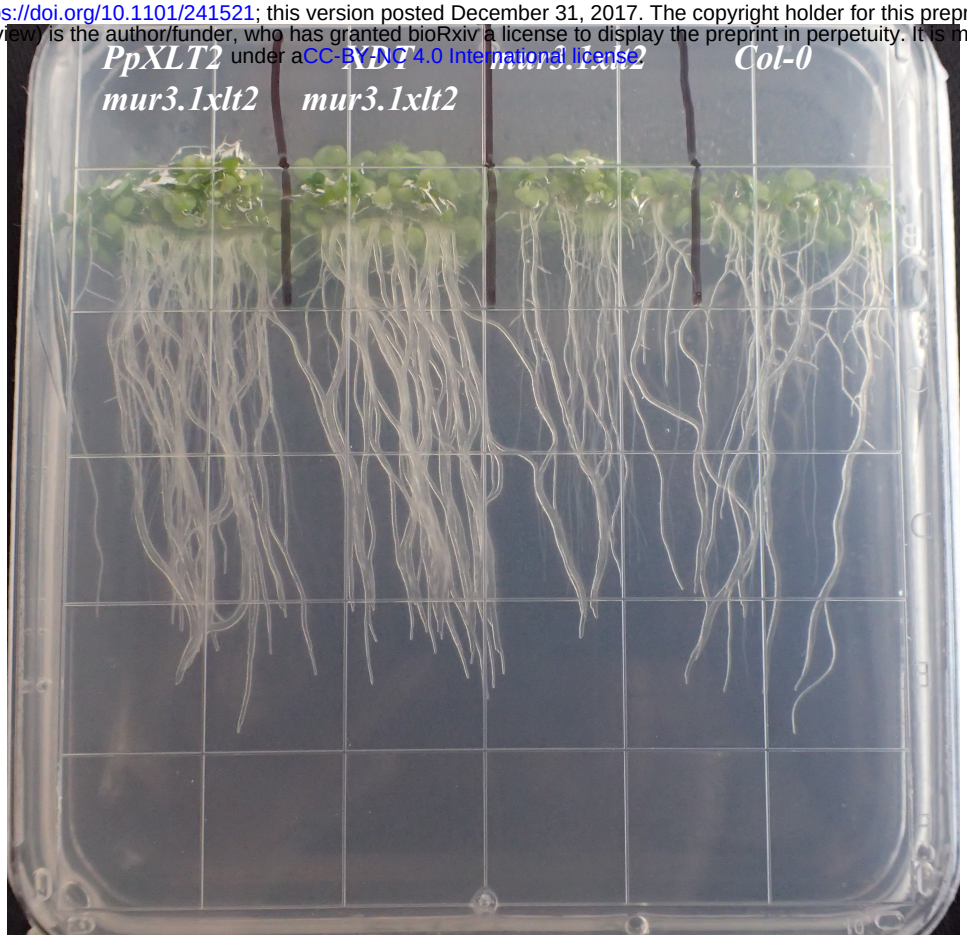


Figure 6. A) Growth habit of 8-week old Arabidopsis plants. B) Height of Inflorescence stems of 8-week old Arabidopsis plants. n≥3, Small letters – ANOVA analysis.

A)



B)

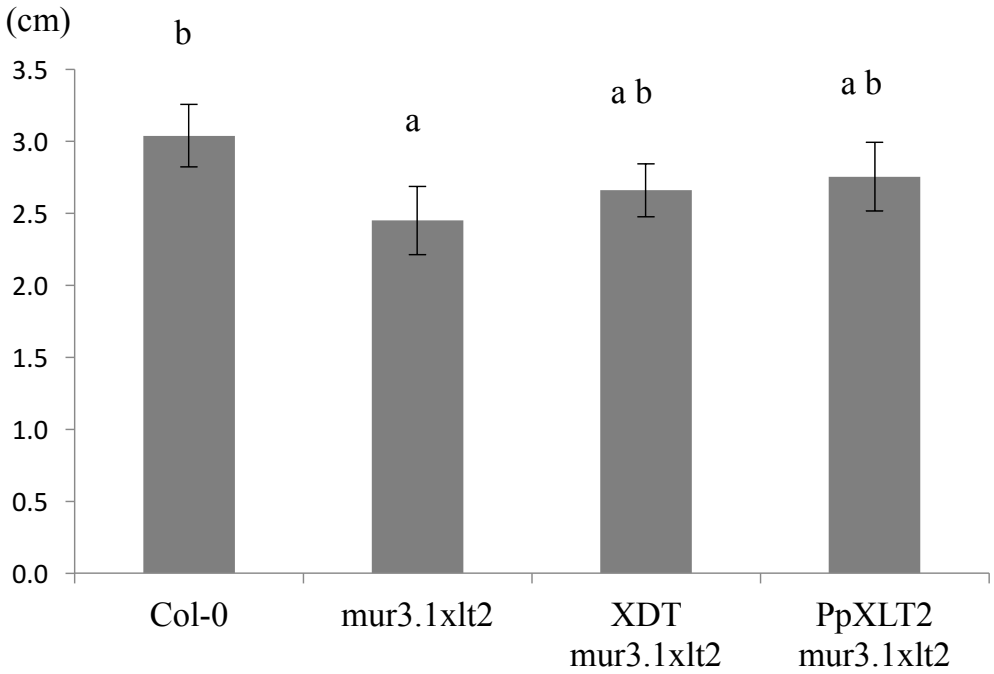


Figure 7. Root growth of Arabidopsis plants. A) Root habit B) Root length of 8-day-old seedlings, n=10. Small letters – ANOVA analysis.

Table S1. The XyG oligosaccharide composition and relative abundance of ions (m/z) seen in Fig. 3.

	Ions (m/z)	Oligosaccharide composition	Predicted Oligosaccharide Structure [a]
PpXDT mur3.1 xlt2	1085	H4P3	XXXG
	1217	H4P4	XXDG/XDXG
	1349	H4P5	XDDG
	1363	$H4P4H_{de}1$	XXEG/XEXG
	1481	H4P6	DDDG
	1495	$H4P5H_{de}1$	XDEG/XEDG
	1627	$H4P6H_{de}1$	DDEG/DEDG
PpXLT2 mur3.1 xlt2	1085	H4P3	XXXG
	1247	H5P3	XXLG/XLXG
	1409	H6P3	XLLG
	1555	$H6P3H_{de}1$	XLFG
	1597	$H6P3H_{de}1Ac1$	XLFG-Ac

P - pentose; H - hexose; H_{de}, - deoxy hexose; Ac - acetate.

[a] Assuming that PpXDT only transfers arabinopyranosyl-residues

Table S2. Primers for PCR amplification of genes and quantitative RT-PCR.

Primers	Sequences
Pp_42620_exon-F-Gibson	5'- AGA ATT CGT CGA CTT TGC ATG
Pp_42620_exon-R-Gibson	5'- AGT AAA AGG TAC CGA GCT TCA
Pp_201625_CDS_Gibson-F	5'- AGG GAT ATC ACT AGT CAA CAA
Pp_201625_CDS_Gibson-R	5'- CAA AAC CCA CCG GAT ATG GGG
Pp201625_CDS-R	5'- TCA ACA ACC ATC CGT GAC CTT
Pp201625_CDS-F	5'- ATG GGG TAT GCA CCG CAA T
GAPDH REV Set 2	5'- ACG GTT GGA ACA CGG AAA GAC A
GAPDH FWD Set 2	5'- ATG AAG GAC TGG AGA GGT GGA A
Pp11_copy# FWD Set 1	5'- TTC GTC CGC TCG GTG GGT AAT TT
Pp11_copy# REV Set 1	5'- GTT CTC AAC GGC CGA GAT CCA TT
Pp12_copy# FWD Set 5	5'- TGC ATC CGA CCG AAG TCG AAA C
Pp12_copy# REV Set 5	5'- ACA GAG CTC CGT GGA GCA AAT G

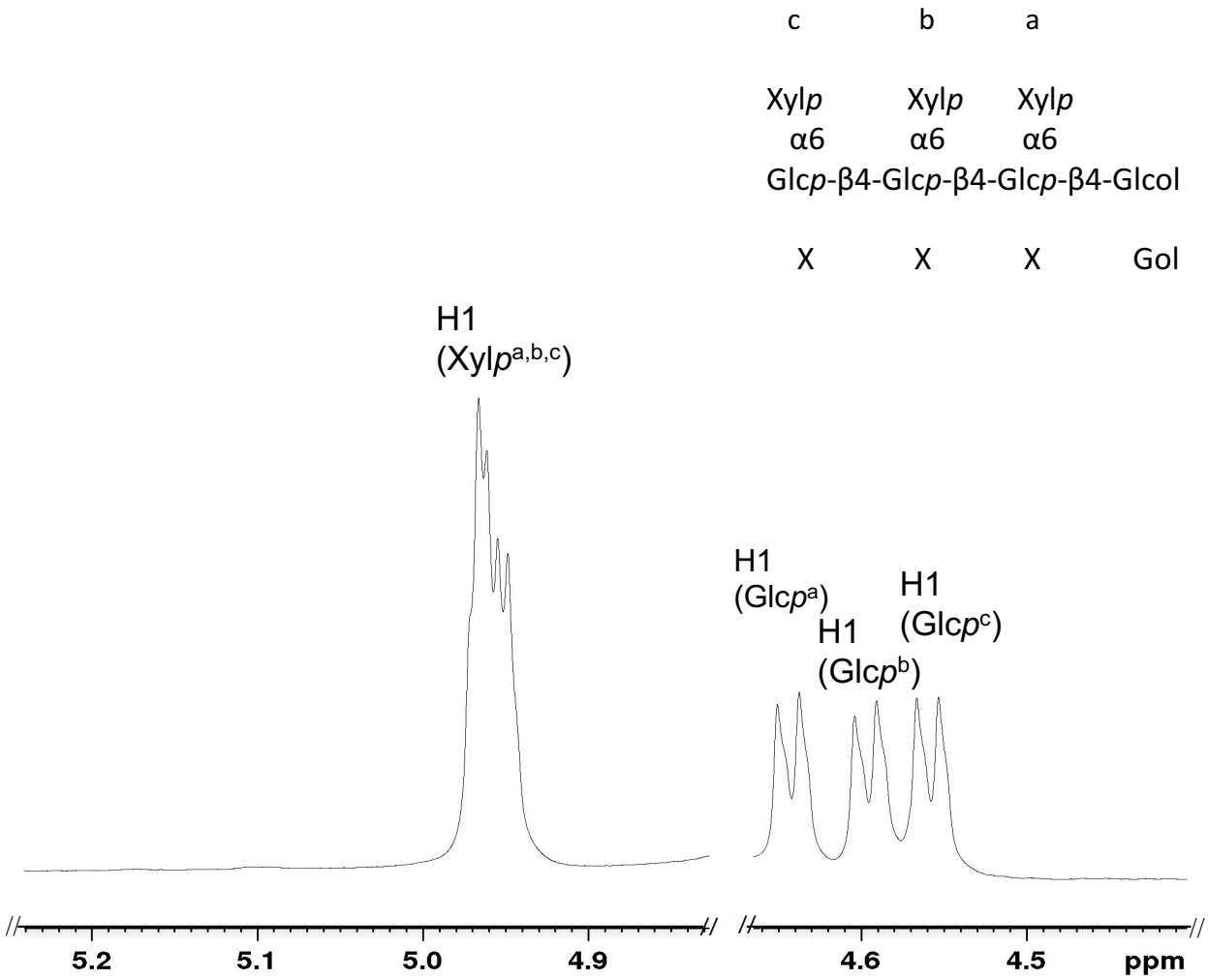


Figure S1. Anomeric region of the ^{1}H NMR spectra of the oligosaccharide XXXG obtained from *mur3.1 xlt2*. ^{1}H NMR peaks were labelled according to CCRC (ccrc.uga.edu) database. The putative structure of the oligosaccharide is shown in the upper right corner indicating the order of the side-chains on top and the one-letter code below.