New Phytologist

Two conifer GUX clades are responsible for distinct glucuronic acid patterns on xylan

Jan J. Lyczakowski[1,2], Li Yu[2], Oliver M. Terrett[2], Christina Fleischmann[3], Henry Temple[2], Glenn Thorlby[3], Mathias Sorieul[3] and Paul Dupree[2]

[1]Department of Plant Biotechnology, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Gronostajowa 7, Krakow 30-387, Poland; [2]Department of Biochemistry, University of Cambridge, Tennis Court Road Cambridge, CB2 1QW, UK; [3]Scion, 49 Sala Street, Rotorua 3020, New Zealand

Authors for correspondence:
Jan J. Lyczakowski
Email: jan.lyczakowski@uj.edu.pl

Paul Dupree
Email: pd101@cam.ac.uk

Received: *19 February 2021*
Accepted: *21 May 2021*

New Phytologist (2021) **231:** 1720–1733
doi: 10.1111/nph.17531

Key words: conifers, glucuronic acid (GlcA), plant cell walls, softwood, xylan.

Summary

- Wood of coniferous trees (softwood), is a globally significant carbon sink and an important source of biomass. Despite that, little is known about the genetic basis of softwood cell wall biosynthesis. Branching of xylan, one of the main hemicelluloses in softwood secondary cell walls, with glucuronic acid (GlcA) is critical for biomass recalcitrance. Here, we investigate the decoration patterns of xylan by conifer GlucUronic acid substitution of Xylan (GUX) enzymes.
- Through molecular phylogenetics we identify two distinct conifer *GUX* clades. Using transcriptional profiling we show that the genes are preferentially expressed in secondary cell wall forming tissues. With *in vitro* and *in planta* assays we demonstrate that conifer GUX enzymes from both clades are active glucuronyltransferases.
- Conifer GUX enzymes from each clade have different specific activities. While members of clade one add evenly spaced GlcA branches, the members of clade two are also capable of glucuronidating two consecutive xyloses. Importantly, these types of xylan patterning are present in softwood.
- As xylan patterning might modulate xylan–cellulose and xylan–lignin interactions, our results further the understanding of softwood cell wall biosynthesis and provide breeding or genetic engineering targets that can be used to modify softwood properties.

Introduction

The majority of terrestrial carbon is stored in forests as wood (Pan *et al.*, 2011; Ramage *et al.*, 2017; Bar-On *et al.*, 2018). Plantation forests for commercial timber production are mainly growing coniferous species. For example, conifers account for as much as 98% of trees planted in the US managed forests (Sedjo, 2001). The abundance and sustainability of softwood position it as excellent feedstocks for use in alternative industrial processes, tackling the climate emergency. Unfortunately, progress in conifer breeding and genetic engineering aiming to improve softwood suitability for industrial application is hindered by the poor understanding of the genetic basis for its biosynthesis.

Softwood consists mainly of tracheid cells (Ramage *et al.*, 2017). Their secondary cell walls are composed of cellulose microfibrils embedded in a matrix of hemicellulose polysaccharides and hydrophobic lignin. Cellulose is believed to be the main load-bearing structure of softwood cell walls. The microfibrils are a crystalline assembly of 18 or 24 glucan chains (Fernandes *et al.*, 2011; Jarvis, 2018). In conifers, lignin is mainly made out of guaiacyl monomers. It impregnates softwood secondary cell walls and enables water transport in trees (Vanholme *et al.*, 2010). The two main types of hemicelluloses; galactoglucomannan (GGM) and arabinoglucuronoxylan (AGX), account for 10–30% and 5–15% of softwood secondary cell walls, respectively (Scheller & Ulvskov, 2010). The GGM backbone consists of β-1,4-linked mannose and glucose monomers. Some mannose subunits are branched with α-1,6-linked galactose or acetyl substituents linked to carbon 2 or 3 (Willför *et al.*, 2008). The AGX backbone is composed of β-1,4-linked xylose monomers with α-1,2-linked 4-*O*-methyl glucuronic acid (MeGlcA) and α-1,3-linked arabinofuranose (Ara) branches (Timell & Syracuse, 1967; Willför *et al.*, 2005).

The exact biological functions of conifer hemicelluloses are still debated. However, they are likely to mediate cross-linking of different cell wall components and modulate the mechanical properties of the cell wall (Terrett & Dupree, 2019; Berglund *et al.*, 2020a,b; Martínez-Abad *et al.*, 2020). Both *in vitro* binding assays (Whitney *et al.*, 1998; Berglund *et al.*, 2020a,b) and *in planta* analysis (Terashima *et al.*, 2009; Yu *et al.*, 2018; Terrett *et al.*, 2019) suggest that GGM can interact with cellulose microfibrils. Recent two-dimensional solid-state nuclear magnetic resonance (ssNMR) analysis of spruce wood demonstrated that GGM may directly interact with cellulose *in planta* and extend its crystalline structure (Terrett *et al.*, 2019). Interestingly, the same study indicated that a majority of AGX can also directly bind to cellulose microfibrils. In addition to their interaction with cellulose, hemicelluloses can also bind lignin (Terrett &

Dupree, 2019). Both AGX and GGM are located < 1 nm away from lignin in spruce wood (Terrett *et al.*, 2019). GGM is likely to form direct covalent linkages with lignin (Nishimura *et al.*, 2018). In softwood, the glucuronic acid sidechains of AGX were proposed to form ester linkages with lignin (Oinonen *et al.*, 2015). However, this xylan–lignin complex was never clearly identified. In Arabidopsis, the presence of GlcA on xylan has been proven to be critical for the maintenance of plant biomass resistance to enzymatic degradation (Lyczakowski *et al.*, 2017). Therefore, this plausible xylan–lignin complex may be important for industrially relevant properties of softwood. Similar principles may extend to other taxa, as both in hardwood poplar (Addison *et al.*, 2020) and in grass species (Kang *et al.*, 2019) lignin was shown, using solid-state NMR, to locate within bonding distance to cell wall polysaccharides.

To understand how xylan contributes to the softwood cell wall properties it is necessary to fully characterise its structure. The addition of the GlcA onto the xylan backbone is carried by GlucUronic acid substitution of Xylan (GUX) enzymes (Mortimer *et al.*, 2010; Rennie *et al.*, 2012). In Arabidopsis, an even pattern of xylan branches allows the hemicellulose to bind onto the hydrophilic faces of cellulose microfibrils via hydrogen bonds (Busse-Wicher *et al.*, 2014; Busse-Wicher *et al.*, 2016a; Grantham *et al.*, 2017). In Arabidopsis the addition of evenly spaced GlcA side branches is catalysed by GUX1 enzymes (Bromley *et al.*, 2013). In addition to this even patterning, some of the GlcA decorations are distributed with uneven spacing. This generates a xylan domain unable to bind in the same way to the hydrophilic surface of the cellulose microfibrils. This incompatible patterning, in Arabidopsis, is laid out by GUX2 enzymes (Bromley *et al.*, 2013).

In conifers (pine and Douglas fir), the GlcA layout on xylan was reported to be simpler than in Arabidopsis (Busse-Wicher *et al.*, 2016b). It consists primarily of an even spacing of MeGlcA branches every six xylose units. In addition to the MeGlcA, conifer xylan also has Ara decorations. These Ara branches are placed predominantly with an even spacing of two xylose units from the MeGlcA. A spacing equal to six xylose units between MeGlcA branches was confirmed also in spruce (Martínez-Abad *et al.*, 2017). Interestingly, the presence of a second pattern consisting in MeGlcA branches on two consecutive xyloses along the AGX backbone was also reported in larch, two cypress species and in spruce (Shimizu *et al.*, 1978; Yamasaki *et al.*, 2011; Martínez-Abad *et al.*, 2017). The biological functions of these two types of xylan patterning in conifers and the enzymes responsible for their biosynthesis are unknown. It is however possible, that similarly to angiosperms, the patterning of gymnosperm xylan with GlcA may be important for the control of its interaction with the cellulose microfibril.

Here, we report the discovery of a second conifer GUX clade, first clade was described by Lyczakowski *et al.*, 2017, and the characterisation of the biosynthetic activities of the two clades identified. The two GUX enzymes have distinct specific activities enabling them to lay out different patterns of GlcA branches. Our results demonstrate the biosynthetic role of GUX enzymes for the patterning of xylan in softwoods. Since xylan patterning is likely to be important for cell wall properties, our results should influence future genetic engineering or breeding approaches aimed at improving softwood suitability for applications in the bioeconomy.

Materials and Methods

Sampling, plant material, growth and Arabidopsis transformation

Arabidopsis thaliana (*At*) plants of the Columbia-0 ecotype were grown in a cabinet maintained at 21°C, with a 16 h : 8 h, light : dark photoperiod. Two previously described *A. thaliana* insertion lines *gux1/2/3* (Mortimer *et al.*, 2015) and *esk-kak* (Grantham *et al.*, 2017) were used for transformation experiments. For *Pinus radiata* (*Pr*) GUX1 and *Pr*GUX2 expression profiling, *Pinus radiata was* grown by Scion, New Zealand. For biomass analysis, alcohol insoluble residues (AIR) were prepared as previously described (Mortimer *et al.*, 2010) from three individual biological replicates of plants with each biological replicate consisting of a pooled sample of 36 plants. Constructs used for *Pinus taeda* (*Pt*) GUX2 overexpression in *gux1/2/3 A. thaliana* and for *Picea glauca* (*Pg*) GUX1 and *Pt*GUX2 overexpression in *esk-kak* were prepared using OpenPlant common syntax GoldenGate assembly (Patron *et al.*, 2015). Arabidopsis transformation was performed using the floral dip method as previously described (Clough & Bent, 1998).

Phylogenetic and bioinformatic analysis

Coding sequences of *Arabidopsis thaliana* (*At*) GUX and *Populus trichocarpa* (*Ptr*) GUX (Kumar *et al.*, 2019) were used in phylogenetic analysis. *At*GUX1, *At*GUX2 and *At*GUX3 were used to identify putative GUX encoding transcripts from Coniferophyta and Gnetophyta transcriptomic data available from 1000 Plant Transcriptomes, Congenie and Gymno Plaza BLAST services (Nystedt *et al.*, 2013; Matasci *et al.*, 2014; Proost *et al.*, 2015). In-house *Pinus radiata* genome and transcriptome resources (Scion, New Zealand) were used to recover full length DNA sequences for *PrGUX1* and *PrGUX2* genes. The sequences were confirmed by amplification and sequencing from genomic DNA and primers for droplet digital (dd) PCR analysis were designed (Supporting Information Table S1). All amino acid sequences were reconstructed from transcripts with ExPASy translate tool. A maximum likelihood phylogenetic tree was constructed using MEGA6 software (Tamura *et al.*, 2013) including the gamma distribution algorithm in the reconstruction. Table S2 summarises codes required to access gymnosperm GUX sequences studied.

For the computational expression profiling of *Picea abies* (*Pa*) GUX enzymes Congenie (Nystedt *et al.*, 2013) and Norwood (Jokipii-Lukkari *et al.*, 2017) bioinformatics tools were used. The putative cellulose, mannan, xylan and lignin biosynthesis genes were identified using Congenie BLAST service using *A. thaliana* protein sequences as a query.

Pinus radiata RNA isolation and cDNA synthesis

Total RNA was extracted from six different tissue types (root, cambium/xylem, phloem, vegetative bud, needles and green stem). Samples were collected from four individual 5-yr-old clonal *Pinus radiata* trees growing in Rotorua, New Zealand. For the collection of developing xylem and phloem, a section of bark was scored and peeled away from the stem. The exposed cambium and youngest layers of xylem were collected by scraping and the phloem peeled from the inside of the bark. All samples were snap frozen and stored at −80°C. The RNA extractions were performed on 50 mg of plant material from each tissue type using the Plant RNA Isolation Mini Kit (Agilent, Santa Clara, CA, USA). Concentration and purity of total RNA was determined fluorometrically (Qubit, Invitrogen) and the integrity of the RNA samples was verified using 1.2% agarose formaldehyde gels. Reverse transcription (RT) was carried out using the qScript XLT cDNA Super Mix Kit (Quantabio, Beverly, MA, USA) with 500 ng of DNaseI-treated total RNA as a template.

ddPCR analysis of *PrGUX1* and *PrGUX2* expression

The ddPCRs were performed using the Bio-Rad QX200 system (Bio-Rad, Hercules, CA, USA). The ddPCR reactions contained 11 µl of ddPCR EvaGreen Supermix (Bio-Rad), primers at a final concentration of 150 nM each, 2 µl of a 1 : 30 cDNA template dilution and ultrapure water to a final volume of 22 µl. Then, 20 µl of the ddPCR reactions along with 70 µl of droplet generation oil for EvaGreen (Bio-Rad) were loaded onto an 8-channel DG8TM cartridge for droplet generation in the QX200 Droplet Generator (Bio-Rad). Droplets were transferred into a 96-well plate (Bio-Rad) and a two-step thermocycling protocol ((enzyme activation at 95°C × 5 min); 40 cycles × (denaturation at 95°C × 30 s, annealing/extension at 60°C × 60 s); signal stabilisation at 4°C × 5 min and 90°C × 5 min, ramp rate set at 2°C s^{-1}) was performed in a thermal cycler (Bio-Rad C1000 Touch). Droplets were quantified in a QX200 Droplet Reader (Bio-Rad) and results were analysed with QuantaSoft™ Analysis Pro Software v.1.0.596. Statistical analysis was performed using R. Estimated marginal means were shown with bars representing the 95% confidence intervals. Four biological replicates were used for all tissue types. No amplification was observed from RNA-only controls (no RT).

Nicotiana benthamiana expression and Western blot analysis

Blunt ended *myc*-tagged GUX and *At*MAGT1 CDSs were amplified from a synthetic construct (synthesised by Genewiz) with Q5 DNA Polymerase (NEB) using primers detailed in Table S1. The PCR product was ligated into *Nru*I (NEB) digested pEAQ-HT *Nicotiana benthamiana* overexpression vector (Sainsbury *et al.*, 2009) using T4 DNA ligase (Thermo-Fisher Scientific). For overexpression in tobacco AGL-1 *Agrobacterium* was used and a previously described infiltration protocol was followed (Sparkes *et al.*, 2006). The membrane collection protocol was adapted after Rennie *et al.* (2012) and performed as described in Lyczakowski *et al.* (2017). Protein concentration in the membranes fraction was quantified using modified Bradford reagent (Expedeon). Western blot analysis, was performed as described in Lyczakowski *et al.* (2017) with anti-*myc* antibody (rabbit; Santa-Cruz, A14) and anti-rabbit HRP linked antibody (mouse; Bio-Rad, 170-6515).

In vitro activity assay for GUX

For each reaction dried aliquots of acetylated heteroxylan (extracted following a protocol described by Busse-Wicher *et al.*, 2014) were resuspended in 30 µl of a reaction master-mix (0.5 mM DTT, 10 mM $MnCl_2$, 10 mM $MgCl_2$, 2% Triton X-100, 10 mM UDP-GlcA). UDP-GlcA was replaced with water in certain reactions to control for nonspecific glucuronidation. The reaction mix was amended with 30 µl of undiluted microsomal proteins extracted from *N. benthamiana* leaves. For most experiments the reaction was performed for 5 h at room temperature and terminated with a 10-min-long 100°C heat treatment. For reactions including acetylxylan esterase (Acetylxylan esterase CE4 from *Clostridium thermocellum*, NzyTech CZ00341) 5 µl (0.2 U) of enzyme suspension was added to the reaction following 3 h from its start. Additional UDP-GlcA (U6751; Sigma-Aldrich) was introduced into the reaction upon addition of the esterase enzyme, leading to doubling of the initial concentration in the reaction tube from 5 mM to 10 mM. Reactions performed in the presence of the esterase enzyme were carried out overnight. Following completion of the reaction, the polysaccharides products were extracted using a methanol : chloroform lipid removal protocol adapted from Bligh & Dyer (1959).

Carbohydrate digestions, polysaccharide analysis by gel electrophoresis (PACE) and saccharification

De-acetylated AIR (0.25 mg) or *in vitro* reaction products were digested with *Neocallimastix patriciarum* GH11 enzyme overnight at 30°C by amending the suspension with 2 µl (20 U) of enzyme stock (Megazyme). These conditions achieved complete digestion of accessible xylan to xylose, xylobiose and glucuronidated oligosaccharides. For GH30 digestion 1 µl of 1 mg ml^{-1} purified *Erwinia chrysanthemi* (*Ec*) enzyme was applied onto *in vitro* reaction products for 1 h at room temperature. *Ec*GH30 used in this work was a kind gift from Dr Kristian Krogh (Novozymes A/S, Denmark). Fungal GH115 glucuronidase used in this work was also received from Dr Kristian Krogh. For de-glucuronidation, 1 µl of 1 mg ml^{-1} fungal GH115 was added to oligosaccharides resuspended in 0.1 M ammonium acetate pH 5.5 buffer and incubated overnight at 55°C. Before GH115 treatment, the GH11 enzyme was heat inactivated by incubating the oligosaccharide sample at 120°C for 10 min. Following digestion the oligosaccharides were analysed with PACE as previously described (Mortimer *et al.*, 2010). PACE band intensity was performed using ImageJ software and the degree of glucuronidation was quantified as described by Mortimer (2017). All saccharification experiments were

performed on AIR as previously described (Lyczakowski *et al.*, 2017) without any pretreatment of the feedstock.

Results

Conifer genomes encode two distinct GUX homologues

In our previous work we identified a single conifer GUX enzyme clade and characterised the *in vitro* activity of its representative from *Picea glauca* (Lyczakowski *et al.*, 2017). In dicots, two clades of GUX enzymes produce specific GlcA distribution patterns on xylan backbone (Bromley *et al.*, 2013). Therefore, the presence of two distinct GlcA patterns on gymnosperm xylan (Shimizu *et al.*, 1978; Yamasaki *et al.*, 2011; Martínez-Abad *et al.*, 2017) prompted us to investigated the presence of additional clades of GUX enzymes within this taxonomic group. We used Arabidopsis GUX1, GUX2 and GUX3 sequences to search for putative gymnosperm xylan glucuronyltransferases. A highly conserved region, containing the GT8 domain, was observed in all gymnosperm protein sequences that had high similarity with Arabidopsis GUX1, GUX2 and GUX3 and a maximum likelihood phylogenetic tree was constructed using these amino acid sequences (Fig. 1). The phylogenetic analysis indicated that, similarly to eudicots, conifers have two distinct clades of putative GUX enzymes. Sequences from the previously characterised conifer clade (Lyczakowski *et al.*, 2017) were observed to be more similar to eudicot GUX1 and GUX3 sequences. The newly identified clade of putative conifer GUX enzymes was more similar to eudicot GUX2/4/5 clade than to the other conifer clade. We labelled the conifer sequences as members of GUX1 and GUX2 clades.

Conifer GUX1 and GUX2 enzymes are likely to be involved in the wood formation process

To determine the possible biological function of GUX1 and GUX2 in conifers, we first analysed their expression profiles. We started with information available in public databases such as Congenie (Nystedt *et al.*, 2013) and Norwood (Jokipii-Lukkari *et al.*, 2017). The Congenie exHeatmap tool indicates that both *Picea abies* (*Pa*) GUX1 and *Pa*GUX2 mRNAs are enriched in phloem and xylem (Figs S1, S2). Developing vegetative shoots also had high expression levels. The Norwood database provides a more specific analysis by allowing the mapping of conifer gene expression across different elements of the stem, from cambium to mature tracheids. Therefore, to further evaluate the likely biological role of putative conifer GUX enzymes, expression of *Pa*GUX1 and *Pa*GUX2, together with the expression of other putative cell wall biosynthesis genes, such as putative cellulose synthase complex members (CesAs) or cellulose synthase like A (CslA) family members, was studied using the Norwood tool (Fig. S3). Most of the studied cell wall biosynthesis genes shared a similar high expression profile predominantly in the developing tracheids. This expression profile, which can be associated with gene function in secondary cell wall biosynthesis, was also observed for both *Pa*GUX1 and *Pa*GUX2.

To validate our *in silico* results and to extend expression profiling to another conifer species, we performed an absolute quantification of conifer *GUX1* and *GUX2* expression in four individual clonal *Pinus radiata* (*Pr*) trees using droplet digital (dd) PCR (Fig. 2). We studied the expression of the genes in xylem/cambium, phloem, root, vegetative bud, green stem and needles. Expression of both *PrGUX1* and *PrGUX2* was detected in all tissues tested, with the exception of phloem, which had only trace amounts of *PrGUX2* mRNA. In line with the *in silico* analysis for *Picea abies*, we have detected strong *PrGUX1* and *PrGUX2* expression in green stems and needles. The strongest expression detected for both *PrGUX1* and *PrGUX2* was in xylem/cambium. In xylem, as in all tissues tested, the expression of *PrGUX1* was six to eleven times higher than that of *PrGUX2*. This suggests that *Pr*GUX1 may be responsible for the glucuronidation of the majority of the xylan in this species. In summary, together with the Congenie and Norwood based analyses, our absolute quantification results support the hypothesis that putative conifer GUX enzymes are strongly expressed in developing xylem and, therefore, are likely to be involved in the conifer secondary cell wall biosynthesis process. Importantly, similar *GUX* expression profiles were obtained for both *Picea abies* and *Pinus radiata,* suggesting that the two GUX enzyme clades may contribute to xylan formation across many conifer taxa. This hypothesis is further supported by a recent report indicating that a *Pinus densiflora* gene homologous to *AtGUX1* is expressed more strongly in xylem than in other conifer tissues (Kim *et al.*, 2021).

Conifer GUX1 and GUX2 are xylan glucuronyltransferases with distinct specificities

We previously demonstrated that *Pg*GUX1 is an active xylan glucuronyltransferase by using transient expression in *N. benthamiana* for an *in vitro* activity assay (Lyczakowski *et al.*, 2017). To extend this analysis and study conifer GUX2 clade representatives, we expressed *myc*-tagged *Pt*GUX1 and *Pt*GUX2 enzymes and *Picea sitchensis* (*Ps*) GUX2 enzyme (Fig. S4). We selected *P. taeda* GUX1 and GUX2 as their protein sequences are identical to those in *P. radiata* (Fig. S5). Therefore, the GUXs enzymatic activity is likely to be identical in these two species. *Pg*GUX1 was used as a positive control. An Arabidopsis glucomannan galactosyltransferase *At*MAGT1 (Yu *et al.*, 2018) was used as a negative control to monitor for tobacco's potential endogenous glucuronyltransferase activity. In contrast with hardwood, softwood's xylan is not acetylated. However, xylan lacking acetylation and GlcA side branches is poorly soluble. Therefore, to maintain the acceptor solubility during the *in vitro* reaction, we have used acetylated xylan extracted from the Arabidopsis *gux1/2* mutant. Membrane preparations enriched with the *myc*-tagged proteins of interest were incubated with acetylated polymeric xylan lacking [Me]GlcA branches in the presence or absence of UDP-GlcA. Following incubation, the reaction products were extracted, deacetylated, digested with a GH11 xylanase and analysed with PACE (Fig. 3a). In the tobacco system, no xylan glucuronidation background activity was detected as the reaction products coming from the incubation in the presence of

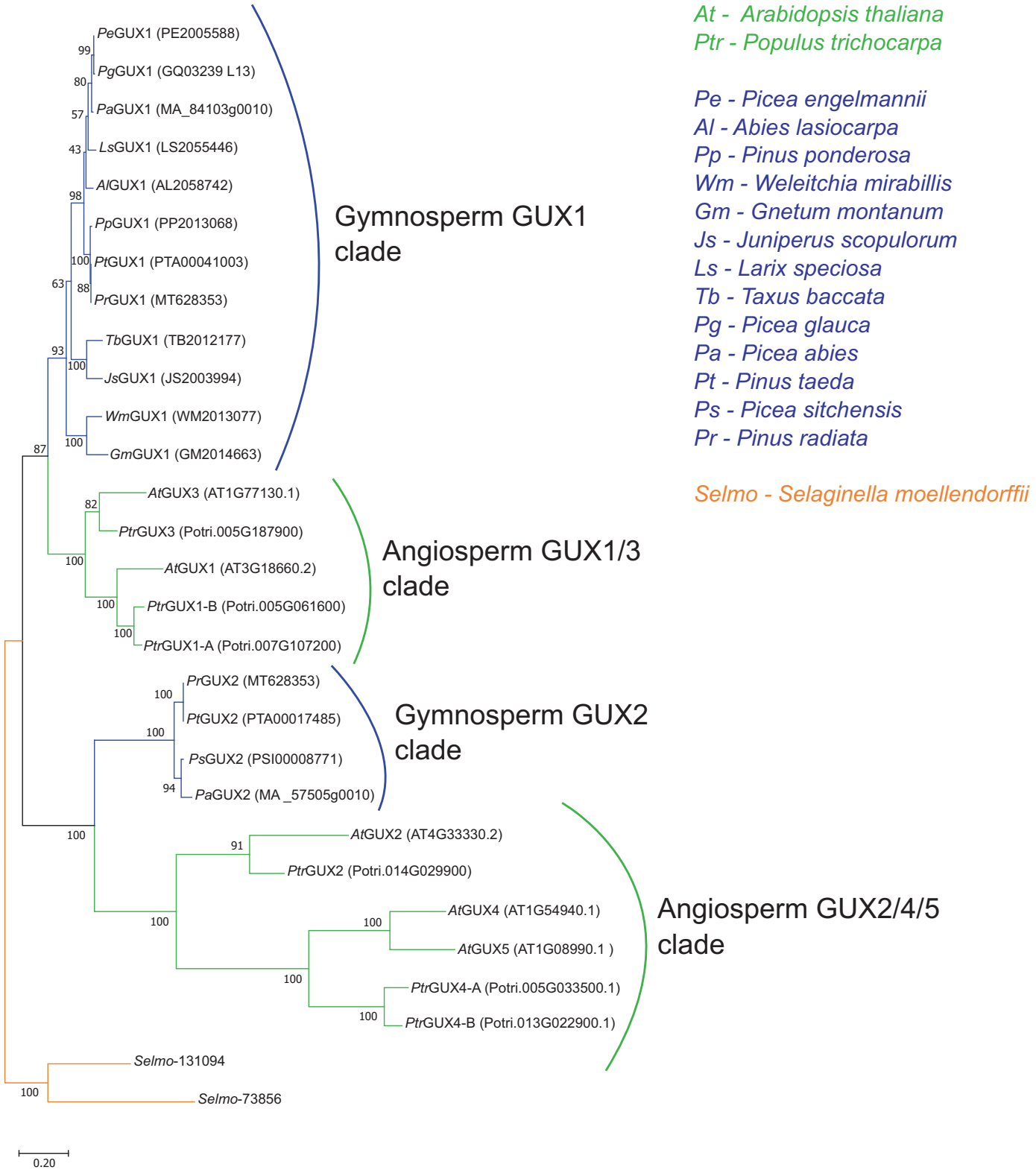

Fig. 1 Phylogenetic analysis of angiosperm and gymnosperm GlucUronic acid substitution of Xylan (GUX) enzymes. Analysis based on amino acid sequences of GT8 domains. Angiosperm and gymnosperm sequences are shown in green and blue, respectively. Specific clades are annotated on the phylogenetic tree. Putative GUXs of the lycophyte *Selaginella moellendorffii*, shown in orange, were used as an outgroup.

*At*MAGT1 were fully converted to xylose and xylobiose. A GlcA-xylotetraose (XUXX) product was observed in all reactions in which putative conifer GUX enzymes were incubated in the presence of the acceptor and UDP-GlcA. This demonstrates that conifer GUX1 and GUX2 enzymes were able to add glucuronic acid onto an acetylated xylan backbone *in vitro*.

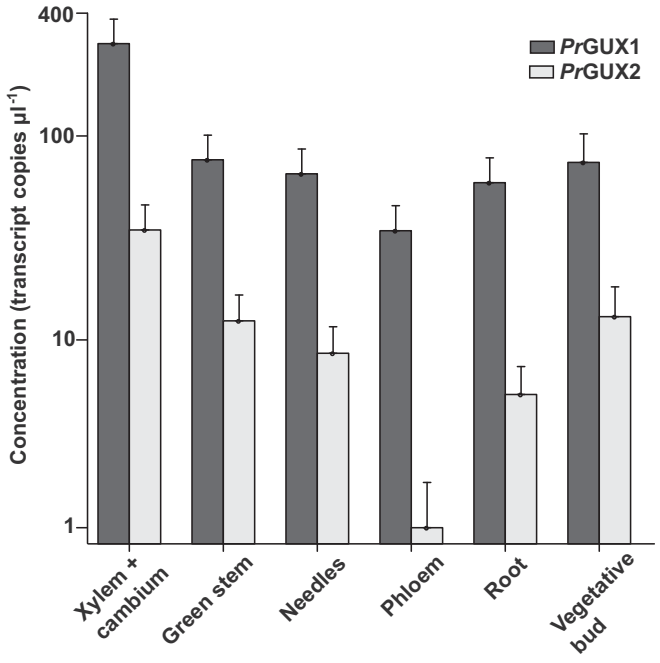

Fig. 2 Absolute quantification of *Pinus radiata* GUX1 and GUX2 expression using ddPCR. Concentrations are presented as copies per µl when an equal amount of cDNA was used for each of six tissue types. Bars represent estimated marginal means ± 95% confidence intervals.

We noticed that the reaction catalysed by both *Pt*GUX2 and *Ps*GUX2 enzymes produced an additional product migrating above XUXX. This product was not detected in GUX1 catalysed reactions. Therefore, the activity enabling its formation was specific to representatives of the conifer GUX2 clade. To characterise this novel structure, GH11 products were further digested with α-glucuronidase GH115 and the resulting oligosaccharides analysed by PACE (Fig. 3b). As expected, the removal of GlcA from the XUXX structure resulted in its co-migration with the xylotetraose (X_4) standard. The GH115 treatment of the uncharacterised oligosaccharide altered its migration profile to match the xylopentaose (X_5) standard, indicating that the structure has five xyloses on its backbone. GH11 needs two unsubstituted xyloses on the −1 and −2 sites and one unsubstituted xylose on the +1 site to digest xylan (Vardakou *et al.*, 2008; Paes *et al.*, 2012). Therefore, oligosaccharides produced by GH11 xylanases have one unsubstituted xylosyl residue at the nonreducing end, and two unsubstituted xylosyl residues at the reducing end. Consequently, a GH11 released oligosaccharide with five xylosyl monomers in the backbone must contain two branches on positions −3 and −4 from the GH11 cleavage site (Fig. 3c). Therefore, the structure of the unknown oligosaccharide must correspond to XUUXX. The identity of the oligosaccharide was supported by its PACE migration profile. The oligosaccharide migrates between X_4 and X_3 oligosaccharide standards, whereas a singly glucuronidated xylopentaose migrates between the X_5 and X_4 standards (Bromley *et al.*, 2013). The increase in the extent of PACE migration originates from the presence of two negatively charged GlcA molecules on the XUUXX backbone. This supports our conclusion that conifer GUX2 enzymes are likely to generate an oligosaccharide with GlcA present on consecutive xyloses. This structure was already reported in softwood (Shimizu *et al.*, 1978; Yamasaki *et al.*, 2011; Martínez-Abad *et al.*, 2017).

Xylan acetylation modifies and reduces conifer GUX activity

As softwood xylan is not acetylated, we wanted to investigate the effect of removing xylan acetyl groups on the activity of conifer GUX1 and GUX2 enzymes. To investigate the impact of xylan acetylation we performed an *in vitro* GlcAT reaction to which we added CE4, a xylan acetylesterase enzyme. The CE4 enzyme removes single 2 or 3 linked acetyl groups from xylose monomers, which do not carry GlcA decorations (Taylor *et al.*, 2006). The enzyme was added 3 h after the start of the *in vitro* glucuronidation reaction to enable some GlcA to be transferred to avoid acceptor precipitation. Following an overnight incubation of the *gux1/2* xylan acceptor with *Pg*GUX1 and *Pt*GUX2 enzymes in the presence and absence of CE4 the reaction products were extracted, digested with GH11 and analysed by PACE (Fig. 4a).

GH11 cleavage of all reactions catalysed by *Pg*GUX1 and *Pt*GUX2 in the presence of UDP-GlcA released xylose, xylobiose and XUXX. Independently of the esterase presence, reactions catalysed by *Pg*GUX1 produced an additional band migrating between xylopentaose and xylotetraose standards. However, the intensity of this novel band was increased in the reactions including the CE4 enzyme. Partial xylan deacetylation increased the amount of this oligosaccharide facilitating analysis of its structure. To investigate the identity of this band, oligosaccharides released by the xylanase GH11 were further digested with alpha-glucuronidase GH115 and analysed by PACE (Fig. S6a). Following the GH115 treatment xylotetraose and xylohexaose were detected. The xylotetraose must have originated from XUXX deglucuronidation. Therefore, the novel oligosaccharide had six xyloses in the backbone. Since the GH11 enzyme requires two unsubstituted xyloses at the −1 and −2 sites and one unsubstituted xylose at the +1 site (Vardakou *et al.*, 2008; Paes *et al.*, 2012), an oligosaccharide with six xylosyl monomers in the backbone is likely to contain two GlcA branches on positions −3 and −5 from the GH11 cleavage sites (Fig. S6b). The presence of this structure is consistent with GH11 specificity, as the enzyme cannot degrade oligosaccharides in which only a single unbranched xylose separates two branched backbone subunits. The identity of the oligosaccharide is also supported by its PACE migration profile between X_5 and X_4 standards, as singly glucuronidated xylohexaose would migrate between X_6 and X_5 oligosaccharides (Bromley *et al.*, 2013). Taking all evidence into consideration, the oligosaccharide is highly likely to be XUXUXX. The reaction catalysed by *Pt*GUX2 produced XUXX and XUUXX in similar frequency in the absence of acetylxylan esterase, while in the presence of the esterase XUUXX became the dominant glucuronidated product (Fig. 4b). Having established the likely identity of all glucuronidated structures observed in the PACE experiment (Fig. 4a) we were able to compare the efficiency of *in vitro* glucuronidation in the presence and absence of xylan

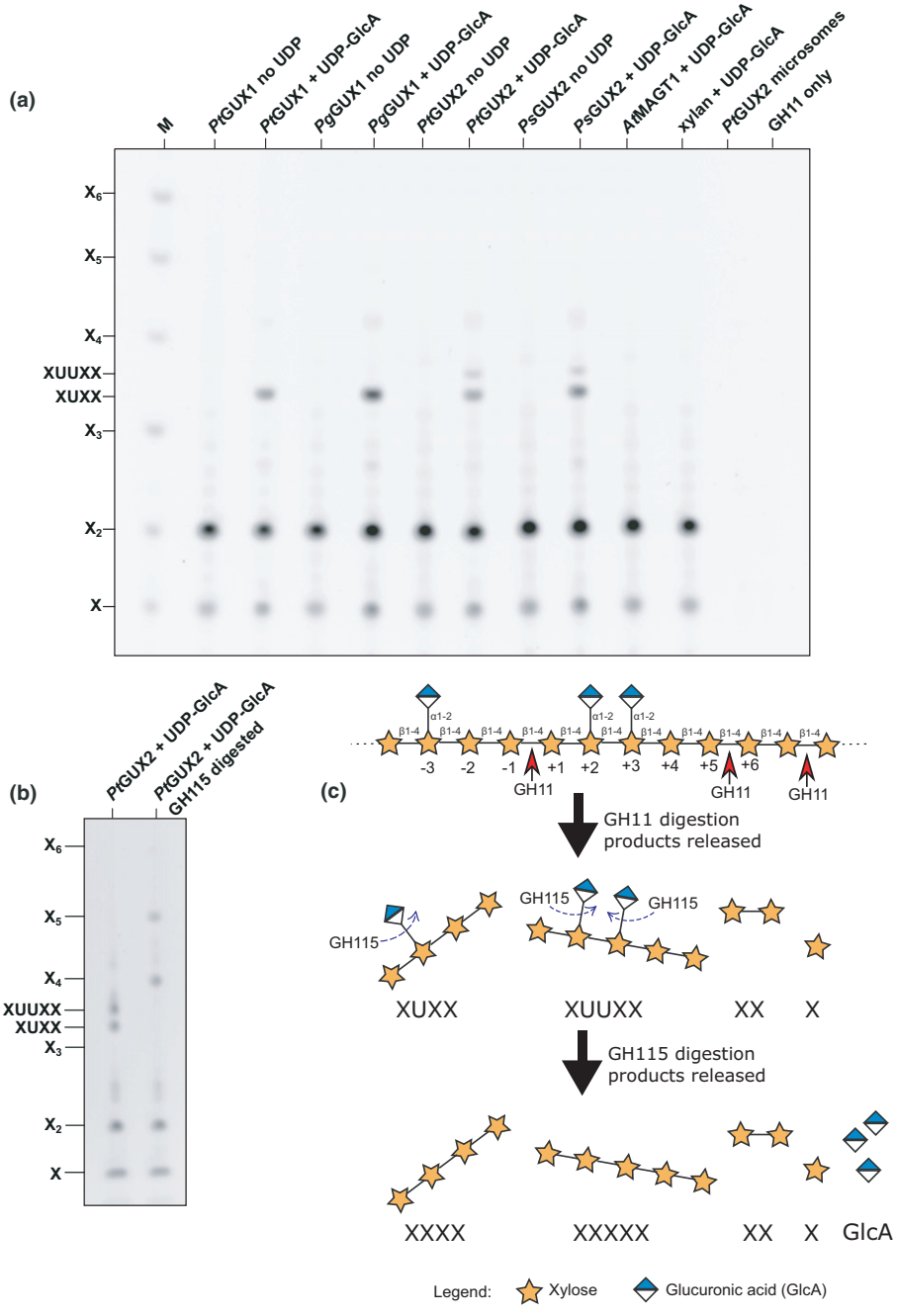

Fig. 3 *In vitro* activity of two conifer GlucUronic acid substitution of Xylan (GUX) enzymes. (a) Results of GH11 digestion of *in vitro* activity products of glucuronidation reaction catalysed by *Pinus taeda* (*Pt*) GUX1 and GUX2, *Picea glauca* (*Pg*) GUX1 and *Picea sitchensis* (*Ps*) GUX2. Glucuronidated oligosaccharides are marked. Control reaction, performed using membranes enriched for Arabidopsis MAGT1 did not release any glucuronidated oligosaccharides. (b) GH115 treatment of *Pt*GUX2 *in vitro* reaction products digested with GH11. Removal of glucuronic acid from the structure unique to clade 2 conifer GUX shifts it to the X5 position. (c) Diagram presenting a GH11 digestion profile expected from a polysaccharide containing some GlcA on consecutive xylosyl units. In (c), red arrow shows GH11 digestion point and blue arrows show GH115 de-glucuronidation. GlcA is not visible in PACE (b) since oligosaccharides are labelled with the ANTS fluorophore only once, after GH11 digestion. GH11 requires three contiguous unsubstituted xyloses on positions −2, −1 and +1 from the digestion point to cleave the xylan backbone. Therefore, the presence of two consecutive decorated xyloses on positions +2 and +3 from one digestion event will lead to a next hydrolysis event happening only between xyloses +5 and +6 from the original hydrolysis point, if these positions are unbranched. This results in formation of xylopentaose oligosaccharide with two glucuronidated xyloses: XUUXX.

acetylesterase (Fig. 4b). Our analysis indicates that the addition of the esterase increases the efficiency and alters the specific glucuronidation of *Pg*GUX1 and *Pt*GUX2. It is possible that using an acceptor with reduced acetylation allows the conifer GUX enzymes to exert *in vitro* glucuronyltransferase activities more similar to their *in planta* function.

Conifer GUX1 and GUX2 branching patterns are different and dependant on the degree of acetylation of the xylan acceptor

As the acetylation level of the acceptor influences the specific *in vitro* glucuronyltransferase activity of *Pg*GUX1 and *Pt*GUX2, we decided to evaluate its impact on the longer distance GlcA patterning. For this purpose, we analysed the GlcA decoration pattern of the xylan products generated during *in vitro* reactions by *Pg*GUX1 and *Pt*GUX2 in the presence or absence of the xylan acetylesterase. The products were deacetylated and digested with xylanase GH30, which enables analysis of GlcA patterning on xylan (Fig. 5a). The released oligosaccharides were derivatised with the 8-aminonaphthalene-1,3,6-trisulfonic acid, disodium salt (ANTS) fluorophore and analysed by PACE (Fig. 5b). In addition to the xylo-oligosaccharide standard, wild-type (WT) *A. thaliana* AIR was digested with GH30 to produce known glucuronidated xylo-oligosaccharide markers.

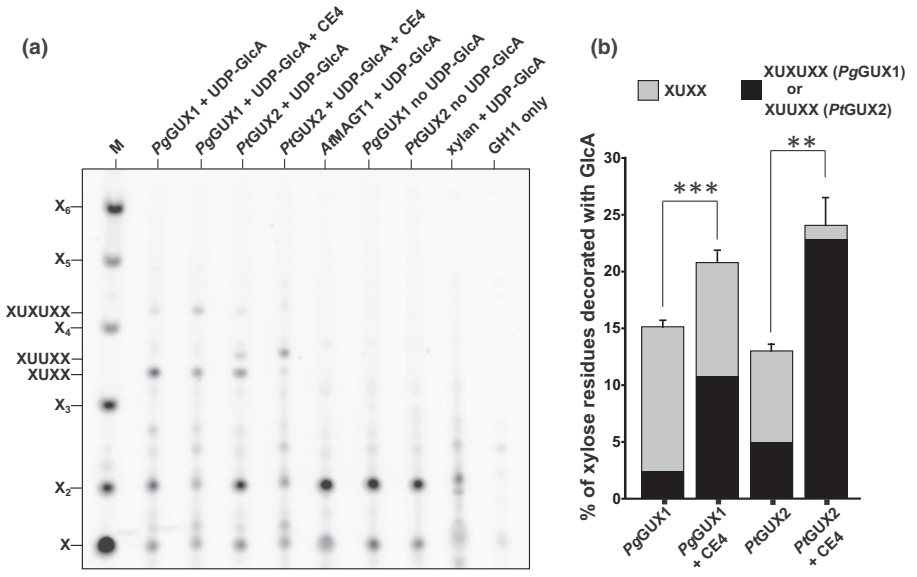

Fig. 4 Partial removal of xylan acetylation increases the activity of conifer GlucUronic acid substitution of Xylan (GUX) enzymes *in vitro*. (a) PACE of a GH11 digest of *Pg*GUX1 and *Pt*GUX2 *in vitro* glucuronidation products obtained with or without the CE4 xylan acetyl esterase (CE4) in the reaction mix. Arabidopsis MAGT1 is used as a negative control. (b) Quantification of the degree of glucuronidation obtained for the reaction catalysed by *Pg*GUX1 and *Pt*GUX2 in the absence or presence of the xylan acetyl esterase in the reaction mix. Three replicates of the reaction were performed and analysed. Error bars represent standard deviation, **, $P \leq 0.01$; ***, $P \leq 0.001$.

For *Pg*GUX1, in the absence of xylan acetylesterase, the most frequent spacing of GlcA was equal to 6 xylose units. Oligosaccharides with a longer degree of polymerisation (DP) were also released. The pattern changed substantially upon the addition of xylan acetylesterase to the reaction. Partial deacetylation of the acceptor resulted in *Pg*GUX1 generating a denser GlcA patterning *in vitro* with spacings of 6, 4 and 2 xylosyl units preferred. For *Pt*GUX2, in the absence of the xylan acetylesterase, a GH30 digestion of the glucuronoxylan released a mixture of single GlcA branched oligosaccharides with a second band migrating close to it. The GH11 digestion of the products of the same reaction resulted in production of singly and doubly glucuronidated oligosaccharides (Fig. 4a). Therefore, the GH30-produced oligosaccharides migrating close to known digestion products might be carrying two GlcA on consecutive xyloses. This structure is unlikely to be digested further by GH30 due to the steric hindrance of the two consecutive decorations. Upon the addition of xylan acetylesterase to a glucuronidation catalysed by *Pt*GUX2 the digestion pattern produced by GH30 was modified. The released oligosaccharides were shorter and almost all of them migrated like putative oligosaccharides bearing two GlcA branches. This alteration in the ratio of doubly and singly glucuronidated oligosaccharides, is in line with the GH11 digestion results (Fig. 4a).

Together with the GH11 data (Fig. 4a), the results of the GH30 analysis provide further evidence that the pattern of GlcA branches added by both *Pg*GUX1 and *Pt*GUX2 is sensitive to the degree of acetylation of the xylan acceptor. For both enzymes, partial removal of acetyl branches resulted in generation of more densely glucuronidated xylan molecules. This is indicated by both a significant increase in the degree of glucuronidation in the GH11 analysis, and the release of shorter oligosaccharides after the GH30 digestion. Moreover, the addition of the acetylesterase enabled *Pt*GUX2 to add GlcA to neighbouring xylose units with higher efficiency, suggesting this might be a preferred mode of glucuronidation for this enzyme.

Conifer GUX enzymes are able to glucuronidate xylan *in planta*

We wanted to observe the *in planta* glucuronyltransferase activity of conifer GUX2 enzymes. For this purpose, we expressed *PtGUX2* under the control of *pIRX3,* a strong secondary cell wall-specific Arabidopsis promoter, in a *gux1/2/3* Arabidopsis background. The *gux1/2/3* plants lack any detectable xylan glucuronidation (Mortimer *et al.*, 2015). Therefore, any GlcA detected on the xylan of *gux1/2/3* plants would be introduced by *PtGUX2*. Two individual lines were obtained (*Pt2gux*.1 and *Pt2gux*.2) and their xylan structure was analysed by PACE after xylanase GH11 digestion (Fig. 6a). The *Pt*GUX2 expression led to restoration of GlcA side chains on *gux1/2/3* xylan. The total degree of xylan glucuronidation in the transgenic plants exceeded the one observed in WT Arabidopsis (Fig. 6b). Moreover, for both transgenic lines the XUXX and the XUUXX oligosaccharides were detected after GH11 digestion of xylan. This further confirmed that the clade 2 conifer GUX enzymes are capable of glucuronidating consecutive xylosyl units.

We then investigated the effect of xylan acetylation on the *in planta* activity of conifer GUX enzymes. To this end, we generated *A. thaliana esk-kak* mutant plants expressing *PgGUX1* or *PtGUX2* (lines *Pg1 esk-kak* and *Pt2 esk-kak,* respectively) under the control of *IRX3* promoter. The *esk-kak* plants have an approximate 50% reduction in xylan acetylation levels (Xiong *et al.*, 2013; Grantham *et al.*, 2017). Therefore, in this background, the activity of conifer GUX enzymes may be higher than within the *gux1/2/3* plants where the xylan is fully acetylated. However, in the Arabidopsis *esk-kak* mutant background the native GUX enzymes contribute to glucuronidation. AIR isolated from basal stem of the *Pg1 esk-kak* and *Pt2 esk-kak* plants was digested with xylanase GH11 and the resulting oligosaccharides were separated by PACE (Fig. 6c). The detected glucuronidated structures were primarily XUXUXX and XUUXX in *esk-kak* plants overexpressing *Pg*GUX1 and *Pt*GUX2, respectively. Interestingly, when

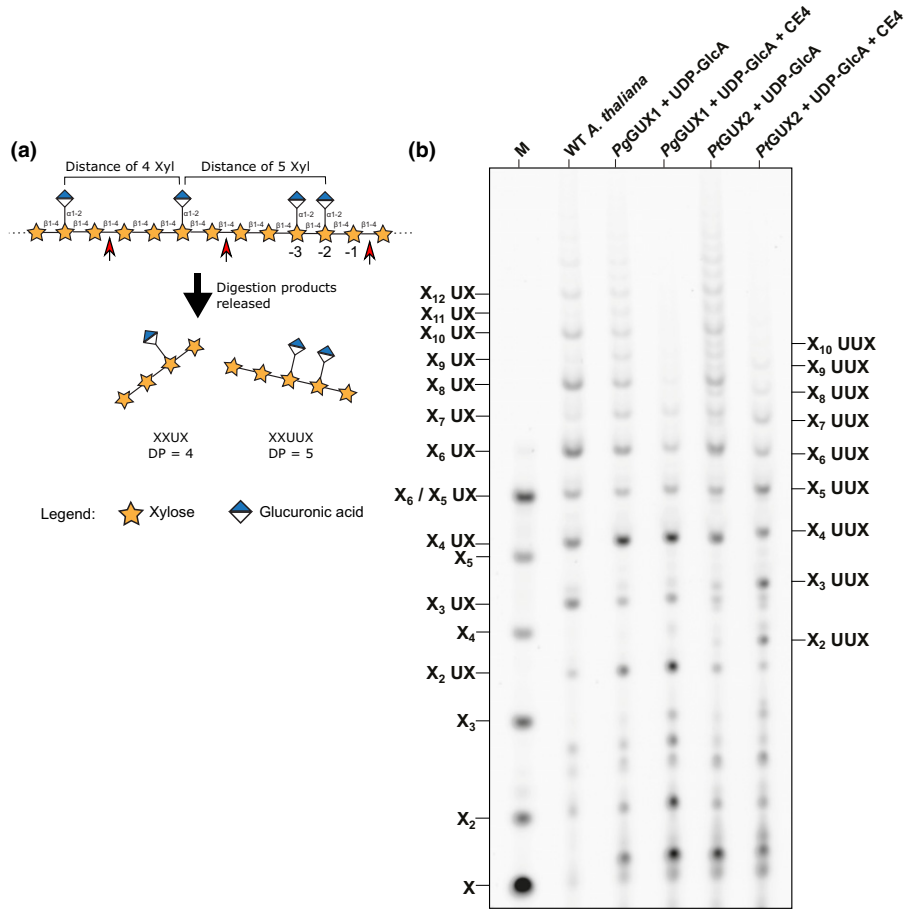

Fig. 5 Supplementation with a xylan acetyl esterase changes the pattern of GlcA branches introduced by conifer GlucUronic acid substitution of Xylan (GUX) enzymes *in vitro*. (a) Diagram demonstrating GH30 activity on xylan with individual or consecutive GlcA branches. Red arrows indicate digestion points by GH30. DP is degree of polymerisation. GH30 enzyme requires presence of glucuronidated xylose on position −2 to catalyse the hydrolysis of the xylan backbone. The presence of glucuronidated xylose also on position −3 appears not to inhibit GH30 digestion. (b) PACE of a GH30 digest of *Pg*GUX1 and *Pt*GUX2 *in vitro* activity products obtained in reactions performed with or without the xylan acetyl esterase. Wild-type *Arabidopsis thaliana* biomass was digested to obtain a standard of singly glucuronidated oligosaccharides (marked on the left edge of the gel, UX on the reducing end). Predicted annotation of putative doubly glucuronidated oligosaccharides (marked on the right edge of the gel, UUX on the reducing end) is proposed based on the identity of the most proximous singly glucuronidated oligosaccharide. Change in band migration is likely to originate from increase in charge (positive effect on band migration) and mass (negative effect on band migration) of the oligosaccharide due to presence of second GlcA branch. *Pt*GUX2 reaction performed in the absence of the esterase enzyme releases a mix of singly and doubly glucuronidated structures.

overexpressed in the *esk-kak* background, the *Pt*GUX2 enzyme also generated structures resulting in XUXUXX formation by GH11. In *esk-kak* plants overexpressing the *Pg*GUX1 and *Pt*GUX2 enzymes the total degree of glucuronidation exceeded 20% of the backbone (Fig. 6d). This is more than observed previously for *Pg*GUX1 (Lyczakowski *et al.*, 2017) or *Pt*GUX2 expressed in *gux1/2/3* plants. This observation strengthens the hypothesis that a low degree of xylan acetylation is favourable to conifer GUX activity.

Impact of *Pt*GUX2 expression on saccharification of *gux1/2/3 A. thaliana* biomass

The absence of [Me]GlcA branches on xylan of *gux1/2/3* Arabidopsis is associated with a significant increase in glucose and xylose release from biomass during enzymatic saccharification (Lyczakowski *et al.*, 2017). Therefore, to investigate if the GlcA branches introduced by the *Pt*GUX2 enzyme are capable of rescuing the mutant phenotype we performed a saccharification assay on cell wall material isolated from WT, *gux1/2/3* and *gux1/2/3* plants expressing *Pt*GUX2. In the *gux1/2/3* plants expressing *Pt*GUX2 the release of both glucose (Fig. 7a) and xylose (Fig. 7b) was reduced to that measured for WT Arabidopsis. This suggested that a softwood-like pattern of GlcA branches, introduced by the *Pt*GUX2 enzyme, can participate in the maintenance of biomass recalcitrance.

Discussion

The progression towards a more sustainable economy requires an optimised use of softwood, which represent a significant proportion of the lignocellulosic biomass processed in biorefineries.

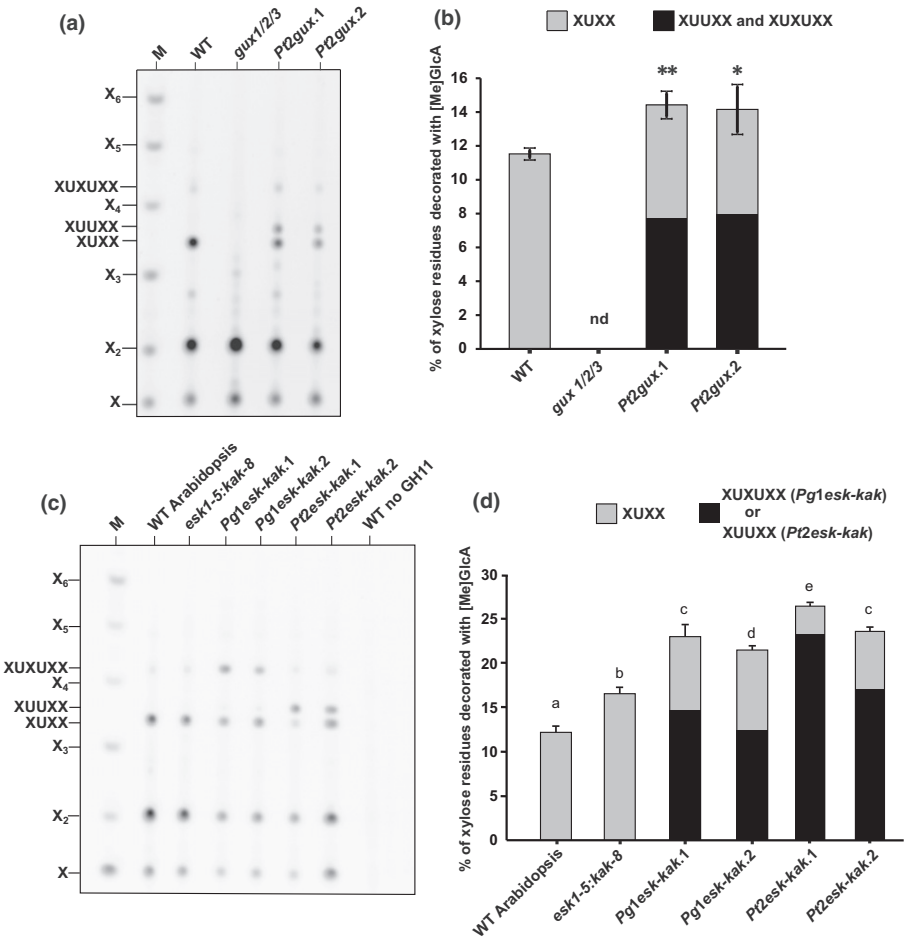

Fig. 6 Analysis of transgenic Arabidopsis plants overexpressing the *Pt*GUX2 enzyme under the control of a secondary cell wall-specific *IRX3* promoter. (a) PACE of a GH11 digest of Arabidopsis wild-type (WT), *gux1/2/3* and two independent *gux1/2/3* lines overexpressing the *Pt*GUX2 enzyme. X1 to X6 marker (M) is used. (b) Quantification of the degree of glucuronidation in WT and transgenic lines. In line with previous reports (Mortimer *et al*., 2015; Lyczakowski *et al*., 2017) *gux1/2/3* plants have no detectable (nd) GlcA on xylan. Student's *t*-test was used to evaluate the significance of the difference between the degree of glucuronidation in WT and *Pt*GUX2 expressing *gux1/2/3* plants, *, $P \leq 0.05$; **, $P \leq 0.01$. Error bars represent standard deviation. (c) PACE of xylan structure in WT, *esk-kak* and *PgGUX1* (lines *Pg*1*esk-kak*) and *PtGUX2* (lines *Pt*2*esk-kak*) expressing *esk-kak* using xylanase GH11. (d) Quantification of the degree of glucuronidation obtained for WT, *esk-kak* and *PgGUX1* and *PtGUX2* expressing *esk-kak*. Three biological replicates of biomass were analysed. Grey coloration of the bars indicates the GlcA contributed by the XUXX structure. Black bars denote GlcA contributed by the XUUXX and XUXUXX structures. Error bars represent standard deviation. Tukey test after ANOVA analysis was used to evaluate the significance of the difference between the degree of glucuronidation across the plant genotypes studies. Letter codes are provided to indicate plant lines with significantly different degrees of xylan glucuronidation.

Despite its importance very little information is known about the molecular basis of softwood cell wall biosynthesis. In this work we investigated the gene expression profiles and biochemical activities of conifer xylan glucuronyltransferases. We found two distinct conifer clades of GUX enzymes expressed during secondary cell wall synthesis. Each clade produced a specific glucuronic acid substitution pattern. The conifer *GUX* genes could be important breeding or engineering targets for modification of conifer biomass, particularly to improve softwood digestibility.

In our previous work, we discovered a single clade of conifer GUX enzymes (Lyczakowski *et al*., 2017). Here, after re-analysing transcriptomic and genomic data of conifer species, we discovered a second clade of conifer GUX enzymes. Our phylogenetic analysis (Fig. 1) indicated that while the previously characterised conifer GUX clade shares a high sequence similarity with hardwood GUX1/3 sequences, the newly identified sequences were more comparable to the hardwood GUX2/4/5 clade. This suggests a potential functional divergence of conifer GUX genes, similar to that seen in dicot species (Bromley *et al*., 2013). Interestingly, conifer GUX sequences share an exceptionally high degree of sequence similarity within each clade (Fig. S5). This is consistent with the theory that most of gymnosperm evolution happened before the lineage split with angiosperms and that after this event the pace of genome evolution has been slower in the gymnosperm lineage than in the angiosperm one (Buschiazzo *et al*., 2012; Pavy *et al*., 2012). Therefore, it is likely that the functional characterisation performed in this work may be an accurate representation of GUX activities for the entire taxon.

We were able to demonstrate, using both *in vitro* assays and *in vivo* complementation, that some conifer genomes encode a

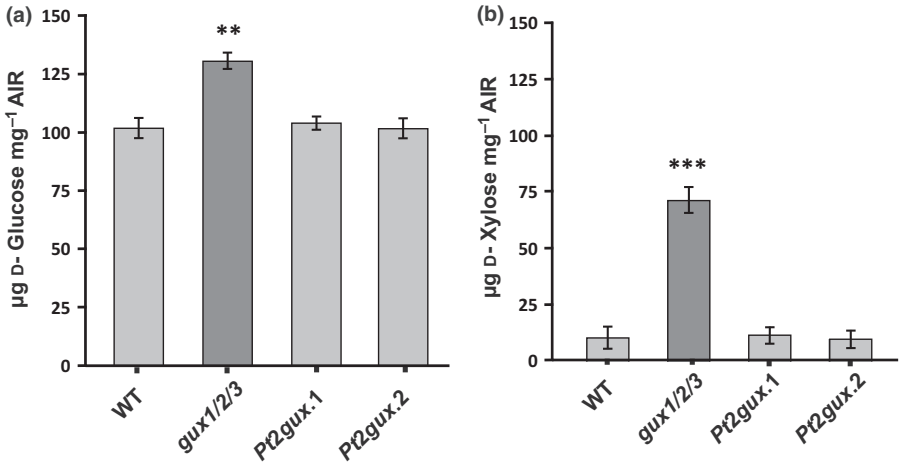

Fig. 7 Effect of *PtGUX2* expression on saccharification of *gux1/2/3 Arabidopsis thaliana* biomass. Saccharification of AIR from Arabidopsis WT, *gux1/2/3* and *Pt2gux* lines. Glucose (a) and xylose (b) release from AIR. Dark grey bars denote datasets that are significantly different from the WT. For all graphs error bars represent standard deviation. Student's *t*-test was used to evaluate the significance of the difference between WT and transgenic lines. **, $P \leq 0.01$; ***, $P \leq 0.001$.

minimum of two distinct active xylan glucuronyltransferases. Importantly, our results indicate that the two conifer GUX enzymes may be responsible for the synthesis of different patterns of GlcA branches. This conclusion is supported by the results of our *in vitro* GUX activity experiments. In particular, by decreasing the acetylation level of the *in vitro* xylan acceptor it was possible to shift conifer GUX activities towards generating softwood specific types of patterning. As the conifer AGX is not acetylated (Scheller & Ulvskov, 2010) it is possible that these low acetylation conditions are more similar to the *in planta* environment. When acting on the xylan acceptor with enzyme-reduced acetylation, clade 1 enzymes were observed to add GlcA predominantly onto every 6th, 4th and 2nd xylosyl unit. In the same conditions, clade 2 enzymes added GlcA mainly onto consecutive xylosyl units. The two adjacent xylose backbone units, each bearing a GlcA branch, were mainly interspaced by 4, 5 or 6 unsubstituted xylosyl units. Softwood xylan is also decorated with arabinose residues. However, these decorations were absent in our *in vitro* and *in planta* acceptors. This difference is probably influencing the specific activity of conifer GUX. Therefore, the relationship between the patterning of glucuronic acid and arabinose substitutions in conifers remains to be investigated.

Our results suggested that the two distinct clades of GUX enzymes in conifers are responsible for the two patterns of MeGlcA branches observed in softwood. The conifer clade 1 enzymes could be responsible for the synthesis of the predominant xylan pattern in softwood (Busse-Wicher *et al.*, 2016b), with a spacing of six xylose units between the GlcA branches. Clade 2 of conifer GUX enzymes could be responsible for addition of GlcA branches onto two consecutive xylose residues, forming the structure described by Martínez-Abad *et al.* (2017). To confirm this hypothesis, we are currently analysing transgenic conifer plants lacking GUX1 or GUX2 activity.

Patterning of GlcA branches is thought to be a determinant of the compatibility of xylan molecules for their interaction with the hydrophilic face of the cellulose fibril (Grantham *et al.*, 2017). In *A. thaliana*, the *At*GUX1 enzyme is responsible for the addition of evenly spaced GlcA branches, whereas the *At*GUX2 enzyme adds GlcA branches without this preference (Bromley *et al.*, 2013). Therefore, with xylan in a two-fold screw conformation, *At*GUX1 enzymes produce a pattern with all the GlcA projecting from the same side of the backbone. The other side of xylan, free of decorations, is therefore compatible with the cellulose fibril and can adsorb onto it (Busse-Wicher *et al.*, 2016a). *At*GUX2 places GlcA decorations projecting on both sides of a potential two-fold screw xylan. This pattern, known as incompatible, is likely to alter the interaction between xylan and the hydrophilic surface of the cellulose microfibril.

Similarly, to the *A. thaliana* model, both compatible and incompatible GlcA patterns may be present in softwood. The majority of softwood AGX has a MeGlcA branch on every 6th monomer (Busse-Wicher *et al.*, 2016b) and as such can be considered compatible with adsorption onto the hydrophilic surface of the cellulose microfibril. A minor proportion of softwood AGX has GlcA branches on two consecutive monomers (Martínez-Abad *et al.*, 2017). According to the current xylan–cellulose interaction model this structure could have a reduced capacity to adsorb onto the hydrophilic face of the cellulose fibril (Busse-Wicher *et al.*, 2016a; Grantham *et al.*, 2017). However, molecular dynamics simulations indicate that xylan with GlcA on consecutive xylosyl monomers may also interact with certain parts of the hydrophilic face of the cellulose microfibril (Martínez-Abad *et al.*, 2017). Convergently, recently published ssNMR experiments on spruce indicated that both two-fold, cellulose bound, and three-fold, soluble, xylan structures are present in softwood (Terrett *et al.*, 2019). The cellulose face(s) upon which the xylan binds remains to be discovered but it is likely that at least some of the softwood AGX has a structure that is incompatible with binding the hydrophilic surface of the cellulose microfibril. Noticeably, the extractability of spruce xylan with consecutive MeGlcA branches is lower than that of xylan with evenly spaced MeGlcA branches (Martínez-Abad *et al.*, 2020). This extractability difference might originate from xylan's interaction capacity with lignin, which can be regulated partially by its glucuronidation (Martínez-Abad *et al.*, 2020). When investigating softwood properties it is also is important to account for the role of the predominant hemicellulose, the GGM. Similar to AGX, GGM can bind to cellulose (Terrett *et al.*, 2019) and

lignin (Nishimura *et al.*, 2018). Like AGX, GGM also shows diverse extractability patterns (Martínez-Abad *et al.*, 2020). It is therefore likely that the potential interactions between AGX and GGM may also contribute to conifer cell wall properties. This intrahemicellulose interaction might also be influenced by AGX and GGM decoration patterns.

Xylan glucuronidation, in addition to determining the capacity of the hemicellulose to interact with the hydrophilic surface of the cellulose microfibril, is critical for the maintenance of plant biomass recalcitrance to enzymatic digestion (Lyczakowski *et al.*, 2017). Importantly, our previous work suggested that reducing the amount of [Me]GlcA branching by 30% had no effect on cell wall recalcitrance (Lyczakowski *et al.*, 2017). In *gux1/2/3* plants overexpressing the *Pt*GUX2 enzyme, we observed a 20%, increase in [Me]GlcA levels compared with WT. However, this did not result in any further increase in biomass recalcitrance. These observations are consistent with the proposed role of GlcA in the formation of ester linkages with lignin (Watanabe & Koshijima, 1988). In this model, a small number of linkages might maintain full xylan–lignin cross-linking (Giummarella & Lawoko, 2016). Therefore, a partial change in the content of [Me]GlcA branches does not result in recalcitrance reduction or further resistance gain to enzymatic digestion. This may be important for conifer GUX mutagenesis experiments aiming to generate softwood material with reduced xylan glucuronidation and possibly with reduced wood recalcitrance. Since the absence of xylan glucuronidation is required to decrease biomass recalcitrance, homologues of both conifer GUX may need to be simultaneously targeted for mutagenesis to achieve a significant reduction in softwood recalcitrance. However, even small incremental improvements in saccharification or pulping efficiency provided by a single GUX mutant could be relevant in the context of industrial processes. For example, the reduction of GlcA during the kraft pulping process would reduce the formation of undesired chemical by-products such as hexenuronic acid, which leads to the creation of hazardous organic halogens (Danielsson *et al.*, 2006).

In summary, our work identifies and describes the function of two clades of conifer GUX enzymes. We were able to demonstrate that clade 1 enzymes are likely to generate compatible pattern of GlcA branches and that clade 2 enzymes have the unique capacity of adding GlcA branches onto two consecutive xylose monomers. Since xylan glucuronidation is important for the maintenance of biomass recalcitrance, our work may guide mutagenesis or breeding approaches aiming to generate softwood biomass that can be more easily converted into pulp or biofuels. Moreover, as xylan–cellulose interaction may influence the mechanical properties of secondary cell walls (Simmons *et al.*, 2016; Grantham *et al.*, 2017) modulating the relative contribution clade 1 and clade 2 GUX activity to the process of xylan glucuronidation in softwood may prove a valid strategy to modify wood strength. As such, our work may be applied for the modification of softwood material to make it more suitable for applications in sustainable industries such as pulping, biofuels or wood-based construction.

Acknowledgements

We would like to acknowledge Prof. George Lomonossoff (John Innes Centre, UK), who developed the pEAQ-HyperTrans expression system used in this study. Plant Bioscience Limited supplied the pEAQ-HT vector that was used in this work. Enzymes used in this study (GH115 and *Ec*GH30) were a kind gift from Kristian Krogh (Novozymes A/S, Denmark). Dr Krogh also acted as a supporting supervisor during OMT doctoral training at the University of Cambridge. This work was supported by the Leverhulme Trust Centre for Natural Material Innovation and by Biotechnology and Biological Sciences Research Council (BBSRC) of the UK as part of the OpenPlant Synthetic Biology Research Centre (Reference BB/L014130/1), Cambridge BBSRC-DTP Programme (Reference BB/J014540/1) and iCASE studentship (Reference BB/M015432/1). Work was also supported by a grant from the National Science Centre Poland awarded to JJL as part of the SONATINA 3 programme (project number 2019/32/C/NZ3/00392). We acknowledge the funding received from the New Zealand Ministry of Business, Innovation and Employment (MBIE) Endeavour Fund (Contract no. C04X1707, Fibre Grand Design) for supporting this work. The authors declare no conflict of interest.

Author contributions

JJL designed the study, performed most polysaccharide analyses and genetic engineering, and co-wrote the paper; LY contributed to biochemical analyses, assignment of new polysaccharide structures and manuscript writing; OMT supported the biochemical analysis, isolation of stable transgenic Arabidopsis lines and co-wrote the paper; CF isolated RNA from radiata pine and performed RT-PCR experiments profiling GUX expression; HT contributed to assignment of new polysaccharide structures and isolation of stable transgenic Arabidopsis lines; GT was involved in the design of RT-PCR experiments and manuscript preparation; MS designed parts of the study and co-wrote the manuscript; PD designed the study, interpreted data and co-wrote the manuscript.

ORCID

Paul Dupree https://orcid.org/0000-0001-9270-6286
Jan J. Lyczakowski https://orcid.org/0000-0002-7694-8629
Mathias Sorieul https://orcid.org/0000-0001-7326-3707
Henry Temple https://orcid.org/0000-0002-0415-6403
Oliver M. Terrett https://orcid.org/0000-0002-3796-2858
Glenn Thorlby https://orcid.org/0000-0002-6742-900X
Li Yu https://orcid.org/0000-0001-8820-6705

Data availability

The data that support the findings of this study are available in the main text and in the Supporting Information for this article. Sequences for newly identified *P. radiata GUX* were deposited in GenBank and access codes are provided in the Supporting

Information. All plant material described in this manuscript is available on a reasonable request from the corresponding authors.

References

Addison B, Stengel D, Bharadwaj VS, Happs RM, Doeppke C, Wang T, Bomble YJ, Holland GP, Harman-Ware AE. 2020. Selective one-dimensional ^{13}C-^{13}C spin-diffusion solid-state nuclear magnetic resonance methods to probe spatial arrangements in biopolymers including plant cell walls, peptides, and spider silk. *Journal of Physical Chemistry B* **124**: 9870–9883.

Bar-On YM, Phillips R, Milo R. 2018. The biomass distribution on Earth. *Proceedings of the National Academy of Sciences, USA* **115**: 6506–6511.

Berglund J, Kishani S, de Carvalho DM, Lawoko M, Wohlert J, Henriksson G, Lindstrom ME, Wagberg L, Vilaplana F. 2020a. Acetylation and sugar composition influence the (in)solubility of plant beta-mannans and their interaction with cellulose surfaces. *ACS Sustainable Chemistry & Engineering* **8**: 10027–10040.

Berglund J, Mikkelsen D, Flanagan BM, Dhital S, Gaunitz S, Henriksson G, Lindström ME, Yakubov GE, Gidley MJ, Vilaplana F. 2020b. Wood hemicelluloses exert distinct biomechanical contributions to cellulose fibrillar networks. *Nature Communications* **11**: 4692.

Bligh EG, Dyer WJ. 1959. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* **37**: 911–917.

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

Buschiazzo E, Ritland C, Bohlmann J, Ritland K. 2012. Slow but not low: genomic comparisons reveal slower evolutionary rate and higher dN/dS in conifers compared to angiosperms. *BMC Evolutionary Biology* **12**: 8.

Busse-Wicher M, Gomes TCF, Tryfona T, Nikolovski N, Stott K, Grantham NJ, Bolam DN, Skaf MS, Dupree P. 2014. The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of *Arabidopsis thaliana*. *The Plant Journal* **79**: 492–506.

Busse-Wicher M, Grantham NJ, Lyczakowski JJ, Nikolovski N, Dupree P. 2016a. Xylan decoration patterns and the plant secondary cell wall molecular architecture. *Biochemical Society Transactions* **44**: 74–78.

Busse-Wicher M, Li A, Silveira RL, Pereira CS, Tryfona T, Gomes TCF, Skaf MS, Dupree P. 2016b. Evolution of xylan substitution patterns in gymnosperms and angiosperms: implications for xylan interaction with cellulose. *Plant Physiology* **171**: 2418–2431.

Clough SJ, Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal* **16**: 735–743.

Danielsson S, Kisara K, Lindstrom ME. 2006. Kinetic study of hexenuronic and methylglucuronic acid reactions in pulp and in dissolved xylan during kraft pulping of hardwood. *Industrial & Engineering Chemistry Research* **45**: 2174–2178.

Fernandes AN, Thomas LH, Altaner CM, Callow P, Forsyth VT, Apperley DC, Kennedy CJ, Jarvis MC. 2011. Nanostructure of cellulose microfibrils in spruce wood. *Proceedings of the National Academy of Sciences, USA* **108**: E1195–E1203.

Giummarella N, Lawoko M. 2016. Structural basis for the formation and regulation of lignin-xylan bonds in birch. *ACS Sustainable Chemistry & Engineering* **4**: 5319–5326.

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

Jarvis MC. 2018. Structure of native cellulose microfibrils, the starting point for nanocellulose manufacture. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences* **376**: 20170045.

Jokipii-Lukkari S, Sundell D, Nilsson O, Hvidsten TR, Street NR, Tuominen H. 2017. NorWood: a gene expression resource for evo-devo studies of conifer wood development. *New Phytologist* **216**: 482–494.

Kang X, Kirui A, Widanage MCD, Mentink-Vigier F, Cosgrove DJ, Wang T. 2019. Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. *Nature Communications* **10**: 347.

Kim M-H, Tran TNA, Cho J-S, Park E-J, Lee H, Kim D-G, Hwang S, Ko J-H. 2021. Wood transcriptome analysis of *Pinus densiflora* identifies genes critical for secondary cell wall formation and NAC transcription factors involved in tracheid formation. *Tree Physiology*. doi: 10.1093/treephys/tpab001.

Kumar V, Hainaut M, Delhomme N, Mannapperuma C, Immerzeel P, Street NR, Henrissat B, Mellerowicz EJ. 2019. Poplar carbohydrate-active enzymes: whole-genome annotation and functional analyses based on RNA expression data. *The Plant Journal* **99**: 589–609.

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

Martinez-Abad A, Berglund J, Toriz G, Gatenholm P, Henriksson G, Lindstrom M, Wohlert J, Vilaplana F. 2017. Regular motifs in xylan modulate molecular flexibility and interactions with cellulose surfaces. *Plant Physiology* **175**: 1579–1592.

Martínez-Abad A, Jiménez-Quero A, Wohlert J, Vilaplana F. 2020. Influence of the molecular motifs of mannan and xylan populations on their recalcitrance and organization in spruce softwoods. *Green Chemistry* **22**: 3956–3970.

Matasci N, Hung LH, Yan ZX, Carpenter EJ, Wickett NJ, Mirarab S, Nguyen N, Warnow T, Ayyampalayam S, Barker M *et al.* 2014. Data access for the 1,000 Plants (1KP) project. *Gigascience* **3**: 17.

Mortimer JC. 2017. Structural analysis of cell wall polysaccharides using PACE. *Xylem: Methods and Protocols*, **1544**, 223–231.

Mortimer JC, Faria-Blanc N, Yu X, Tryfona T, Sorieul M, Ng YZ, Zhang Z, Stott K, Anders N, Dupree P. 2015. An unusual xylan in Arabidopsis primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14. *The Plant Journal* **83**: 413–527.

Mortimer JC, Miles GP, Brown DM, Zhang Z, Segura MP, Weimar T, Yu X, Seffen KA, Stephens E, Turner SR *et al.* 2010. Absence of branches from xylan in Arabidopsis *gux* mutants reveals potential for simplification of lignocellulosic biomass. *Proceedings of the National Academy of Sciences, USA* **107**: 17409–17414.

Nishimura H, Kamiya A, Nagata T, Katahira M, Watanabe T. 2018. Direct evidence for alpha ether linkage between lignin and carbohydrates in wood cell walls. *Scientific Reports* **8**: 6538.

Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin YC, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A *et al.* 2013. The Norway spruce genome sequence and conifer genome evolution. *Nature* **497**: 579–584.

Oinonen P, Zhang LM, Lawoko M, Henriksson G. 2015. On the formation of lignin polysaccharide networks in Norway spruce. *Phytochemistry* **111**: 177–184.

Paes G, Berrin J-G, Beaugrand J. 2012. GH11 xylanases: Structure/function/properties relationships and applications. *Biotechnology Advances* **30**: 564–592.

Pan YD, Birdsey RA, Fang JY, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG *et al.* 2011. A large and persistent carbon sink in the world's forests. *Science* **333**: 988–993.

Patron N, Orzaez D, Marillonnet S, Warzecha H, Matthewman C, Youles M, Raitskin O, Leveau A, Farre G, Rogers C *et al.* 2015. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytologist* **208**: 13–19.

Pavy N, Pelgas B, Laroche J, Rigault P, Isabel N, Bousquet J. 2012. A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. *Bmc Biology* **10**: 84.

Proost S, van Bel M, Vaneechoutte D, van de Peer Y, Inze D, Mueller-Roeber B, Vandepoele K. 2015. PLAZA 3.0: an access point for plant comparative genomics. *Nucleic Acids Research* **43**: D974–D981.

Ramage MH, Burridge H, Busse-Wicher M, Fereday G, Reynolds T, Shah DU, Wu GL, Yu L, Fleming P, Densley-Tingley D *et al.* 2017. The wood from the trees: The use of timber in construction. *Renewable & Sustainable Energy Reviews* **68**: 333–359.

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

Sainsbury F, Thuenemann EC, Lomonossoff GP. 2009. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnology Journal* **7**: 682–693.

Scheller HV, Ulvskov P. 2010. Hemicelluloses. *Annual Review of Plant Biology* **61**: 263–289.

Sedjo R. 2001. *Biotechnology's potential contribution to global wood supply and forest conservation*. Discussion Paper 01-51. Washington, DC, USA: Resources for the Future.

Shimizu K, Hashi M, Sakurai K. 1978. Isolation from softwood xylan of oligosaccharides containint two 4-O-methyl glucuronic acid residues. *Carbohydrate Research* **62**: 117–126.

Simmons TJ, Mortimer JC, Bernardinelli OD, Poppler AC, Brown SP, Deazevedo ER, Dupree R, Dupree P. 2016. Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nature Communications* **7**: 13902.

Sparkes IA, Runions J, Kearns A, Hawes C. 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature Protocols* **1**: 2019–2025.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725–2729.

Taylor EJ, Gloster TM, Turkenburg JP, Vincent F, Brzozowski AM, Dupont C, Shareck F, Centeno MSJ, Prates JAM, Puchart V *et al.* 2006. Structure and activity of two metal ion-dependent acetylxylan esterases involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases. *Journal of Biological Chemistry* **281**: 10968–10975.

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

Terrett O, Dupree P. 2019. Covalent interactions between lignin and hemicelluloses in plant secondary cell walls. *Current Opinion in Biotechnology* **56**: 97–104.

Terrett OM, Lyczakowski JJ, Yu L, Iuga D, Franks WT, Brown SP, Dupree R, Dupree P. 2019. Molecular architecture of softwood revealed by solid-state NMR. *Nature Communications* **10**: 4978.

Timell TE, Syracuse NY. 1967. Recent progress in the chemistry of wood hemicelluloses. *Wood Science and Technology* **1**: 45–70.

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. 2010. Lignin biosynthesis and structure. *Plant Physiology* **153**: 895–905.

Vardakou M, Dumon C, Murray JW, Christakopoulos P, Weiner DP, Juge N, Lewis RJ, Gilbert HJ, Flint JE. 2008. Understanding the structural basis for substrate and inhibitor recognition in eukaryotic GH11 xylanases. *Journal of Molecular Biology* **375**: 1293–1305.

Watanabe T, Koshijima T. 1988. Evidence for an ester linkage between lignin and glucuronic-acid in lignin carbohydrate complexes by DDQ-oxidation. *Agricultural and Biological Chemistry* **52**: 2953–2955.

Whitney SEC, Brigham JE, Darke AH, Reid JSG, Gidley MJ. 1998. Structural aspects of the interaction of mannan-based polysaccharides with bacterial cellulose. *Carbohydrate Research* **307**: 299–309.

Willför S, Sundberg A, Hemming J, Holmbom B. 2005. Polysaccharides in some industrially important softwood species. *Wood Science and Technology* **39**: 245–257.

Willfor S, Sundberg K, Tenkanen M, Holmbom B. 2008. Spruce-derived mannans – a potential raw material for hydrocolloids and novel advanced natural materials. *Carbohydrate Polymers* **72**: 197–210.

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

Yamasaki T, Enomoto A, Kato A, Ishii T, Shimizu K. 2011. Structural unit of xylans from sugi (*Cryptomeria japonica*) and hinoki (*Chamaecyparis obtusa*). *Journal of Wood Science* **57**: 76–84.

Yu L, Lyczakowski JJ, Pereira CS, Kotake T, Yu XL, Li A, Mogelsvang S, Skaf MS, Dupree P. 2018. The patterned structure of galactoglucomannan suggests it may bind to cellulose in seed mucilage. *Plant Physiology* **178**: 1011–1026.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Expression profiling for *Picea abies* (*Pa*) GUX1.

Fig. S2 Expression profiling for *Picea abies* (*Pa*) GUX2.

Fig. S3 Expression profiles of conifer genes likely to be involved in softwood biosynthesis.

Fig. S4 Western blot analysis.

Fig. S5 Sequence similarity matrix generated based on whole length sequences of conifer GUX1 and GUX2 enzymes analysed in this study.

Fig. S6 Annotation of the XUXUXX structure.

Table S1 Primers used in this study.

Table S2 Database codes for gymnosperm GUX sequences used in this study.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.