nature biotechnology

Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis

Ajaya K Biswal[1–3], Melani A Atmodjo[1–3], Mi Li[3–5], Holly L Baxter[3,6], Chang Geun Yoo[3–5], Yunqiao Pu[3–5], Yi-Ching Lee[3,7,13], Mitra Mazarei[3,6], Ian M Black[2], Ji-Yi Zhang[3,7,13], Hema Ramanna[3,7,13], Adam L Bray[3,8], Zachary R King[3,8], Peter R LaFayette[3,8], Sivakumar Pattathil[2,3,13], Bryon S Donohoe[3,9], Sushree S Mohanty[2,3], David Ryno[2,3], Kelsey Yee[3,4], Olivia A Thompson[3,4], Miguel Rodriguez Jr.[3,4], Alexandru Dumitrache[3,4], Jace Natzke[3,4], Kim Winkeler[3,10], Cassandra Collins[3,10], Xiaohan Yang[3,4], Li Tan[1–3], Robert W Sykes[3,9,13], Erica L Gjersing[3,9], Angela Ziebell[3,9], Geoffrey B Turner[3,9,13], Stephen R Decker[3,9], Michael G Hahn[2,3,11], Brian H Davison[3,4], Michael K Udvardi[3,7], Jonathan R Mielenz[3,4], Mark F Davis[3,9], Richard S Nelson[3,7], Wayne A Parrott[3,8], Arthur J Ragauskas[3–5,12], C Neal Stewart Jr[3,6] & Debra Mohnen[1–3]

Cell walls in crops and trees have been engineered for production of biofuels and commodity chemicals, but engineered varieties often fail multi-year field trials and are not commercialized. We engineered reduced expression of a pectin biosynthesis gene (*Galacturonosyltransferase 4*, *GAUT4*) in switchgrass and poplar, and find that this improves biomass yields and sugar release from biomass processing. Both traits were maintained in a 3-year field trial of *GAUT4*-knockdown switchgrass, with up to sevenfold increased saccharification and ethanol production and sixfold increased biomass yield compared with control plants. We show that GAUT4 is an α-1,4-galacturonosyltransferase that synthesizes homogalacturonan (HG). Downregulation of *GAUT4* reduces HG and rhamnogalacturonan II (RGII), reduces wall calcium and boron, and increases extractability of cell wall sugars. Decreased recalcitrance in biomass processing and increased growth are likely due to reduced HG and RGII cross-linking in the cell wall.

The efficient conversion of non-food and non-animal-feed lignocellulosic biomass into biofuels[1,2] could decrease the greenhouse gas emissions that are associated with fossil fuel use. Plant cellulosic biomass is structurally complex and recalcitrant to deconstruction, making it necessary to use chemical and enzymic pretreatments to convert cell wall polymers into oligo- and monosaccharides for processing by microorganisms into fuel[3]. One strategy to overcome biomass recalcitrance is to engineer feedstocks with cell walls that are more easily broken down to release sugars while maintaining plant architecture and biomass yields. Progress has been made by modifying lignin[4–7] and polysaccharide[8] structures for improved biofuel yield in switchgrass (*Panicum virgatum*) and poplar (*Populus deltoides*) without compromising plant growth (**Supplementary Table 1**). In some studies, transgenic plants show preliminary evidence of increased growth[8]. But conclusive evidence that this approach has substantive potential for commercial translation with increased yield in woody and grass species has been lacking.

Plant cell walls comprise polysaccharide, lignin, proteoglycan, and glycoprotein matrices that differ among species, organs, tissues, and even cell types[9]. The cell wall matrix is thought to be composed of the polysaccharides cellulose, hemicellulose, and pectin, with the polyphenolic lignin present in secondary walls. However, there is increasing evidence that some wall glycans exist as both polysaccharides and proteoglycans, and that each specific type of pectic and hemicellulosic polysaccharide may comprise multiple unique polymers with unique functions in primary versus secondary walls or in different cell types[10–13]. Woody biomass is composed largely of secondary walls that comprise cellulose, the hemicellulose glucuronoxylan, and lignin[14], and substantial amounts of pectin-rich primary walls[8]. Grass biomass is abundant in cellulose and the hemicellulose arabinoxylan, with very small amounts of pectin[15,16].

Pectic polysaccharides are the most structurally complex plant cell wall glycans[10]. An estimated 67 different transferase activities are

[1]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia, USA. [2]Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia, USA. [3]DOE-BioEnergy Science Center (BESC), Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA. [4]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA. [5]UT-ORNL Joint Institute for Biological Sciences, Oak Ridge, Tennessee, USA. [6]Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee, USA. [7]Noble Research Institute, Ardmore, Oklahoma, USA. [8]Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Athens, Georgia, USA. [9]National Renewable Energy Laboratory, Golden, Colorado, USA. [10]ArborGen, Inc., Ridgeville, South Carolina, USA. [11]Department of Plant Biology, University of Georgia, Athens, Georgia, USA. [12]Department of Chemical and Biomolecular Engineering & Department of Forestry, Wildlife, and Fisheries, University of Tennessee, Knoxville, Tennessee, USA. [13]Current addresses: National Synchrotron Radiation Research Center, Hsinchu Science Park, Hsinchu City, Taiwan (Y.-C.L.); Bayer CropScience Division, Research Triangle Park, North Carolina, USA (J.-Y.Z.); H-IVs-204, National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi, India (H.R.); Mascoma LLC (Lallemand Inc.), Lebanon, New Hampshire, USA (S.P.); Nuclear Materials Science, Los Alamos National Laboratory, Los Alamos, New Mexico, USA (R.W.S); Nu Mark LLC, Richmond, Virginia, USA (G.B.T.). Correspondence should be addressed to D.M. (dmohnen@ccrc.uga.edu).

Received 14 June 2017; accepted 2 January 2018; published online 12 February 2018; doi:10.1038/nbt.4067

needed to synthesize branched rhamnogalacturonan I and II (RGI and RGII) and non-branched homogalacturonan (HG), which are the main pectic glycans. The number and type of unique pectic polymers in the wall remain unknown, with estimates ranging from one matrix of covalently linked HG, RGI, and RGII to multiple numbers and types of unique polymers, each with specific pectic glycan backbones and side chains, and with unique structural and functional roles[10].

We set out to produce reduced recalcitrance in biomass in both grass and woody feedstocks. We focused on the minor cell wall component pectin, and identified a putative pectin biosynthetic gene *GAUT4*, present in poplar and switchgrass, that might be involved in the synthesis of structures associated with cell wall recalcitrance. Engineering of poplar and switchgrass to express less *GAUT4* produced plants with reduced recalcitrance and increased growth in both the greenhouse and field. We also characterized cell wall structural changes in *GAUT4*-knockdown (KD) plants and analyzed the enzymatic activity of GAUT4, to understand why downregulation of GAUT4-synthesized pectic glycan(s) improves cell wall extractability and growth of both woody and grass biofuel feedstocks.

RESULTS

GAUT4-KD increases sugar release

GAUT4, a member of the GAUT family[17] of putative[18] and proven[17,19] pectin biosynthetic galacturonosyltransferases, was identified as a candidate recalcitrance-associated gene due to high expression in wood xylem[14], dicot stems[18], and grass shoots[20]. GAUT4 shares the GAUT family clade A with GAUT1, a proven homogalacturonan:α-1,4-galacturonosyltransferase (HG:GalAT)[17,19]. *GAUT1*, *GAUT4*, and *GAUT7* are the only *GAUT* genes with more than two homologs in rice, which might indicate a more prominent role in grasses for these genes than for other *GAUT* genes.

We used RNA interference (RNAi) to downregulate the expression of switchgrass (*Panicum virgatum, Pavir.J36995*; *Pv*) and rice (*Oryza sativa, LOC_Os08g23780*) *GAUT4* homologs that are most similar to *Arabidopsis GAUT4*, and of the *Populus deltoides GAUT4* homolog most similar to the most highly expressed *GAUT4* in *Populus trichocarpa* (*Pt*) wood (*Potri.016G001700*) (**Supplementary Fig. 1a–c**). The *PvGAUT4* RNAi construct was used in both switchgrass and rice, and the *PtGAUT4* construct was used in *P. deltoides* (**Supplementary Fig. 1b**,**c**). Two *PvGAUT4*-KD (2A/2B and 4A), two *OsGAUT4*-KD (2A/2B and 7A/7B) and 15 *PdGAUT4*-KD (AB23.1 to AB23.15) transgenic lines were produced and grown in the greenhouse (Online Methods). Transcript expression of respective target genes was reduced 56–68% (average 62%, $P < 0.001$) in *PvGAUT4*-KD, 66–75% (average 71%, $P < 0.001$) in *OsGAUT4*-KD, and 49–78% (average 64%, $P < 0.001$) in *PdGAUT4*-KD lines compared to their corresponding controls, while the non-target *GAUT4* homologs were unaffected (**Supplementary Fig. 1d–f**).

Saccharification of transgenic and wild-type (WT) biomass was measured using hot water pretreatment and enzymatic hydrolysis. Glucose release increased by 10–15% (average 13%, $P < 0.05$) in *PvGAUT4*-KD lines, 24–38% (average 31%, $P < 0.001$) in *OsGAUT4*-KD lines, and 4–9% (average 7%, $P < 0.05$) in 12 *PdGAUT4*-KD lines compared to controls (**Fig. 1a**,**d**,**g** and **Supplementary Fig. 2a**). Xylose release increased by 14–23% (average 18%, $P < 0.05$) in *PvGAUT4*-KD lines, and 5–8% (average 6%, $P < 0.05$) in 12 *PdGAUT4*-KD lines; however, no increase in xylose was observed in *OsGAUT4*-KD lines (**Fig. 1b**,**e**,**h** and **Supplementary Fig. 2b**). Total sugar release increased by 14–15% (average 15%, $P < 0.05$) in *PvGAUT4*-KD lines, 11–17% (average 14%, $P < 0.001$) in *OsGAUT4*-KD lines, and 7–8% (average 7%, $P < 0.05$) in 13 *PdGAUT4*-KD lines compared to controls (**Fig. 1c**,**f**,**i** and **Supplementary Fig. 2c**). The lignin content was unchanged in the transgenic lines, whereas the syringyl/guaiacyl (S/G) ratios were increased only in *PdGAUT4*-KD lines (**Supplementary Fig. 2d–g**).

Microbial bioconversion of *PvGAUT4*-KD and WT biomass was assessed using yeast-based separate hydrolysis and fermentation (SHF) (**Fig. 1j–l**). A switchgrass transgenic line downregulated in caffeic acid 3-*O*-methyltransferase (*PvCOMT3*-TG)[21] was also analyzed as a comparator reduced-recalcitrance line. Relative to WT, the *PvGAUT4*-KD lines showed significantly improved bioconversion to ethanol as good as, or better than, the *PvCOMT3*-TG line (**Fig. 1j–k**), and faster ethanol production rates than the controls' (**Fig. 1l**).

GAUT4-KD plants grow better than controls

Unexpectedly, all *GAUT4*-KD grasses and trees showed increased growth compared to controls in the greenhouse (**Fig. 2a–c** and **Supplementary Figs. 3** and **4**). Switchgrass *PvGAUT4*-KD plants had 12–19% (average 16%, $P < 0.05$) increased plant height compared to WT (**Fig. 2d**). *PvGAUT4*-KD plants also produced significantly more and larger tillers beginning in week 5 (**Supplementary Fig. 3i–k**), resulting in overall 150–190% (average 170%, $P < 0.001$) more aerial biomass than WT (**Fig. 2e**,**f** and **Supplementary Fig. 3**). Similarly, the rice *OsGAUT4*-KD lines had 21–24% (average 23%, $P < 0.001$) increased plant height, 42–70% (average 59%, $P < 0.001$) increased tiller number, and 43–69% (average 56%, $P < 0.001$) increased dry biomass yield compared to WT (**Fig. 2g–i**).

Plant height and stem diameter were increased 8–46% and 11–49% (average 31%, $P < 0.001$), respectively, in 3-month-old *PdGAUT4*-KD trees compared with controls (**Fig. 2j**,**k** and **Supplementary Fig. 4a–c**), continued to increase throughout the 9-month growth period, and were inversely correlated with *GAUT4* transcript levels (**Supplementary Fig. 4d**,**e**). Three-month-old *PdGAUT4*-KD trees also had larger leaves (average 48%, $P < 0.001$) and higher relative water content (average 9%, $P < 0.05$), accompanied by 38–49% (average 44%, $P < 0.001$) greater total aerial biomass, 19–26% (average 23%, $P < 0.05$) greater stem fresh weight, and 36–49% (average 43%, $P < 0.001$) greater stem dry weight compared to controls (**Fig. 2l**, **Supplementary Fig. 4f–o** and **Supplementary Note 1**).

Switchgrass *PvGAUT4*-KD traits in the field

In a 3-year field study, *PvGAUT4* transcript expression remained reduced in the *PvGAUT4*-KD lines by 48–55% (average 52%), 54–64% (average 60%), and 65–75% (average 70%), ($P < 0.001$) in years 1, 2, and 3, respectively, compared to WT (**Supplementary Figs. 1g** and **3l**), indicating stable RNA silencing over time. No growth differences were observed among plants in year 1 when they were being established in the field. In subsequent years, transgenic plants had 15–64% (year-2 average 18%, $P < 0.05$; year-3 average 57%, $P < 0.001$) increased plant width with 28–102% (year-2 average 38%, $P < 0.05$; year-3 average 91%, $P < 0.001$) greater tiller numbers, resulting in doubled biomass yield in year 2 (year-2 average 89%, $P < 0.001$) and sixfold more in year 3 (year-3 average 435%, $P < 0.001$), compared to WT (**Fig. 2m–p**). When assessed in years 2 and 3, *PvGAUT4*-KD plants (year-2 average 1.29 ± 0.3; year-3 average 3.0 ± 3.9) were similar ($P > 0.05$) to WT (year-2 average 0.2 ± 0.4; year-3 average 3.0 ± 3.9) in rust disease (*Puccinia emaculata*) severity, measured as percentage leaf area covered by uredia (**Supplementary Fig. 3m–o**), which may indicate that *PvGAUT4* downregulation does not reduce resistance to this pathogen. Furthermore, field-grown *PvGAUT4*-KD biomass had higher saccharification efficiency (year-3 average 483%, $P < 0.001$) and ethanol

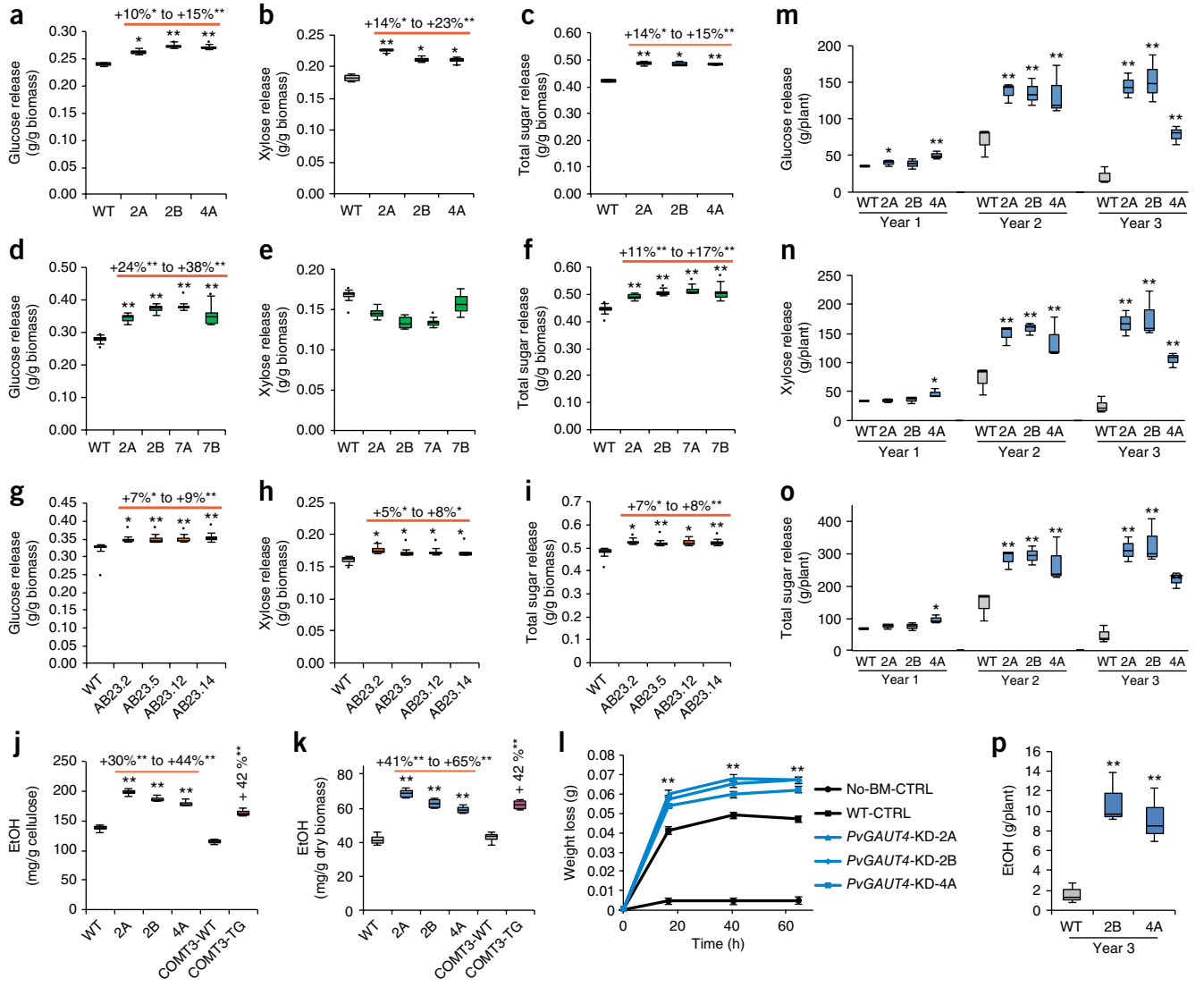

Figure 1 Saccharification yield from switchgrass, rice, and poplar *GAUT4*-KD lines and bioconversion of switchgrass *GAUT4*-KD to ethanol. (**a**–**i**) Glucose, xylose, and total sugar release, respectively, from greenhouse-grown *PvGAUT4*-KD (**a**–**c**), *OsGAUT4*-KD (**d**–**f**), and *PdGAUT4*-KD (**g**–**i**) lines compared to controls. Materials used were 3-month-old switchgrass tillers, 3-month-old rice aerial biomass, and 9-month-old poplar wood. (**j**,**k**) Ethanol yield from yeast-based separate hydrolysis and fermentation (SHF) of hot-water-pretreated greenhouse-grown switchgrass *PvGAUT4*-KD lines 2A, 2B and 4A and wild-type (WT) biomass, shown as mg EtOH/g cellulose (**j**) and mg EtOH/g dry biomass (**k**). Biomass hydrolyses were carried out with commercial enzyme blends (Ctec2, Htec2, and 188; Novozymes) at 50 °C, pH 4.5 for 5 d; fermentation with *S. cerevisiae* D5A at 35 °C, pH 4.5, 150 r.p.m. Switchgrass caffeic acid 3-*O*-methyltransferase knockdown transgenic (*COMT3*-TG) and corresponding WT lines (*COMT3*-WT) were included as a reference[21]. (**l**) Time-course fermentation of whole biomass in **k** measured as weight loss of fermentation bottles. Fermentation was performed with three technical replicates of two independently grown biomass samples of three genetically identical clones of a single transgenic event ($n = 6$). (**m**–**o**) Glucose, xylose, and total sugar release, respectively, from field-grown *PvGAUT4*-KD compared to controls in the first (2013), second (2014), and third (2015) growing seasons. (**p**) Ethanol yield from yeast-based separate hydrolysis and fermentation (SHF) of hot water-pretreated field-grown year 3 switchgrass *PvGAUT4*-KD and WT biomass, shown as grams EtOH/plant. For greenhouse-grown biomass: WT and *PvGAUT4*-KD lines, $n = 4$; WT and *OsGAUT4*-KD lines, $n = 6$; WT *P. deltoides*, $n = 25$; *PdGAUT4*-KD lines, $n = 15$. For field-grown biomass *PvGAUT4*-KD and WT: $n = 3$. Data are presented as box plots showing the median as well as the 25th and 75th percentiles. Ends of whiskers are set at 1.5 * IQR (interquartile range) above and below the third and first quartiles, respectively. Statistical analysis was done using one-way analysis of variance (ANOVA) followed by Tukey's multiple comparison test for greenhouse-grown biomass, or by Fisher's least significant difference method for field-grown biomass (*$P < 0.05$, **$P < 0.001$).

yield (year-3 average 517%, $P < 0.001$) per plant throughout the 3-year field trial, being approximately sevenfold greater than WT's in year 3 (**Fig. 1m**–**p**). Of note, while unchanged in the greenhouse, the field-grown *PvGAUT4*-KD lines had 10–12% higher ($P < 0.05$) lignin S/G ratio than WT in year 3 (**Supplementary Fig. 2h**,**i**). Overall, the field trial reproduced the *PvGAUT4*-KD phenotypes observed in the greenhouse.

GAUT4-KD reduces homogalacturonan content

To understand the mechanism underlying improved saccharification and growth of *GAUT4*-KD plants, we first analyzed GAUT4 enzymatic function. Transient expression in *Nicotiana benthamiana* leaves of *Arabidopsis*, poplar, and switchgrass GAUT4 increased HG:GalAT activity in both microsomes and partially purified protein preparations (**Fig. 3a**, **Supplementary Note 2** and **Supplementary Fig. 5**).

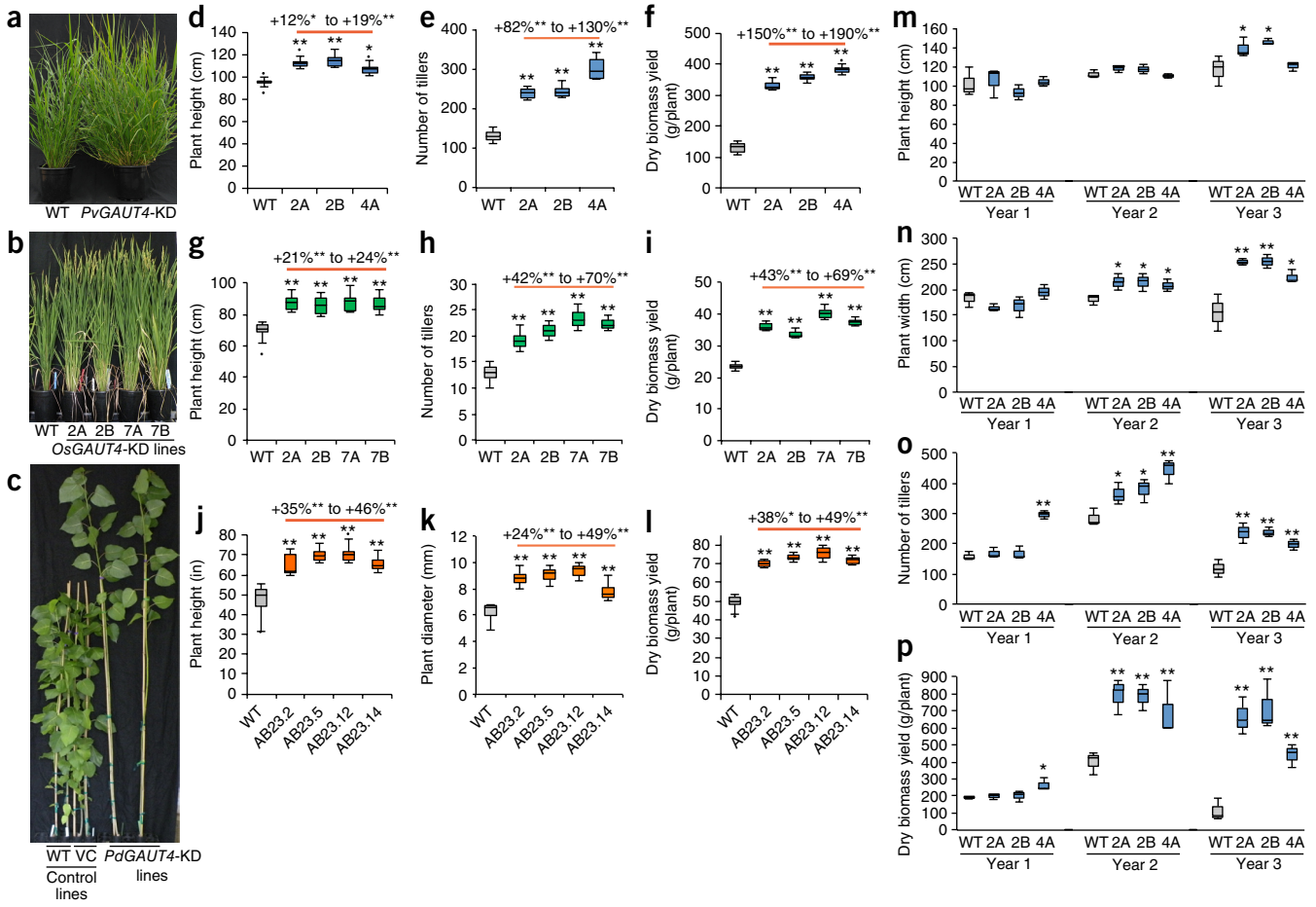

Figure 2 Growth and yield of switchgrass, rice, and poplar *GAUT4*-KD transgenic lines. (**a**–**c**) Growth of greenhouse-grown 3-month-old WT and *PvGAUT4*-KD switchgrass grown in 19-liter pots (**a**), 3-month-old WT control and *OsGAUT4*-KD rice (**b**), and 3-month-old WT (left two plants), vector control (VC, middle two plants) and *PdGAUT4*-KD (right two plants) poplar (**c**). (**d**–**i**) Height, number of tillers and dry aerial biomass of greenhouse-grown switchgrass ($n = 10$) (**d**–**f**) and rice ($n = 15$) plants (**g**–**i**). (**j**–**l**) Height (**j**), radial growth (**k**), and dry aerial biomass (**l**) of greenhouse-grown poplar plants (WT, $n = 25$ for height and radial growth, $n = 15$ for dry biomass yield; *PdGAUT4*-KD lines, $n = 15$). (**m**–**p**) End-of-season tiller height (**m**), plant width (**n**), number of tillers (**o**), and dry aerial biomass (**p**) per plant of field-grown switchgrass ($n = 3$) over the 3-year field trial. The year 1 and 2 dry biomass data for lines 2B and 4A were also part of a study by Dumitrache *et al.*[46]. Data are presented as box plots showing the median as well as the 25th and 75th percentiles. Ends of whiskers are set at 1.5 * IQR above and below the third and first quartiles, respectively. Statistical analysis was done using one-way analysis of variance (ANOVA) followed by Tukey's multiple comparison test for the greenhouse-grown biomass, or by Fisher's least significant difference method for field-grown biomass (*$P < 0.05$, **$P < 0.001$).

Furthermore, *PvGAUT4*-KD leaf microsomes had substantially reduced total HG:GalAT activity (by 26% at 45 min, $P < 0.05$) compared to WT (**Fig. 3b**). These results support the view that GAUT4 is an HG:GalAT, and suggest that decreased recalcitrance and increased growth correlate with reduced cell wall HG content.

Prior analyses of cell wall structural changes in multiple GAUT mutants[8,18,22] led to the hypothesis that different GAUTs synthesize specific HG domains in distinct polysaccharides and/or proteoglycans[10,11]. To determine the effect of downregulated *GAUT4* expression, we isolated cell walls as alcohol-insoluble residues (AIR) from *GAUT4*-KD lines and analyzed them for the glycosyl residue composition (**Table 1**). The only glycosyl residue significantly decreased in total AIR from the *GAUT4*-KD lines was galacturonic acid (GalA), the pectin signature sugar residue and the sole sugar residue in HG. The total AIR GalA content was reduced 74–91% (average 81%, $P < 0.001$) in switchgrass, 69–78% (average 74%, $P < 0.001$) in rice, and 48–57% (average 52%, $P < 0.001$) in poplar *GAUT4*-KD lines compared to controls, consistent with a function of GAUT4 in the synthesis of major HG glycans in all three species.

To determine the type of HG-containing polymer affected in the *GAUT4*-KD biomass, we sequentially extracted switchgrass and poplar AIR with increasingly harsh solvents[8,23] (Online Methods). GalA was the only sugar significantly reduced in all extracts and in the final insoluble pellet, reduced by 64–97% (average 78%, $P < 0.001$) for switchgrass and 56–77% (average 64%, $P < 0.001$) for poplar *GAUT4*-KD biomass compared to controls (**Supplementary Tables 2–5**). Several other sugars were also significantly reduced in one or more, but not all, wall extracts from transgenic lines compared to controls, including mannose (Man) and galactose (Gal) for switchgrass, and Man, Gal, xylose (Xyl), and glucuronic acid (GlcA) for poplar (**Supplementary Note 3**). It is worth noting that in a study of *GAUT4*-silenced tomato[24], using a smaller set of cell wall analysis techniques on only two cell wall extracts, no changes were observed in leaf cell wall sugar composition, although there were reduced amounts of rhamnose (Rha), arabinose (Ara), and Gal in tomato fruits.

Glycosyl residue linkage analysis of switchgrass and poplar *GAUT4*-KD biomass (**Supplementary Tables 6** and **7**) revealed 4-linked and terminal GalA, the constituents of the HG backbone, as the major linkages

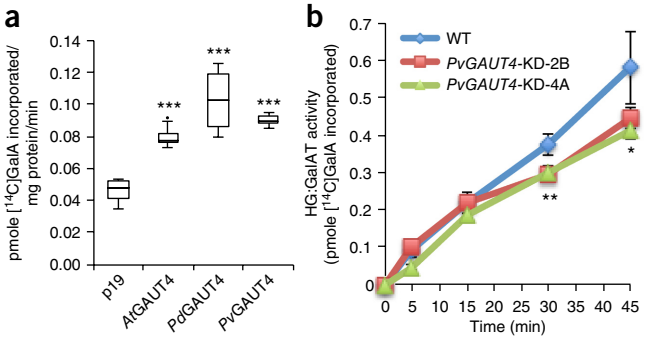

Figure 3 HG:GalAT activity of *Arabidopsis* (*At*GAUT4), poplar (*Pd*GAUT4), and switchgrass (*Pv*GAUT4) recombinant GAUT4 transiently expressed in *N. benthamiana* and HG:GalAT activity in switchgrass WT and KD lines. (**a**) Pmole [^{14}C]GalA-radiolabeled HG synthesized by microsomal membranes (100 µg protein) prepared from *N. benthamiana* leaves transiently co-expressing *GAUT4* constructs and the silencing suppressor p19 (ref. 52), and leaves expressing p19 alone, from 3-h reactions containing UDP-[^{14}C]GalA and exogenous HG acceptors. (**b**) HG:GalAT activity in leaf microsomes prepared from switchgrass WT and *PvGAUT4*-KD lines. Data are duplicate samples from two independent experiments ($n = 4$), presented as box plots showing the median as well as the 25th and 75th percentiles. Ends of whiskers are set at 1.5 * IQR above and below the third and first quartiles, respectively. Statistical analysis was done using one-way analysis of variance (ANOVA) followed by Tukey's multiple comparison test (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

reduced in all wall extracts tested. Furthermore, glycome profiling analysis using non-cellulosic wall-polysaccharide-specific monoclonal antibodies (mAbs)[13,23] showed substantially reduced epitope recognition by multiple HG-backbone-specific mAbs in all *GAUT4*-KD wall extracts compared to WT (**Supplementary Figs. 6** and **7** and **Supplementary Note 4**). Lastly, immunogold-labeling of switchgrass stem cross-sections with HG-specific mAbs decreased significantly in *PvGAUT4*-KD samples compared with WT (**Supplementary Fig. 8**). Taken together, these data support the conclusion that downregulation of *GAUT4* reduces cell wall HG content in both switchgrass and poplar.

GAUT4-KD increases polymer accessibility and cell growth

We tested whether cell wall integrity was compromised in *GAUT4*-KD biomass. Yield data of sequential AIR extracts showed that although the transgenic biomass and control biomass had comparable total amounts of AIR, substantially more material (9–49%, average 25%, $P < 0.001$, for *PvGAUT4*-KD; 10–32%, average 20%, $P < 0.001$, for *PdGAUT4*-KD) was recovered in the wall extracts from *GAUT4*-KD lines than from WT, and less material (10–20%, $P < 0.001$) remained in the insoluble pellet (**Supplementary Fig. 9** and **Supplementary Note 3**). Furthermore, glycome profiles of both *PvGAUT4*-KD and *PdGAUT4*-KD biomass showed altered extraction patterns for multiple wall polysaccharides in addition to HG, including xylan, RGI, and arabinogalactan (**Supplementary Figs. 6** and **7** and **Supplementary Note 4**). These results show that reduced *GAUT4* expression and the associated GAUT4-catalyzed HG synthesis compromise cell wall integrity, resulting in more easily extracted walls with less material retained in the final pellet.

We next tested whether *GAUT4*-KD cell walls might facilitate enzyme access for saccharification and ethanol production. First, we showed that *PvGAUT4*-KD biomass had a greater water-swelling response, with a greater volume difference between dry and wet switchgrass stem sections than WT (**Supplementary Fig. 10a**). Second, scanning electron microscopy of hot-water-pretreated biomass revealed more damage, seen as torn cell walls, in *PvGAUT4*-KD biomass than in WT (**Supplementary Fig. 10b**,**c**). More readily fragmented walls could increase enzyme access to biomass during saccharification. We tested accessibility to cellulose in two *PvGAUT4*-KD lines using the Simon's stain procedure[25] and found that it was increased compared to WT (**Supplementary Fig. 11** and **Supplementary Note 5**). Conversely, neither cellulose content nor chain length was affected in the *PvGAUT4*-KD lines. Altogether, these results indicate that *PvGAUT4*-KD biomass has looser, more porous cell walls with better enzyme accessibility and improved digestion.

We next tested how *GAUT4*-KD cell walls responded to turgor pressure. Microscopic analysis of stem cross-sections and isolated xylem cells revealed increased amounts of vascular tissues and larger xylem cells in *PdGAUT4*-KD compared with WT (**Supplementary Figs. 12**

Table 1 Glycosyl residue composition of total cell walls from 3-month-old switchgrass tillers, 9-month-old poplar stems and 3-month-old rice aerial biomass

	Glycosyl residue composition of AIR (mole % ± SE)								
	Ara	Rha	Fuc	Xyl	GlcA	GalA	Man	Gal	Glc
Switchgrass									
WT	15.93 ± 0.53	1.32 ± 0.11	0.88 ± 0.09	59.17 ± 1.56	0.86 ± 0.04	2.47 ± 0.11	1.44 ± 0.08	5.31 ± 0.21	12.62 ± 0.56
2A	**17.59 ± 0.43***	**2.83 ± 0.13***	1.00 ± 0.08	62.41 ± 1.32	1.01 ± 0.03	**0.65 ± 0.05****	**1.79 ± 0.09***	4.90 ± 0.23	**7.82 ± 0.67****
2B	**18.12 ± 0.47***	**2.10 ± 0.16***	**0.62 ± 0.07***	59.51 ± 1.45	1.03 ± 0.02	**0.55 ± 0.05****	1.59 ± 0.11	**4.73 ± 0.27***	11.75 ± 0.78
4A	**18.41 ± 0.61***	**2.31 ± 0.11***	**0.21 ± 0.02***	61.92 ± 1.74	**2.24 ± 0.03****	**0.21 ± 0.03****	**2.49 ± 0.12****	4.98 ± 0.31	**7.03 ± 0.81****
Poplar									
WT	3.23 ± 0.09	1.96 ± 0.08	0.10 ± 0.010	45.86 ± 1.13	0.32 ± 0.010	7.01 ± 0.26	5.89 ± 0.39	3.25 ± 0.26	32.38 ± 1.01
V control-1	3.14 ± 0.10	1.84 ± 0.09	0.09 ± 0.005	46.01 ± 1.02	0.28 ± 0.005	6.88 ± 0.24	6.02 ± 0.31	3.18 ± 0.24	32.57 ± 1.11
AB23.2	3.69 ± 0.12	2.26 ± 0.07	0.10 ± 0.006	48.23 ± 0.99	**0.45 ± 0.006****	**3.62 ± 0.10****	5.97 ± 0.29	3.41 ± 0.21	32.27 ± 0.98
AB23.5	3.51 ± 0.13	2.19 ± 0.08	0.08 ± 0.007	**49.34 ± 0.89***	**0.56 ± 0.003****	**3.35 ± 0.15****	**5.31 ± 0.22***	3.55 ± 0.31	32.11 ± 1.07
AB23.12	3.68 ± 0.19	2.37 ± 0.10	0.09 ± 0.010	47.36 ± 0.95	**0.39 ± 0.008***	**3.04 ± 0.09****	5.78 ± 0.38	**4.10 ± 0.11****	33.19 ± 1.19
AB23.14	3.49 ± 0.11	2.33 ± 0.08	0.10 ± 0.008	48.71 ± 1.29	**0.41 ± 0.005****	**3.18 ± 0.17****	5.99 ± 0.45	**3.74 ± 0.13***	32.05 ± 0.88
Rice									
WT	18.21 ± 0.61	1.78 ± 0.05	0.90 ± 0.02	55.66 ± 1.21	1.54 ± 0.07	4.10 ± 0.11	1.29 ± 0.04	5.91 ± 0.21	10.61 ± 0.62
2A	**23.86 ± 0.55****	**2.28 ± 0.03***	0.62 ± 0.02	55.78 ± 1.11	**2.17 ± 0.02***	**1.29 ± 0.05****	**1.93 ± 0.05***	5.11 ± 0.18	**6.96 ± 0.34***
2B	**22.51 ± 0.69***	**2.32 ± 0.04***	0.62 ± 0.03	56.19 ± 1.08	**2.46 ± 0.04****	**0.96 ± 0.06****	1.21 ± 0.07	**4.38 ± 0.37***	9.35 ± 0.69
7A	**21.52 ± 0.61***	1.77 ± 0.05	**0.45 ± 0.03***	**62.53 ± 0.88***	**2.14 ± 0.05***	**0.89 ± 0.06****	1.31 ± 0.03	**4.08 ± 0.35****	**5.31 ± 0.52****
7B	**20.12 ± 0.38***	1.71 ± 0.07	0.74 ± 0.02	**63.06 ± 0.91***	1.77 ± 0.02	**1.07 ± 0.08****	1.58 ± 0.06	**4.84 ± 0.11***	**5.11 ± 0.65****

Data are average mole % sugars in alcohol insoluble residue (AIR) from switchgrass WT and *PvGAUT4*-KD lines 2A, 2B, and 4A; poplar WT, vector control, and *PdGAUT4*-KD lines; and rice WT and *OsGAUT4*-KD lines as measured by GC-MS of tetramethylsilane (TMS) derivatives. Data are three technical replicates of each of four independently grown biomass samples of three genetically identical clones of a single transgenic event ($n = 12$). Bold values with asterisk(s) are *GAUT4*-KD line values significantly different from controls. *$P < 0.05$, **$P < 0.001$ (one-way ANOVA followed by Tukey's multiple comparison test).

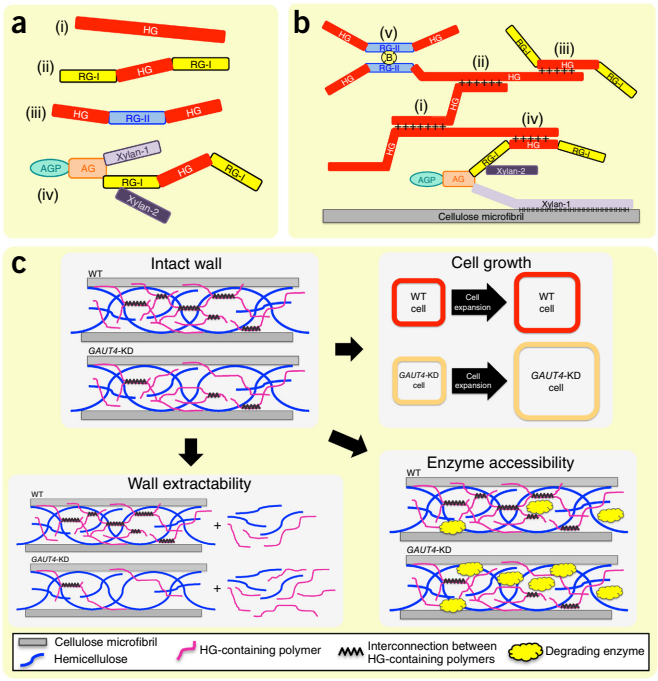

Figure 4 Model of GAUT4 function in cell wall extractability, porosity and cell size, and hypothesis for mechanism of *GAUT4*-KD in recalcitrance and plant growth. (**a**) Wall polymers known to contain HG glycan: (**a**, **i**) HG polysaccharide/oligosaccharide, (**a**, **ii**) HG domain connected to RGI through the backbone glycans, (**a**, **iii**) HG connected to RGII through their backbone glycans, (**a**, **iv**) HG glycan as part of the proteoglycan APAP1, whose whole structure is shown here. (**b**) Covalent and non-covalent bonds that can exist among HG-containing polymers in the wall, including (**a**, **i**–**iv**) ionic HG-calcium salt bridges between HG regions of independent HG molecules or between HG and polymers that contain HG[26], and (**a**, **v**) borate diester cross-linking between RGII regions[38]. The HG-containing polymers may further interact with other wall polymers, for example, the hemicellulose (e.g., xylan) moieties in HG-containing proteoglycans (e.g., APAP1 (ref. 11)) and cellulose microfibrils[34]. (**c**) Model depicting reduction of GAUT4-synthesized HG in *GAUT4*-KD versus WT biomass and hypothesized consequences on wall extractability, wall enzyme accessibility, and cell growth. A hypothesis: a reduction in HG in *GAUT4*-KD biomass leads to reduced *in muro* interconnections between HG-containing wall polymers resulting in reduced cell wall integrity and the associated (1) increased wall extractability, wall porosity, and accessibility to enzymes resulting in reduced recalcitrance for deconstruction and (2) loosened and/or reactive walls that are more pliable when turgor pressure is increased leading to greater cell growth.

and **13** and **Supplementary Note 6**). These results support a model in which the compromised cell wall integrity in the *GAUT4*-KD lines enables greater cell expansion, which in turn results in increased plant growth.

Less HG- and RGII-mediated cross-linking in *GAUT4*-KD walls

To analyze the mechanism(s) underlying increased growth and biomass saccharification, we assessed HG structure in the *GAUT4*-KD wall. HG is a pectic homopolymer of α-1,4-linked GalA residues partially modified by methylesterification and *O*-acetylation[10]. HG is synthesized in the Golgi, secreted into the wall in a highly methylesterified form, and partially demethylesterified *in muro* by pectin methylesterases[26–28]. Stretches of demethylesterified GalA residues in HG can form Ca^{2+}-salt bridges, resulting in HG dimers or larger complexes[29–31]. Some cell wall models depict HG as part of a single pectic structure covalently connected to the other pectic polysaccharides RGI and RGII[32,33]. However, the pectic polysaccharides that have been structurally defined have all been extracted from the wall using chemical and/or enzymatic treatments that fragment native pectin. Thus, a complete, intact native pectin has never been isolated. Based on structural analyses of isolated pectic fragments, four types of HG-containing polymers have been structurally confirmed[10] (**Fig. 4a**): (**i**) HG as an independent polysaccharide, (**ii**) polymers with HG and RGI connected via their backbones, (**iii**) polymers with HG and RGII connected via their backbones, and (**iv**) proteoglycans, such as, the xylan- and pectin-containing arabinogalactan protein APAP1 (ref. 11). Each of these pectic polymers may interact non-covalently and covalently with each other (**Fig. 4b**, **i**–**v**) and/or with other wall polymers[34], thus contributing to overall cell wall strength, flexibility, and integrity. Perturbations of these interactions are likely to affect wall performance[35,36].

Based on the confirmed pectin structures (**Fig. 4a**) and the data presented here, we propose a model (**Fig. 4c**) to explain how reduced production of HG decreases recalcitrance and increases growth in the *GAUT4*-KD lines. In WT plants, the HG made by GAUT4 associates either covalently or non-covalently (e.g., calcium cross-linking between HG molecules) with multiple wall polymers, leading to an integrated and cohesive wall that restricts the extent of cell wall expansion and limits growth, as well as resists biomass deconstruction. In *GAUT4*-KD plants, fewer GAUT4-synthesized HG-containing polymers are produced, leading to less polymer–polymer cross-linking and thus permitting more extensive cell expansion and greater plant growth. This model is consistent with the data presented here, and also with the taller plants and larger palisade parenchyma cells upon silencing *GAUT4* in tomato[24]. We propose that the decreased recalcitrance phenotypes of the *GAUT4*-KD biomass results from less polymer–polymer cross-linking, leading to increased wall porosity and reduced wall cohesiveness, which allow better movement of wall-degrading enzymes within the wall, including greater access to cellulose, and, therefore, a more easily deconstructed wall (**Fig. 4c**). Fewer GAUT4-synthesized HG-containing polymers could lead to less pectin coating of the cellulose microfibrils[34,37], making them more accessible to degrading enzymes.

To test the proposed model, we first asked whether other types of HG-containing polymers, besides independent HG polysaccharides (**Fig. 4a**, **i**), were also reduced in *GAUT4*-KD biomass. The glycosyl residue composition and linkage data (**Table 1** and **Supplementary Tables 2**–**7**) suggested that the pectic RGI content was not reduced, since the amount of RGI-specific sugars (i.e., Rha, Ara, and Gal) and linkages were not reduced consistently across different wall extracts and plant species, and even were increased in certain wall extracts. We therefore concluded that RGI-containing polymers (**Fig. 4a**, **ii**,**iv**) were not negatively affected in *GAUT4*-KD plants.

Contrary to RGI, which has an alternating Rha/GalA backbone, RGII has a backbone of HG and thus could have been affected in *GAUT4*-KD plants. Since RGII makes up only ~10% of total pectin and has a very complex structure with 12 different sugars in over 20 different linkages[38], the signals for its unique and low-abundance sugars are undetectable in standard glycosyl residue composition and linkage analyses of total AIR and wall extracts (**Table 1** and **Supplementary Tables 2** and **7**). We assessed RGII content by measuring the amount of two RGII-specific sugars, 2-*O*-Me-α-Xyl and 2-*O*-Me-α-L-Fuc[39], in the *PvGAUT4*-KD and WT oxalate and carbonate extracts that typically contain the bulk of extracted RGII[38,40]. The two RGII signature sugar residues were significantly reduced in the *PvGAUT4*-KD compared to WT wall extracts (**Fig. 5a**,**b**), indicating

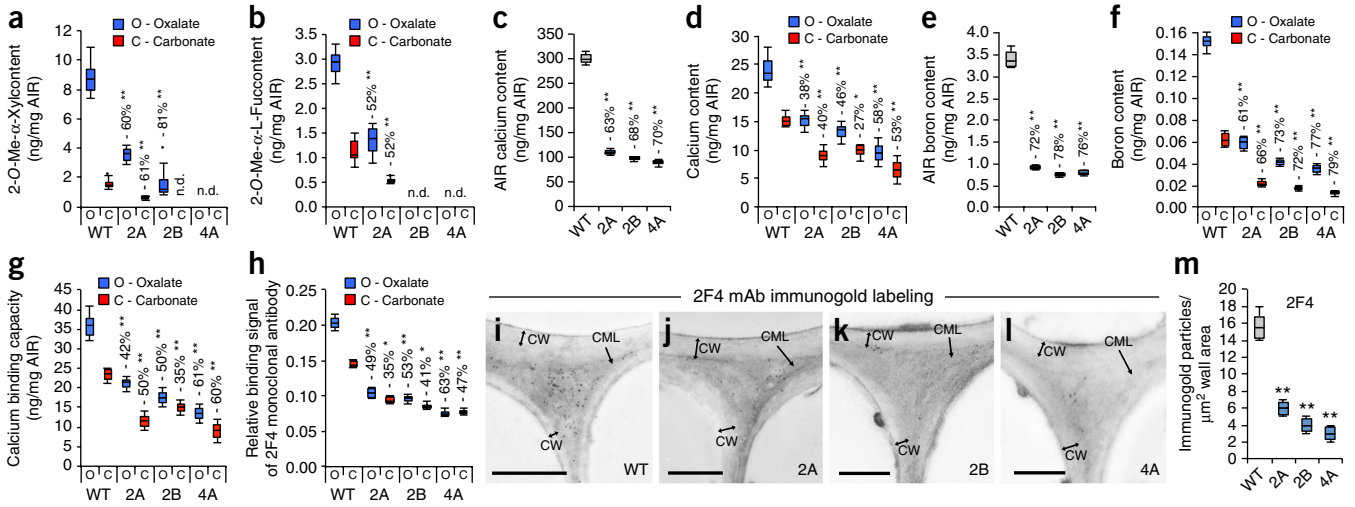

Figure 5 Pectin-mediated wall cross-linking is reduced in the *PvGAUT4*-KD cell walls. (**a**,**b**) RGII content is reduced in *PvGAUT4*-KD walls, as indicated by reduced contents of RGII-specific glycosyl residue 2-*O*-Me-α-Xyl (**a**) and 2-*O*-Me-α-L-Fuc (**b**) in the pectin-rich wall extracts from *PvGAUT4*-KD compared to WT biomass. (**c**–**f**) Calcium and boron contents measured in total AIR (**c**,**e**) and pectin-enriched wall extracts (**d**,**f**) from *PvGAUT4*-KD and WT biomass. (**g**) Calcium binding capacity of pectin-rich wall extracts as an estimate of the amount of calcium involved in cross-linking with HG. (**h**) ELISA of pectin-enriched wall extracts using 2F4 antibody. (**i**–**l**) Transmission electron microscopy of phloem cell wall cross-section of stems from WT and *PvGAUT4*-KD lines immunolabeled with the 2F4 antibody specific to calcium-cross-linked-HG. CW, cell wall; CML, compound middle lamellae. Black double-ended arrows indicate cell wall thickness; immunogold particles represent 2F4 antibody labeling. Scale bars, 0.5 μm. (**m**) Number of immunogold particles per μm^2 wall area observed as representative of 2F4 antibody epitope abundance in the WT and *PvGAUT4*-KD cell wall cross-sections shown in **a**–**d**. Data ($n = 6$) are presented as box plots showing the median as well as the 25th and 75th percentiles. Ends of whiskers are set at 1.5 * IQR above and below the third and first quartiles, respectively. Statistical analysis was done using one-way analysis of variance (ANOVA) followed by Tukey's multiple comparison test (*$P < 0.05$, **$P < 0.001$).

that RGII content was also greatly reduced in *PvGAUT4*-KD plants. We proposed that *PvGAUT4* silencing affects both HG (**Fig. 4a**, **i**) and HG-RGII (**Fig. 4a**, **iii**) polymers.

HG and RGII function in wall polysaccharide cross-linking via interactions with calcium and boron, respectively, and both are required for normal plant growth and development[27,38,41]. We tested whether HG–calcium and RGII–boron cross-linking (**Fig. 4b**, **i**,**ii**,**v**) decreased in *GAUT4*-KD cell walls. The calcium and boron content of total AIR and pectin-enriched extracts (oxalate and carbonate) were substantially decreased (**Fig. 5c**–**f**) in the *PvGAUT4*-KD walls compared to WT, strongly supporting the hypothesis that down-regulation of *GAUT4* reduces both HG–calcium and RGII–boron pectin cross-linking. Furthermore, we determined the relative amount of HG engaged in cross-linking with Ca^{2+} in the *PvGAUT4*-KD and WT biomass by measuring the wall calcium binding capacity[42] and the relative abundance of the 2F4 mAb epitopes (specific for calcium-cross-linked HG)[43,44] in the oxalate and carbonate extracts. Both parameters were significantly decreased in *PvGAUT4*-KD wall extracts compared to WT (**Fig. 5g**,**h**). Transmission electron microscopy analysis of immunogold-labeled switchgrass stem cross-sections (**Fig. 5i**–**m**) also revealed fewer 2F4 immunogold particles in *PvGAUT4*-KD than in WT phloem cell walls.

These results show that *GAUT4* downregulation decreases HG-Ca^{2+}-salt bridges and RGII–borate diester cross-linking in the wall, supporting the hypothesis that the observed increased wall porosity, enzyme accessibility to the cell wall, and plant growth in the *GAUT4*-KD lines are due to reduced pectin cross-linking in the wall. These results also support a role for unique GAUTs in the synthesis of specific cell wall HG-containing polymers[10] and the conclusion that GAUT4-synthesized HG resides in independent HG and/or HG-RGII wall polymers.

DISCUSSION

Our efforts to improve grass and woody biomass feedstock quality by RNA silencing of a pectin biosynthetic *GAUT4* gene increased biomass saccharification in greenhouse-grown engineered switchgrass, rice, and poplar, and increased ethanol production from switchgrass *GAUT4*-KD lines compared with WT. Biomass yield increased by 190%, 69%, and 49% in greenhouse-grown transgenic switchgrass, rice, and poplar, respectively. Importantly, improved traits were maintained in a 3-year field trial of switchgrass *GAUT4*-KD plants. Sugar composition and linkage analyses of *GAUT4*-KD biomass revealed 50–90% reduced HG content, indicating that one or more HG-containing pectic polymers were affected.

We showed that GAUT4 has pectin biosynthetic HG:GalAT activity and that knockdown expression of *GAUT4* in feedstocks reduces HG:GalAT activity and consequently cell wall pectic HG and RGII content. In turn, reduced HG- and RGII-mediated wall polymer cross-linking compromised overall wall integrity, leading to greater biomass saccharification and plant growth. These results support the proposal that at least some pectins have comparable roles in both grass and dicot cell walls, and in growth[36], and that *GAUT4* is an effective gene target for production of high-yielding biomass with improved properties for lignocellulosic biofuel production.

The use of lignocellulosic biomass for biofuels is hindered by its inherent recalcitrance to processing, owing to structurally complex cell walls[1,3], and suboptimal biomass yields on marginal land[45]. The phenotypes of the *GAUT4*-downregulated transgenic feedstocks address both issues simultaneously, and consistently across multi-year field trials, which is an improvement compared with other engineered trees and grasses (**Supplementary Table 1**). The *GAUT4*-KD transgenic switchgrass provided the greatest absolute ethanol and biomass yield compared to reduced-recalcitrance feedstock lines modified in lignin and C-1 metabolism[46]. Furthermore, compared with >35 other

feedstock lines modified in unique lignin, cellulose, hemicellulose, and pectin-associated genes, the switchgrass *GAUT4*-KD lines had the highest increase in biomass, sugar, and ethanol yield (g/plant basis; **Supplementary Table 1**). The performance of the transgenic switchgrass during the 3-year field study demonstrated the stability of the RNA-silencing-mediated genetic modification during multiple years of field growth[4,47]. In future, genome editing technologies for biofuel crop improvement may be preferred for genetic and regulatory reasons, as no foreign DNA would remain in the final feedstocks[48].

The increased saccharification of *GAUT4*-KD biomass provides direct evidence that controlled synthesis of a specific pectic polymer, in contrast to polymers containing both pectin and xylan constituents[8], or to post-synthesis modification of pectin in the wall[49–51], affects recalcitrance. Silencing of poplar *GAUT12*, another *GAUT* gene family member, also increased biomass saccharification and growth in *PdGAUT12*-KD lines[8], similar to the *GAUT4*-KD transgenic phenotypes. However, the *PdGAUT12*-KD biomass had concomitant reductions of both pectin (HG and RGI) and xylan, indicating that GAUT12 has a different biochemical function than GAUT4. This agrees with our hypothesis that the different GAUT proteins synthesize HG in unique pectin-containing glycans, each presumably with a unique function in the wall[10]. Interestingly, HG post-synthesis modification *in muro,* by overexpressing enzymes that degrade HG or reduce HG demethylesterification, also resulted in enhanced saccharification, supporting a role for pectin in cell wall integrity[49–51]. However, many of these post-synthesis modification strategies negatively affected plant growth, except for overexpression of a pectin methylesterase inhibitor (PMEI) in *Arabidopsis*[50]. The increased growth effect, however, was not observed when PMEI was overexpressed in wheat, leading the authors to conclude that HG methylesterification has a minor role in monocots[50]. This contrasts to our finding here that *GAUT4* downregulation increases saccharification and growth in both dicots and monocots.

The reduction of both HG and RGII in the *GAUT4*-KD biomass suggests that GAUT4-synthesized HG is used as the backbone for RGII synthesis, making GAUT4 the first HG:GalAT implicated in RGII synthesis. The *in vitro* synthesis of RGII by co-expression of recombinant GAUT4 along with RGII side-chain-synthesizing enzymes will be required to identify the mechanism of GAUT4 in RGII synthesis and in overall HG synthesis. For example, the current results cannot inform whether GAUT4 synthesizes both independent HG and HG-RGII polymers, or only an HG-RGII polymer. *In vitro* synthesis, genetic, and cell wall structural studies will be required to distinguish these possibilities.

Our data indicate that HG has multiple roles in diverse wall polymers, that different HG regions are synthesized by unique GAUTs, and that the functions of HG-containing polymers include cell wall structural integrity and plant growth. We establish that targeted engineering of plant cell wall polymers in general, and HG-containing wall polymers in particular, can loosen the walls by reducing inter- and intrapolymer cross-linking. This strategy can improve plant biomass yield and quality and is a step toward a bio-based economy.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank CCRC Analytical Services for glycosyl residue linkage analysis, W. Rottmann for leading the *Populus* transformation, L. Gunter for validation of *Populus* constructs, I. Gelineo-Albersheim for submission of BESC transformation pipeline file and, along with K. Hunt, for help in establishing the poplar greenhouse growth conditions, E. Chandler, R. Amos, and K. Engle for preparation of endopolygalacturonase, and B. Rockwell and C. Treager for assistance with cell wall isolation and extraction. We also thank B. Wolfe, M. Laxton, and the UT field staff for assistance with data collection and general field maintenance, and R. Millwood for assistance with the USDA APHIS BRS permit regulations. The work was primarily supported by BioEnergy Science Center grant DE-PS02-06ER64304, and partially by the Center for Bioenergy Innovation. The BioEnergy Science Center and the Center for Bioenergy Innovation are US Department of Energy Bioenergy Research Centers, supported by the Office of Biological and Environmental Research in the Department of Energy's Office of Science. The research was also partially funded by the Department of Energy Center Grant DE-SC0015662. The CCRC series of plant cell wall glycan-directed antibodies were generated with the support from the US National Science Foundation Plant Genome Program (grants DBI-0421683 and IOS-0923992).

AUTHOR CONTRIBUTIONS

A.K.B. participated in all aspects of the study, including line selection, plant phenotyping, expression study, tissue handling and distribution, and cell wall analysis, and wrote the manuscript. M.A.A. performed molecular cloning and generated heterologous expression constructs for poplar and switchgrass genes; designed and performed heterologous expression and enzymatic activity assays; and wrote the manuscript. M.L., C.G.Y., and Y.P. performed the cellulose analyses. H.L.B., M.M., and C.N.S. carried out and analyzed the switchgrass field study. Y.-C.L., J.-Y.Z., and H.R. carried out molecular cloning and production of RNAi plasmid for switchgrass and rice. I.M.B. performed some of the cell wall analyses. A.L.B., Z.R.K., and P.R.L. performed switchgrass and rice transformations and propagated transgenic plants. S.P. and M.G.H. performed and analyzed the glycome profiling analysis. B.S.D. performed the stereomicrograph measurement of switchgrass biomass water uptake. S.S.M. and D.R. participated in growth, sampling and analysis of the plants. K.Y., O.A.T., M.R., A.D., and J.N. carried out the ethanol fermentation analyses. K.W. carried out molecular cloning and production of the RNAi plasmid for poplar. C.C. performed *Populus* transformation and propagated transgenic plants. X.Y. contributed bioinformatic information for construction of poplar gene constructs. L.T. performed molecular cloning and produced heterologous expression construct of *Arabidopsis* gene. R.W.S. conducted high-throughput pyrolysis molecular beam mass spectrometry (py-MBMS) lignin assays. E.L.G. and A.Z. coordinated analysis of samples through the BioEnergy Science Center (BESC) high-throughput MBMS and saccharification pipelines. G.B.T. performed high-throughput recalcitrance pipeline through BESC. S.R.D. guided overall high-throughput saccharification pipeline through BESC and provided data analysis. W.A.P. guided the switchgrass and rice transformation pipeline and provided data analysis. M.K.U. guided cloning and production of RNAi vectors for switchgrass and rice. J.R.M. and B.H.D. guided the overall ethanol assay and interpreted the ethanol data. M.F.D. developed and provided leadership for the MBMS pipeline through BESC. R.S.N. directed the BESC transformation pipeline and coordinated analyses through the pipeline. A.J.R. coordinated and analyzed cellulose research. D.M. conceived of the study, coordinated the research, and contributed to the interpretation of results, and drafting and finalizing of the manuscript. All authors read and approved the final manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Lynd, L.R. *et al.* How biotech can transform biofuels. *Nat. Biotechnol.* **26**, 169–172 (2008).
2. Tan, H.T., Corbin, K.R. & Fincher, G.B. Emerging technologies for the production of renewable liquid transport fuels from biomass sources enriched in plant cell walls. *Front. Plant Sci.* **7**, 1854 (2016).
3. Himmel, M.E. *et al.* Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804–807 (2007).
4. Baxter, H.L. *et al.* Two-year field analysis of reduced recalcitrance transgenic switchgrass. *Plant Biotechnol. J.* **12**, 914–924 (2014).

5. Bryan, A.C. *et al.* Knockdown of a laccase in *Populus deltoides* confers altered cell wall chemistry and increased sugar release. *Plant Biotechnol. J.* **14**, 2010–2020 (2016).
6. Shen, H. *et al.* A genomics approach to deciphering lignin biosynthesis in switchgrass. *Plant Cell* **25**, 4342–4361 (2013).
7. Van Acker, R. *et al.* Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. *Proc. Natl. Acad. Sci. USA* **111**, 845–850 (2014).
8. Biswal, A.K. *et al.* Downregulation of *GAUT12* in *Populus deltoides* by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. *Biotechnol. Biofuels* **8**, 41 (2015).
9. Peña, M.J. *et al.* Structural diversity of xylans in the cell walls of monocots. *Planta* **244**, 589–606 (2016).
10. Atmodjo, M.A., Hao, Z. & Mohnen, D. Evolving views of pectin biosynthesis. *Annu. Rev. Plant Biol.* **64**, 747–779 (2013).
11. Tan, L. *et al.* An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* **25**, 270–287 (2013).
12. Mortimer, J.C. *et al.* An unusual xylan in *Arabidopsis* primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14. *Plant J.* **83**, 413–426 (2015).
13. Pattathil, S. *et al.* A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol.* **153**, 514–525 (2010).
14. Mellerowicz, E.J. & Sundberg, B. Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties. *Curr. Opin. Plant Biol.* **11**, 293–300 (2008).
15. Carpita, N.C. & Gibeaut, D.M. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* **3**, 1–30 (1993).
16. Biswal, A.K. *et al.* Comparison of four glycosyl residue composition methods for effectiveness in detecting sugars from cell walls of dicot and grass tissues. *Biotechnol. Biofuels* **10**, 182 (2017).
17. Sterling, J.D. *et al.* Functional identification of an *Arabidopsis* pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proc. Natl. Acad. Sci. USA* **103**, 5236–5241 (2006).
18. Caffall, K.H., Pattathil, S., Phillips, S.E., Hahn, M.G. & Mohnen, D. *Arabidopsis thaliana* T-DNA mutants implicate *GAUT* genes in the biosynthesis of pectin and xylan in cell walls and seed testa. *Mol. Plant* **2**, 1000–1014 (2009).
19. Atmodjo, M.A. *et al.* Galacturonosyltransferase (GAUT)1 and GAUT7 are the core of a plant cell wall pectin biosynthetic homogalacturonan:galacturonosyltransferase complex. *Proc. Natl. Acad. Sci. USA* **108**, 20225–20230 (2011).
20. Zhang, J.Y. *et al.* Development of an integrated transcript sequence database and a gene expression atlas for gene discovery and analysis in switchgrass (*Panicum virgatum* L.). *Plant J.* **74**, 160–173 (2013).
21. Fu, C. *et al.* Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc. Natl. Acad. Sci. USA* **108**, 3803–3808 (2011).
22. Hao, Z. *et al.* Loss of *Arabidopsis GAUT12/IRX8* causes anther indehiscence and leads to reduced G lignin associated with altered matrix polysaccharide deposition. *Front. Plant Sci.* **5**, 357 (2014).
23. Pattathil, S., Avci, U., Miller, J.S. & Hahn, M.G. Immunological approaches to plant cell wall and biomass characterization: Glycome Profiling. *Methods Mol. Biol.* **908**, 61–72 (2012).
24. de Godoy, F. *et al.* Galacturonosyltransferase 4 silencing alters pectin composition and carbon partitioning in tomato. *J. Exp. Bot.* **64**, 2449–2466 (2013).
25. Chandra, R.P. & Saddler, J.N. Use of the Simons' staining technique to assess cellulose accessibility in pretreated substrates. *Ind. Biotechnol. (New Rochelle N.Y.)* **8**, 230–237 (2012).
26. Caffall, K.H. & Mohnen, D. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr. Res.* **344**, 1879–1900 (2009).
27. Levesque-Tremblay, G., Pelloux, J., Braybrook, S.A. & Müller, K. Tuning of pectin methylesterification: consequences for cell wall biomechanics and development. *Planta* **242**, 791–811 (2015).
28. Hocq, L., Pelloux, J. & Lefebvre, V. Connecting homogalacturonan-type pectin remodeling to acid growth. *Trends Plant Sci.* **22**, 20–29 (2017).
29. Thakur, B.R., Singh, R.K. & Handa, A.K. Chemistry and uses of pectin–a review. *Crit. Rev. Food Sci. Nutr.* **37**, 47–73 (1997).
30. Peaucelle, A., Braybrook, S. & Höfte, H. Cell wall mechanics and growth control in plants: the role of pectins revisited. *Front. Plant Sci.* **3**, 121 (2012).
31. Goldberg, R., Morvan, C., Jauneau, A. & Jarvis, M.C. in *Pectins and Pectinases* vol. 14. (eds. Visser, J. & Voragen, A.G.J.) 151–172 (Elsevier Science B.V., Amsterdam, 1996).
32. Mohnen, D. Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* **11**, 266–277 (2008).
33. Harholt, J., Suttangkakul, A. & Vibe Scheller, H. Biosynthesis of pectin. *Plant Physiol.* **153**, 384–395 (2010).
34. Wang, T. & Hong, M. Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *J. Exp. Bot.* **67**, 503–514 (2016).
35. Domozych, D.S. *et al.* Pectin metabolism and assembly in the cell wall of the charophyte green alga *Penium margaritaceum*. *Plant Physiol.* **165**, 105–118 (2014).
36. Voxeur, A. & Höfte, H. Cell wall integrity signaling in plants: "To grow or not to grow that's the question". *Glycobiology* **26**, 950–960 (2016).
37. Singh, B. *et al.* A specialized outer layer of the primary cell wall joins elongating cotton fibers into tissue-like bundles. *Plant Physiol.* **150**, 684–699 (2009).
38. O'Neill, M.A., Ishii, T., Albersheim, P. & Darvill, A.G. Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. *Annu. Rev. Plant Biol.* **55**, 109–139 (2004).
39. Darvill, A.G., McNeil, M. & Albersheim, P. Structure of plant cell walls: VIII. A new pectic polysaccharide. *Plant Physiol.* **62**, 418–422 (1978).
40. Liu, X.L. *et al. Male gametophyte defective 4* encodes a rhamnogalacturonan II xylosyltransferase and is important for growth of pollen tubes and roots in *Arabidopsis*. *Plant J.* **65**, 647–660 (2011).
41. O'Neill, M.A., Eberhard, S., Albersheim, P. & Darvill, A.G. Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth. *Science* **294**, 846–849 (2001).
42. Lamport, D.T. & Várnai, P. Periplasmic arabinogalactan glycoproteins act as a calcium capacitor that regulates plant growth and development. *New Phytol.* **197**, 58–64 (2013).
43. Knox, J.P., Linstead, P.J., King, J., Cooper, C. & Roberts, K. Pectin esterification is spatially regulated both within cell walls and between developing tissues of root apices. *Planta* **181**, 512–521 (1990).
44. Willats, W.G. *et al.* Analysis of pectic epitopes recognised by hybridoma and phage display monoclonal antibodies using defined oligosaccharides, polysaccharides, and enzymatic degradation. *Carbohydr. Res.* **327**, 309–320 (2000).
45. Tilman, D., Balzer, C., Hill, J. & Befort, B.L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. USA* **108**, 20260–20264 (2011).
46. Dumitrache, A. *et al.* Transgenic switchgrass (*Panicum virgatum* L.) targeted for reduced recalcitrance to bioconversion: a 2-year comparative analysis of field-grown lines modified for target gene or genetic element expression. *Plant Biotechnol. J.* **15**, 688–697 (2017).
47. Zhou, Y. *et al.* RNAi-directed down-regulation of RSV results in increased resistance in rice (*Oryza sativa* L.). *Biotechnol. Lett.* **34**, 965–972 (2012).
48. Noman, A., Aqeel, M. & He, S. CRISPR-Cas9: tool for qualitative and quantitative plant genome editing. *Front. Plant Sci.* **7**, 1740 (2016).
49. Biswal, A.K. *et al.* Aspen pectate lyase *Ptxt*PL1-27 mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. *Biotechnol. Biofuels* **7**, 11 (2014).
50. Lionetti, V. *et al.* Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. *Proc. Natl. Acad. Sci. USA* **107**, 616–621 (2010).
51. Tomassetti, S. *et al.* Controlled expression of pectic enzymes in *Arabidopsis* thaliana enhances biomass conversion without adverse effects on growth. *Phytochemistry* **112**, 221–230 (2015).
52. Voinnet, O., Pinto, Y.M. & Baulcombe, D.C. Suppression of gene silencing: a general strategy used by diverse DNA and RNA viruses of plants. *Proc. Natl. Acad. Sci. USA* **96**, 14147–14152 (1999).

ONLINE METHODS

Gene isolation, cloning, transformation and generation of switchgrass, rice and poplar *GAUT4*-KD transgenics. *Switchgrass*. TBLASTN was used to identify the homologous *GAUT4* (*Galacturonosyltransferase4*) gene sequences from the switchgrass EST databases[20] and the draft whole genome sequence (*Panicum virgatum* v1.1, Phytozome 11.0 DOE-JGI) at Phytozome using GAUT4 amino acid sequence from *Arabidopsis thaliana* (TAIR10) as a heterologous probe. Switchgrass contig AP13CTG20100 was identified as the best candidate for *PvGAUT4* based on cluster and multiple sequence alignment analyses, and was used to design primers for cloning: *PvGAUT4*-ORF-F (5′-CTGGCCTCTGGTCAATCAAT-3′) and *PvGAUT4*-ORF-R (5′-CCAGTAGCTTCGGAACTTGG-3′). The RNA interference (RNAi) cassettes were constructed by isolating a 443-bp *GAUT4* fragment from switchgrass (*Panicum virgatum* L.) cDNAs of the variety 'Alamo' clonal genotype ST1 using gene-specific primers for the *PvGAUT4* gene, followed by cloning into pCR8 entry vector for sequence confirmation. The entry clone was recombined into pANIC12A as described[53]. Embryogenic callus derived from immature inflorescences of SA7 was transformed with the expression vector pANIC12A-*PvGAUT4* (**Supplementary Fig. 1**) construct through biolistic transformation[54]. SA7 was developed at the University of Georgia for improved tissue culture response, and selected from a cross of two regenerable 'Alamo' genotypes, ST1 and Alamo2 (refs. 54,55). After antibiotic selection (60 mg/L hygromycin), the resistant calli were regenerated as described[54]. Transgenic lines were confirmed by PCR with vector primers flanking the *GAUT4* RNAi hairpin inserts: Zmubi1900-F (5′- TTTAGCCCTGCCTTCATACG-3′), guslink-R (5′- CCAAAGCCAGTAAAGTAGAACG-3′), guslink-F (5′-GTC GGTGAACAGGTATGGAAT-3′), nosT-R (5′-TGATAATCATCGCA AGACC-3′).

Rice. The putative rice *GAUT4* homolog (*OsGAUT4*, *LOC_Os08g23780*) (**Supplementary Fig. 1**) was identified by a similar process as described above for switchgrass. The *Arabidopsis thaliana* (TAIR10) GAUT4 amino acid sequence was used to perform tblastn/blastp against rice genome (*Oryza sativa* v7.0; Phytozome 11.0 DOE-JGI) followed by cluster and multiple sequence alignment analyses. The 443-bp *Pv*RNAi target has 88% identity to *LOC Os08g23780*, thus the switchgrass pANIC12A-*Pv*GAUT4 construct was used for transformation of rice (var. Taipei 309 from the USDA National Plant Germplasm System) as well, which was performed as described[56]. Rice transformation events were identified by real-time tracking of *in vivo pporRFP*[57] expression using fluorescence microscopy. Transgenic lines were confirmed by PCR with vector primers flanking the *GAUT4* RNAi hairpin inserts as above for switchgrass.

Poplar. P. deltoides genome information was not available at the time the RNAi construct was developed, thus the *P. trichocarpa* genome sequence was used to identify a suitable RNAi target sequence to transform *P. deltoides*. A 200-bp fragment comprising portions of the coding region and 3′-untranslated region of *Potri.016G001700* (**Supplementary Fig. 1**) was amplified via PCR from a *P. trichocarpa* cDNA library using gene-specific primers and cloned into pENTR/D-TOPO entry vector. The primers were: *PdGAUT4*-F (5′- C ACCCCCGGGATAGGCATACCCAAGTATCGAA-3′) and *PdGAUT4*-R (5′-TCTAGAAAAATGTTGGCACAGGATGTG-3′). Further cloning steps, as previously described[8], generated the binary transformation vector pAGW633 (GenBank accession number KX683862), which was used for transformation of eastern cottonwood (*P. deltoides*) clone WV94 as described[8]. The presence of the *GAUT4* transgene was verified by PCR across the *GAUT4*::spacer junction using primers pAGW633F (5′-TTTCAATCAAATGAAGAGCCAAT-3′) and pAGW633R (5′-TCGTGGTGTCATCCATTCAT-3′) on genomic DNA isolated from tissue culture shoots. More recently, a preliminary/early version of the genome sequence of *P. deltoides* became available (https://phytozome.jgi.doe.gov/pz/portal.html#). As part of this work, the *P. deltoides GAUT4* coding sequence was also cloned for heterologous expression (see below). This allowed comparison of the *P. trichocarpa GAUT4* homolog *Potri.016G001700.1* and *P. deltoides* counterpart *Podel.16G001400.1*, revealing 99.24% identity at the protein level and 99.29% identity at the nucleotide level.

Plant growth conditions in the greenhouse. *Switchgrass*. Both WT and transgenic switchgrass plants were grown in a greenhouse under a 16-h light/8-h dark cycle at 25–32 °C, depending on the season. The plants were grown in a mixture of two parts Fafard 3B (two 2.8 cubic feet (79.3 l) bags, GroSouth Inc., Atlanta, GA), one part sand (one 2.8 cubic feet (79.3 l) of River Bottom Sand, Redland Sand, Watkinsville, GA) and 118 mL of Osmocote Plus granular fertilizer (18-9-12 minors, 8–9 month release). After planting, plants were fertilized once a week with 440 mg/L of the Peat Lite Special 20-10-20 (nitrogen-phosphorus-potassium; GroSouth Inc, Atlanta, GA). Fifteen tillers were propagated from both control and transgenic lines to produce five clonal replicates for measurement of growth parameters in 3.8-l (1 gallon) and 19-l (5 gallon) pots for 4 months[58].

Rice. Rice seeds were maintained and cultured as previously described[56]. After transfer to soil, control and transgenic rice lines were grown in a greenhouse under a 16-h light/8 h dark cycle at 25–32 °C, depending on the season. The soil composition for rice was the same as that used for switchgrass. Fertilizer applied at planting was 1.2 mL (1/4 teaspoon) Sprint 330 Iron Chelate and 3.75 g Jack's Peat Lite Special 20-10-20 (nitrogen-phosphorus-potassium) per 2 l water. After planting, the plants were fertilized once a week as for switchgrass.

Poplar. Populus deltoides plants were grown in a greenhouse as previously described[8]. Measurement of plant growth was carried out on 25 WT plants and 12–15 plants for each vector control and *PdGAUT4*-KD line. The relative water content (RWC) and dry matter measurements were performed as previously described[8].

Switchgrass field experiment. A field plot (14.5 m × 23.6 m) on a University of Tennessee-Knoxville farm was divided into three subplots, one of which was used for the field trial of non-transgenic WT (the 'Alamo'-derived parental clone SA7 and an additional clone SA37) and the three *PvGAUT4*-KD lines (2A, 2B, and 4A) (**Supplementary Fig. 14**). The subplot was divided into three replicate-plots for each genotype that were arranged in a completely randomized design within the subplot. Four vegetatively propagated clones of each line were planted 76 cm apart within each replicate-plot, and the replicate-plots were spaced 152 cm apart. A border of 'Alamo'-derived ST1-genotype plants was included to control for shading effects. The plants were transplanted to the field in July of 2013 and grown over three consecutive growing seasons (2013, 2014, and 2015). The field was maintained as described previously[59]. Plants began producing panicles in late June or early July each season. As required by the USDA APHIS BRS release-into-the-environment permit, panicles were removed from all tillers once reaching the first reproductive (R1) developmental stage[58] by cutting the tiller from below the top node.

Measurements of tiller height, plant width, tiller number, and dry weight biomass yield were carried out after plants reached full senescence at the end of each growing season, as described previously[4]. Biomass from the four clonal plants within each replicate were pooled to represent the total dry weight yield from each replicate. Incidence of *Puccinia emaculata* infection, which causes switchgrass rust, was evaluated weekly between late July and late August during the second (2014) and third (2015) growing seasons as described previously[59]. Briefly, randomly selected tillers from two clonal plants within each replicate subplot were tagged and visually assessed for rust severity using the following scale: 0 = 0%, 1 = ≤5%, 2 = ≤10%, 3 = ≤25%, 4 = ≤40%, 5 = ≤55%, 6 = ≤70%, and 7 = ≤100% of leaf area coverage with uredia. In late August of 2014 and 2015, all plants were treated with fungicides 'Quilt' (Syngenta Canada Inc., Guelph, Ontario) at a rate of 0.21 mL/m^2 and 'Heritage' (Syngenta Crop Protection, Greensboro, NC) at a rate of 20 mL/m^2.

Sample preparation for molecular and biochemical analyses. Switchgrass tiller samples were collected and immediately frozen in liquid nitrogen for transcript analysis and enzymatic assay. For field-grown switchgrass, samples for transcript analysis were collected from each replicate-plot at approximately mid-growing season of each year. For saccharification, fermentation, analytical pyrolysis, and cell wall analyses of greenhouse-grown plants, samples were harvested as follows: switchgrass whole tiller at the R1 stage from 3-month-old plants[60], rice whole aerial biomass from 3-month-old plants, and poplar debarked bottom (6 cm from the ground) stem from 9-month-old plants. The plant materials were air-dried at room temperature for 3–4 weeks (depending on the season) and milled to particle size of 20 mesh (0.85 mm) using a Wiley Mini-Mill (model number: 3383L10, Thomas Scientific). For saccharification, fermentation, and analytical pyrolysis analyses of field-grown switchgrass,

doi:10.1038/nbt.4067

whole aboveground senesced biomass was harvested in December of each year, dried for ~168 h at 43 °C, and subsamples from each replicate-plot taken and milled through a 1-mm screen.

RNA isolation and Quantitative Real Time PCR. RNA isolation and transcript analyses were performed as previously described[8]. Briefly, shoot tips from switchgrass, the first leaf from rice, and xylem samples scraped from debarked frozen poplar stems were used for transcript analysis. All tissues were ground to a fine powder in liquid nitrogen. Total RNA from switchgrass and rice tissue was isolated using an RNeasy Plant Mini Kit (Qiagen, Valencia, CA), and from *Populus* using a CTAB (hexadecyl-trimethylammonium bromide) method[8] followed by removal of genomic DNA with DNase (Qiagen, Valencia, CA). First-strand cDNA was synthesized from 1 µg total RNA using the iScript cDNA Synthesis kit (BioRad, Hercules, CA). Quantitative RT-PCR reactions were performed in triplicate using the iQ SYBR Green Supermix (Bio-Rad, Hercules, CA) and the following primers: switchgrass *PvGAUT4*-qRT-F (5′-CCGGAGAAATCACGAATCA-3′) and *PvGAUT4*-qRT-R (5′-CCAGTTCATCCGAGAAGAC-3′), rice *OsGAUT4*-qRT-F (5′-CTCTGCCTCCAGGTTTGATTAC-3′) and *OsGAUT4*-qRT-R (5′-GCGGCACGTTCAATATCTCT-3′), poplar *PdGAUT4*-qRT-F (5′-CCACCTGGCCTCATTACTTT-3′) and *PdGAUT4*-qRT-R (5′-GCTGCTCGTTCAATTTCTCTTT-3′). For reference genes, the following primers were used: switchgrass *CYP5* (*Cyclophilin 5*, Pavir.J00040, GenBank Accession number FE633090) primers *PvCYP5*-qRT-F (5′-CACTACAAGGGAAGCACATTCCA-3′) and *PvCYP5*-qRT-R (5′-TTCACCACCCCTTCCATCAC-3′), rice *Actin* primers *OsActin*-qRT-F (5′-CTTCATAGGAATGGAAGCTGCGGGTA-3′) and *OsActin*-qRT-R (5′-CGACCACCTTGATCTTCATGCTGCTA-3′), and poplar *18S rRNA* primers *Pd18S*-qRT-F (5′-AATTGTTGGTCTTCAACGAGGAA-3′) and *Pd18S*-qRT-R (5′-AAAGGGCAGGGACGTAGTCAA-3′). The method used for analysis of relative transcript expression was as previously described[8,61].

Determination of saccharification efficiency. Dried, milled switchgrass, rice, and poplar was used for analysis of sugar release efficiency via NREL high-throughput sugar release assays[8,62,63]. Biomass was extracted with glucoamylase (Spirizyme Ultra - 0.25% (v/v)) and alpha-amylase (Liquozyme SC DS - 1.5% (v/v)) in 0.1 M sodium acetate pH 5.0 at 55 °C for 24 h to remove starch (16 mL enzyme solution per 1 g biomass) followed by an additional 24 h Soxhlet extraction with ethanol[64]. Water (250 µL) was added to 5 mg of overnight dried samples and hot water pretreatment was carried out at 180 °C for 17.5 min as previously described[8]. The hot water pretreatment conditions were chosen to be suboptimal for maximum glucose release to enable a clear distinction between samples with high and low recalcitrance. Buffer–enzyme stock (40 µL of 8% (v/v) Novozymes CTec2 in 1.0 M sodium citrate buffer, pH 5.0) was added to cooled pretreated sample. After 70 h of incubation at 50 °C, the diluted saccharified hydrolysate was analyzed using Megazyme's GOPOD (glucose oxidase/peroxidase) and XDH (xylose dehydrogenase) assays, with mixed glucose/xylose standards as previously described[8,64]. The methods were used for the data shown in **Figure 1a–i** and **Supplementary Figure 2a–c**.

Pretreatment and fermentation of switchgrass biomass assayed for ethanol yield. Prior to fermentation, the switchgrass ground biomass was subjected to a hot water pretreatment as described earlier[21,65,66]. Briefly, ground biomass was soaked overnight (~18 h) in Milli-Q water at a ratio of 9 mL water per gram of dry biomass. After centrifugation, the biomass was loaded at a ratio of 2.5 g dry biomass per tube into 10 cm × 1 cm hastelloy steel tubular pretreatment reactors (Industrial Alloys Plus, Inc., Utica, KY). The reactors were heated in boiling water for 2 min before heating in the sand bath (Omega FSB1, Techne Co.) at 180 °C for 25 min[21,67], followed by quenching in an ice bath. The solid residual biomass was removed from the reactors, washed with Milli-Q water (100 mL/g dry biomass), and stored at −20 °C until fermentation.

Ethanol fermentation for greenhouse-grown plants was carried out by a separate hydrolysis and fermentation (SHF) method using *Saccharomyces cerevisiae* D5A, ATCC 200062 (ref. 68). The hot-water-pretreated control and transgenic switchgrass biomass samples were subjected to 5-d hydrolysis with commercial blends of Ctec2, Htec2, and Novozyme 188 (Novozymes), pH 4.5 at 50 °C. The weight loss was used to monitor the progress of the fermentation with *S. cerevisiae* D5A at 35 °C, 150 r.p.m., pH 4.5 as described previously[65]. This method was used for data shown in **Figure 1j–l**. For field trial biomass, ethanol fermentation was carried out as described[46] (data shown in **Fig. 1p**).

Analysis of lignin content and composition by analytical pyrolysis. Approximately 4 mg of ground switchgrass and poplar biomass was analyzed in duplicate by analytical pyrolysis for the quantification of lignin content and S/G lignin monomer ratio using the NREL high-throughput py-MBMS method[8,69,70]. Total lignin content was determined by summing the intensities of the major lignin precursors as previously described[8]. S/G ratios were determined by dividing the sum of syringyl peaks by the sum of guaiacyl peaks. Since some of the individual peaks are associated with both S and G precursors, several lignin peaks were omitted from both syringyl or guaiacyl summations for the S/G ratio[69].

Heterologous expression of recombinant GAUT4 in *N. benthamiana* leaves. *Cloning and transformation*. The coding regions of *Arabidopsis*, *P. deltoides*, and switchgrass *GAUT4* were cloned by PCR from cDNA preparations made from RNA isolated from suspension cultured cells, leaves, and first internodes, respectively. The primers, which simultaneously added 6xHis-tag and restriction sites for subcloning, were as follows (6x-His-tag and restriction sites are underlined, coding region sequences are italicized): *Arabidopsis AtGAUT4*-F (5′-CAATAGGATCC*ATGATGGTGAAGCTTCGCAATC*-3′) and *AtGAUT4*-R (5′-GAATTCGAGCTC*CCTAGTGGTGGTGATGGTGGTGAGGATTGATGTTGCATTC*-3′), poplar *PdGAUT4*–F (5′-CCCGGG*ATGAATGCAGTTTCTTTCTCTCATACATCC*-3′) and *PdGAUT4*-R (5′-GAGCTC*TTA*ATGATGATGATGATGATG*TGGATTGATATTGCACTCTC*-3′), switchgrass *PvGAUT4*-F (5′-CCCGGG*ATGGTGAGGCCGAGGAGGATCCTCC*-3′) and *PvGAUT4*-R (5′-GAGCTC*TTA*ATGATGATGATGATGATG*TGGATTGATGTTGCATTCCCGCAGG*-3′). After sequence verification in cloning vectors, the coding regions (GenBank accession numbers: *PdGAUT4* - KX369395, *PvGAUT4* - KX369394) were subcloned into the pBI121 vector[71] to replace the β-glucuronidase gene via the BamHI-SacI sites for *Arabidopsis GAUT4* and via the SmaI-SacI sites for the poplar and switchgrass *GAUT4*. The resulting constructs were introduced into *Agrobacterium tumefaciens* strain GV3010::pMP90 by electroporation, and the resulting *Agrobacterium* strains subsequently used for transient expression in *Nicotiana benthamiana* leaves using previously described methods[19,72] with modifications. Briefly, *N. benthamiana* plants were grown for 8 weeks in a growth chamber at 24 °C, 14 h day and 20 °C, 10 h night. Overnight *Agrobacterium* cultures bearing a *Tomato bushy stunt virus* p19 construct (a gift from Henrik Scheller, JBEI) and the *GAUT4* constructs were harvested, washed once, and resuspended in buffer (100 µm acetosyringone, 10 mM $MgCl_2$, 10 mM MES, pH 5.6) with optical densities (OD_{600}) adjusted to 1.0 for p19 and to 0.1 or 1.0 for the *GAUT4* constructs. The *Agrobacterium* cell suspension mixtures were incubated for 2 h at room temperature before infiltration into *N. benthamiana* leaves. The infiltrated leaves were harvested 4 d later and either directly used for microsome preparation, or flash-frozen in liquid nitrogen and stored at −80 °C until use. Transcript expression of the transgenes was assessed by quantitative real time PCR, using primers *AtGAUT4*-qRT-F (5′-GGATGAATGGAAGAGGCAAAAC-3′) and *AtGAUT4*-qRT-R (5′-AATGTGATTAGACCAGGCGG-3′) for *Arabidopsis*[10] and as described above for poplar and switchgrass.

Preparation and solubilization of microsomes and affinity purification of recombinant GAUT4. Preparation of microsomal membranes was as described[22] with modification. All steps were done on ice or at 4 °C. Harvested tobacco leaves were ground in pre-chilled homogenizing buffer (50 mM HEPES pH 7.3, 0.25 mM $MnCl_2$, 0.1% (v/v) β-mercaptoethanol, 25 mM KCl, 50% (v/v) glycerol, and EDTA-free protease inhibitor cocktail (Roche Diagnostics, Indianapolis, IN); 2 mL buffer/g tissue) using a mortar and pestle until pureed. The homogenate was filtered through three layers of Miracloth (Millipore, Billerica, MA) and centrifuged at 4,000*g*, 4 °C for 30 min. The supernatant was centrifuged at 110,000*g* for 1 h at 4 °C, and the resulting pellet resuspended in storage buffer (50 mM HEPES pH 7.3, 0.25 mM $MnCl_2$, 25 mM KCl, 25% (v/v) glycerol, and EDTA-free protease inhibitor cocktail; ~30 µL buffer/g fresh weight) using a glass homogenizer. The microsomes were kept at −80 °C until use. Protein concentration was determined using the Bradford assay (BioRad, Hercules, CA) and expression analyzed by western blotting using a mouse anti-His-tag antibody (BioRad, Raleigh, NC).

Microsomes were detergent-solubilized as previously described[73], except that CHAPS was used instead of Triton X-100. CHAPS was applied at 40 mM during solubilization[74] and subsequently diluted to 5 mM by adding seven volumes of buffer without detergent (50 mM HEPES, pH 7.3, 0.25 mM $MnCl_2$, 2 mM EDTA, 25% (v/v) glycerol) to the collected supernatant after ultracentrifugation. Solubilized recombinant GAUT4 proteins were affinity-purified via the 6xHis-tag using Talon Metal Affinity Resin (Clontech Laboratories Inc., Mountain View, CA) following batch/gravity-flow column purification essentially as described by the manufacturer with a few modifications. Prior to affinity purification, EDTA was removed from the solubilized protein solutions by passage through PD-10 columns (GE Healthcare Bio-Sciences, Marlborough, MA). The buffers for resin equilibration/wash and elution were prepared as per manufacturer's instruction but augmented with 0.25 mM $MnCl_2$, 5 mM CHAPS, and 20% (v/v) glycerol. Eluted recombinant proteins were concentrated and buffer-exchanged using Amicon Ultra-2 centrifugal filter units with an Ultracel-30 membrane (EMD Millipore, Billerica, MA) into buffer (50 mM HEPES, pH 7.3, 0.25 mM $MnCl_2$, 2 mM EDTA, 5 mM CHAPS, and 25% (v/v) glycerol). Protein concentration was determined by the Bradford assay. Purified proteins were kept at −80 °C until use.

HG:GalAT assay. HG:GalAT assays were carried out in 30-µL reactions containing 50 mM HEPES (pH 7.8), 0.2 M sucrose, 0.05% (w/v) BSA, 25 mM KCl, 1.25 mM $MnCl_2$, 1 mM oligogalacturonides (OGA) DP 7-23 (ref. 74), and 5 µM UDP-[^{14}C]GalA (specific activity 249 mCi/mmol; 1 Ci = 37 GBq). UDP-[^{14}C]GalA was synthesized from UDP-[^{14}C]GlcA (PerkinElmer, Boston, MA) as previously described[19]. Enzyme was added at either 100 µg protein/reaction for microsomes (with addition of 0.1% Triton X-100) or 1 µg protein/reaction for purified recombinant GAUT4. Following incubation for 3 h (unless otherwise noted), reactions were terminated by adding 5 µL of 400 mM NaOH and subsequently processed using the filter assay method[75]. For product characterization, reaction products were subjected to digestion by endopolygalacturonase (EPG, E.C. 3.2.1.15, 9.5 mg/mL, 2,256 units/mg; 1 unit = 1 µmol of reducing sugar produced per minute) from *Aspergillus niger* as previously described[17].

Preparation of alcohol insoluble residues (AIR) and wall extracts, and analysis of glycosyl residue composition and linkage. AIR and fractionated cell walls from ground biomass were prepared as previously described[8]. The sequential fractionation of AIR using the increasingly harsh chemicals was expected to yield the following polymer enrichments: ammonium oxalate and sodium carbonate to isolate HG-enriched polymers, 1M and 4M KOH to isolate pectin-containing fractions enriched for hemicelluloses, chlorite and 4M KOH-post chlorite to identify polymers potentially associated with lignin, leaving a final insoluble pellet containing cellulose and other polymers highly resistant to extraction and potentially involved in foundational cell wall architecture.

Glycosyl residue composition of the AIR (~2 mg) and of each cell wall extract (100–300 µg) was determined by GC-MS of trimethylsilyl (TMS) derivatives[8,76]. For glycosyl residue linkage analysis, approximately 1.3 mg of each wall extract from switchgrass and poplar WT and the transgenics were subjected to the procedure as described[77] with a slight modification. Briefly, wall extracts were suspended in 200 µL dimethyl sulfoxide and stirred for 1 day at 25 °C. The permethylation was performed using potassium dimsyl anion and iodomethane, and permethylated uronic acids were reduced using lithium borodeuteride. Permethylation was by two rounds of treatment with sodium hydroxide for 15 min and methyl iodide for 45 min. The permethylated material was hydrolyzed using 2 M trifluoroacetic acid (TFA) for 2 h in sealed tubes at 121 °C, reduced with $NaBD_4$, and acetylated using acetic anhydride/TFA. The resulting partially methylated alditol acetates (PMAAs) were analyzed by gas chromatography-mass spectrometry (GC-MS) using an Agilent 7890A GC interfaced to a 5975C MSD (mass selective detector, electron impact ionization mode). Separation was performed on a Supelco SP-2331 bonded phase fused silica capillary column (30 m × 0.25 mm ID). Linkage analysis was carried out by the CCRC Analytical Services (University of Georgia, Athens, GA) (http://www.ccrc.uga.edu/services).

Glycosyl residue composition analysis for RGII-specific sugars was performed by combined GC/MS of the alditol acetates (AAs) as described[78]. Briefly, lyophilized samples (1 mg of wall fractions with 20 µg myo-inositol added as an internal standard) were hydrolyzed using 2M TFA for 2 h at 121 °C, reduced using NaBD4, and acetylated using acetic anhydride/TFA. The resulting alditol acetates were analyzed by GC-MS as described above for the glycosyl linkage analysis, with comparison to standards 2-*O*-Me-α-Xyl and 2-*O*-Me-α-L-Fuc (a kind gift from M. O'Neill, Complex Carbohydrate Research Center, University of Georgia).

Glycome profiling and ELISA using 2F4 monoclonal antibody. Cell wall extracts from switchgrass and poplar WT and *GAUT4*-KD lines were subjected to glycome profiling analysis as previously described[8,23]. Glycome profiling is an enzyme-linked immunosorbent assay (ELISA)-based method that makes use of over 150 monoclonal antibodies (mAb) directed against cell wall carbohydrates[13,23] (http://www.wallmabdb.net) to screen wall extracts. Additionally, ELISA was also carried out on selected, pectin-rich extracts using the 2F4 mAb (a kind gift from P. Knox, PlantProbes, www.plantprobes.net), which recognizes largely demethylesterified and Ca^{2+}-cross-linked HG[43,44]. The 2F4 primary antibody was used at 1:10 dilutions in combination with a goat-anti-mouse secondary antibody conjugated with Alexa Fluor 488 (Invitrogen).

Microscopy. Light microscopy of stem cross-sections was carried out as previously described[8], using poplar 10th stem internode from the apex as the starting material. Sample preparation and light microscopy of individual xylem fiber and vessel cells were performed as previously described[8]. At least 100 fibers and vessel cells from each individual plant (three independent plants per genotype) were observed using a Nikon Eclipse 80i fluorescence microscope (Nikon, Melville, NY). Images were captured using a Nikon DS-Ri2 camera (Nikon, Melville, NY) and analyzed with NIS-Elements Basic Research software to calculate the length and width of fiber and vessel cells.

For immunogold labeling and transmission electron microscopy of switchgrass stem sections, stem from R1 stage tillers was fixed, embedded in resin, and cut into sections (60 nm thick) as described[8]. For labeling with JIM5 and JIM7 antibodies, sections were hydrated in potassium phosphate buffered saline (KPBS, 100 mM KPB, 500 mM NaCl, pH 7.2) for 10 min, and the non-specific antibody binding sites blocked by incubation with 3% (w/v) non-fat dry skimmed milk in KPBS buffer (10–15 µl) for 1 h. Sections were subsequently incubated with undiluted primary antibodies for 1 h, washed three times (5 min each) with KPBS buffer, and incubated with the secondary antibody at 1:20 dilution for 1 h. For labeling with the 2F4 antibody, sections were processed using the same procedure except for the following: blocking buffer was 5% (w/v) non-fat dry skim milk in Tris-HCl 20 mM buffer, pH 8.2, with 0.5 mM $CaCl_2$ and 150 mM NaCl, incubation with the primary antibody was for 1 h at RT followed by overnight at 4 °C, and the secondary antibody was used at 1:10 dilution. The secondary antibodies were 12 nm-gold particle-conjugated goat anti-rat (for JIM5 and JIM7) and anti-mouse (for 2F4) IgG (H+L) (Jackson ImmunoResearch Laboratories Inc., West Grove, PA). Samples were viewed using a JEOL JEM 1011 transmission electron microscope (JEOL USA Inc., Peabody, MA) operating at 80 kV and documented using Image Capture Engine V700 software (Advanced Microscopy Techniques, Woburn, MA). TEM work was conducted in the Georgia Electron Microscope Lab at the University of Georgia, Athens, GA, USA.

For scanning electron microscopy of pretreated biomass, the switchgrass WT and *PvGAUT4*-KD biomass was subjected to the hot water pretreatment as described[65]. Briefly, 1 g of biomass (20 mesh, 0.85 mm) was soaked in 9 mL of nanopure water overnight and then incubated at 180 °C for 15 min. The pretreated samples were air-dried, mounted on aluminum specimen stubs using carbon adhesive tabs, coated with gold palladium particles (10 nm thick) using a Leica EM ACE200 sputter coater, and imaged using a JEOL JSM-6010LV scanning electron microscope with InTouchScope software (JEOL USA Inc., Peabody, MA).

Analyses of cellulose content, crystallinity, and degree of polymerization. *Preparation of extractives-free switchgrass biomass samples and measurement of cellulose content.* Extractives-free switchgrass biomass was prepared according to previously published procedures[79,80], by subjecting the ground biomass samples to air-drying followed by Soxhlet-extraction using an ethanol:toluene mixture (1:2, v:v) for 24 h. The cellulose content of switchgrass biomass was measured according to a modification of the NREL protocol http://www.nrel.gov/docs/gen/fy13/42618.pdf for chemical composition analysis[81]. Extractives-free

doi:10.1038/nbt.4067

switchgrass ground biomass samples were subjected to a two-step hydrolysis: (1) ~150 mg samples were hydrolyzed with 1.5 mL 72% H_2SO_4 for 1 h at 30 °C; (2) the hydrolysates were diluted to 4% H_2SO_4 with deionized (DI) water and a second hydrolysis carried out at 121 °C for another 1 h in an autoclave. The hydrolysate was cooled and filtered through a porcelain crucible. Glucose units in the filtrate were quantified by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) using Dionex ICS-3000 (Dionex Corp., USA)[82]. The amount of glucose was taken as cellulose content and reported as weight percentages of extractives-free biomass.

Isolation. Cellulose isolation was according to a previously published procedure[79,80]. The extractives-free switchgrass ground biomass was delignified by peracetic acid at 5 g loading per g biomass. The solution was adjusted to 5% with DI water and holopulping was conducted at room temperature for 24 h with stirring. The solid residue, designated as holocellulose, was washed excessively with DI water and air-dried at room temperature for 24 h. A sub-portion of the air-dried holocellulose (150 mg) was treated with 2.5 M hydrochloric acid (HCl) for 2 h to remove hemicellulose. The isolated cellulose was collected by filtration, rinsed with an excess of DI water, stored in sealed containers to prevent moisture loss, and used to analyze cellulose crystallinity by solid-state NMR. Another sub-portion of holocellulose (100 mg) was consecutively extracted at 25 °C with 17.5% (wt/v) NaOH solution (5 mL) for 2 h, followed by 8.75% NaOH solution (10 mL) for an additional 2 h. The alkaline slurry was filtered, and the solid residue rinsed with 5 mL of 1% acetic acid, washed with an excess of DI water, and air-dried. This solid residue, namely α-cellulose, was analyzed using gel permeation chromatography (GPC) to determine the cellulose degree of polymerization (DP).

Solid-state NMR analysis. Solid-state NMR analysis to determine cellulose crystallinity was performed as previously described[80]. The NMR samples were prepared by packing the moisturized cellulose into 4-mm cylindrical Zirconia MAS rotors. Cross-polarization magic angle spinning (CP/MAS) NMR analysis of cellulose was carried out on a Bruker Advance-400 spectrometer operating at 100.59 MHz for ^{13}C in a Bruker double-resonance MAS probe head at spinning speeds of 8 kHz. CP/MAS experiments used a 5 ms (90°) proton pulse, 1.5 ms contact pulse, 4 s recycle delay and 4,000 scans. The cellulose crystallinity index (CrI) was determined from the areas of the crystalline and amorphous C_4 signals using the following formula:

$$\mathrm{CrI}=\frac{A_{86-92\ \mathrm{ppm}}}{A_{86-92\,\mathrm{ppm}}+A_{79-86\,\mathrm{ppm}}}\times 100$$

Gel permeation chromatographic (GPC) analysis. The weight-average molecular weight (M_w) and number-average molecular weight (M_n) of cellulose were measured by GPC after tricarbanilation[79]. The isolated α-cellulose was derivatized with phenyl isocyanate in an anhydrous pyridine system before GPC analysis. Size-exclusion separation was performed on an Agilent 1200 HPLC system (Agilent Technologies, Inc, Santa Clara, CA) equipped with Waters Styragel columns (HR1, HR4, and HR5; Waters Corporation, Milford, MA). Number-average degree of polymerization (DP_n) and weight-average degree of polymerization (DP_w) of cellulose were calculated by dividing M_n and M_w by the molecular weight of the tricarbanilated cellulose repeat unit (519 g/mol). Polydispersity index (PDI) of cellulose was calculated as the ratio of DP_w over DP_n.

Analysis of cellulose accessibility by Simons' stain. The cellulose accessibility of switchgrass biomass was measured by a down-scaled Simons' stain procedure using Direct Blue (DB) 1 (Pontamine Fast Sky Blue 6BX) and Direct Orange (DO) 15 (Pontamine Fast Orange 6RN) dyes as previously described[25,83]. The extractives-free switchgrass biomass (10 mg oven-dry basis) was aliquoted into individual 2-mL microcentrifuge tubes and incubated with increasing amounts of DO and DB dyes (0.10, 0.20, 0.30, 0.40, 0.60, 0.80, and 1.00 mg/mL) in phosphate-buffered saline solution (pH 6.0, 30 mM PO_4^{3-}, 140 µM NaCl in a total volume of 1 mL). The mixtures were incubated at 70 °C for 6 h with shaking at 200 r.p.m. to allow the dye adsorption to reach equilibrium. Absorbance of the supernatant after centrifugation was measured using a Lambda 35 UV-Vis spectrophotometer at 410 nm and 599 nm, which represent the wavelengths of maximum absorbance for DO and DB, respectively. The amount of each dye adsorbed by the biomass was determined from the difference between the concentration of initial added dye and concentration of dye in the supernatant. The maximum amount of DO or DB adsorbed was calculated by nonlinear fitting to the Langmuir adsorption model. The surface concentration of adsorbed dye (Γ) was given by the equation:

$$\Gamma=\frac{\Gamma_{\max}\times K\times C}{1+K\times C}=\frac{\Gamma_{\max}\times C}{1/K+C}$$

where $\Gamma_{\max}$ is the surface concentration of dyes at full coverage (mg/g substrate), K is the Langmuir constant (mL/mg), and C is the free dyes concentration in the bulk solution (mg/mL). Distribution coefficient R (mL/g) is calculated as a product of $\Gamma_{\max}$ and K.

Biomass hydration, imaging, and image analysis. Switchgrass stems (R1 stage) from greenhouse-grown plants were fully hydrated by soaking in water under vacuum (21 psi) overnight then hand-sectioned with a razor blade. These saturated stem sections were immediately imaged on a Nikon SMZ1500 stereomicroscope and captured with a Nikon DS-Fi1 CCD camera operated by a Nikon Digital Sight system (Nikon, Melville, NY). The stem sections were then dried overnight and then re-imaged. Fiji (ImageJ[84]) was used to measure section edge dimensions on calibrated images. Two stem sections from two individual stems were measured, and an average of three separate measurements of stem section length and a Feret diameter measurement from a best-fit ellipse were used to calculate section volume change. Since most of the stem particles were hollow, the volume of the hollow center was subtracted from the total particle volume to derive the volume of only the plant tissue.

Measurement of calcium and boron content and calcium binding assay. The total calcium and boron content of AIR and selected wall fractions was determined by inductively coupled plasma–mass spectrometry (ICP-MS) analysis at the Center for Applied Isotope Studies, University of Georgia, Athens, GA. The amount of Ca^{2+} that forms cross-links with HG was estimated by a calcium binding assay as described[42] with a few modifications. Briefly, Ca^{2+}-free ammonium oxalate and sodium carbonate wall extracts from switchgrass *PvGAUT4*-KD and WT biomass were prepared by resuspending the extracts (1 mg/mL) in ddH_2O followed by titration with HCl to pH 2.0. The resulting mixture was ultrafiltrated over a 0.45 µm membrane filter (Nanosep MF, Pall Life Sciences) with two washes of ddH_2O to remove excess HCl and released Ca^{2+}. The material retained on the membrane filter was incubated with 1 mL of 2% $CaCl_2$ for 15 min to equilibrate it with Ca^{2+}, and the final mixture transferred into an ultracentrifugation device (Amicon - 30 kDa molecular weight cut-off) for filtration at 4,000*g* for 35 min followed by two washes with ddH_2O to remove excess Ca^{2+}. The ultrafiltrate was analyzed by ICP-MS as above to quantify the amount of Ca^{2+} not bound by the wall material. The Ca^{2+} binding capacity of each wall extract was determined by subtracting the Ca^{2+} content of the ultrafiltrate from the starting amount of Ca^{2+} used for initial equilibration (i.e., 1 mL of 2% $CaCl_2$) of the starting Ca^{2+}-free wall extracts.

Statistical analysis. One-way ANOVAs were used to test for statistical differences among means among controls and transgenic lines. When significant differences were detected at the $P < 0.05$ level, multiple comparison analysis was performed using Tukey's multiple comparison (Statistica 5.0 software) for greenhouse data sets and by Fisher's least significant difference method (SAS version 9.4, SAS Institute Inc., Cary, NC) for field trial data sets.

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Data availability. Sequence data that support the findings of this study have been deposited in GenBank with the accession codes KX369394, KX369395, and KX683862. All the other data generated or analyzed during this study are included in this published article and its supplementary information files.

53. Mann, D.G. *et al.* Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum* L.) and other monocot species. *Plant Biotechnol. J.* **10**, 226–236 (2012).
54. King, Z.R., Bray, A.L., Lafayette, P.R. & Parrott, W.A. Biolistic transformation of elite genotypes of switchgrass (*Panicum virgatum* L.). *Plant Cell Rep.* **33**, 313–322 (2014).

55. Wuddineh, W.A. *et al.* Identification and molecular characterization of the switchgrass AP2/ERF transcription factor superfamily, and overexpression of *PvERF001* for improvement of biomass characteristics for biofuel. *Front. Bioeng. Biotechnol.* **3**, 101 (2015).
56. Mann, D.G. *et al.* Switchgrass (*Panicum virgatum* L.) polyubiquitin gene (*PvUbi1* and *PvUbi2*) promoters for use in plant transformation. *BMC Biotechnol.* **11**, 74 (2011).
57. Mann, D.G. *et al.* Very bright orange fluorescent plants: endoplasmic reticulum targeting of orange fluorescent proteins as visual reporters in transgenic plants. *BMC Biotechnol.* **12**, 17 (2012).
58. Hardin, C.F. *et al.* Standardization of switchgrass sample collection for cell wall and biomass trait analysis. *Bioenerg Res* **6**, 755–762 (2013).
59. Baxter, H.L. *et al.* Transgenic miR156 switchgrass in the field: growth, recalcitrance and rust susceptibility. *Plant Biotechnol. J.* 1**6**, 39–49 (2017).
60. Moore, K.J. *et al.* Describing and quantifying growth stages of perennial forage grasses. *Agron. J.* **83**, 1073–1077 (1991).
61. Gimeno, J., Eattock, N., Van Deynze, A. & Blumwald, E. Selection and validation of reference genes for gene expression analysis in switchgrass (*Panicum virgatum*) using quantitative real-time RT-PCR. *PLoS One* **9**, e91474 (2014).
62. Selig, M.J. *et al.* Lignocellulose recalcitrance screening by integrated high-throughput hydrothermal pretreatment and enzymatic saccharification. *Ind. Biotechnol. (New Rochelle N.Y.)* **6**, 104–111 (2010).
63. Decker, S.R., Brunecky, R., Tucker, M.P., Himmel, M.E. & Selig, M.J. High-throughput screening techniques for biomass conversion. *Bioenerg. Res.* **2**, 179–192 (2009).
64. Decker, S.R. *et al.* in *Biomass Conversion. Methods in Molecular Biology*, vol. 908 (ed. Himmel, M.E.) 181–195 (Humana Press, New York, 2012).
65. Yee, K.L. *et al.* Evaluation of the bioconversion of genetically modified switchgrass using simultaneous saccharification and fermentation and a consolidated bioprocessing approach. *Biotechnol. Biofuels* **5**, 81 (2012).
66. Studer, M.H., DeMartini, J.D., Brethauer, S., McKenzie, H.L. & Wyman, C.E. Engineering of a high-throughput screening system to identify cellulosic biomass, pretreatments, and enzyme formulations that enhance sugar release. *Biotechnol. Bioeng.* **105**, 231–238 (2010).
67. Yang, B. & Wyman, C.E. in *Biofuels: Methods and Protocols, Methods in Molecular Biology*, vol. 581 (ed. Mielenz, J.R.) 103–114 (Humana Press, New York, 2009).
68. Mielenz, J.R., Rodriguez, M. Jr., Thompson, O.A., Yang, X. & Yin, H. Development of *Agave* as a dedicated biomass source: production of biofuels from whole plants. *Biotechnol. Biofuels* **8**, 79 (2015).
69. Evans, R.J. & Milne, T.A. Molecular characterization of the pyrolysis of biomass. *Energy Fuels* **1**, 123–137 (1987).
70. Sykes, R. *et al.* in *Biofuels: Methods and Protocols, Methods in Molecular Biology*, vol. 581 (ed. Mielenz, J.R.) 169–183 (Humana Press, New York, 2009).
71. Jefferson, R.A., Kavanagh, T.A. & Bevan, M.W. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907 (1987).
72. Sparkes, I.A., Runions, J., Kearns, A. & Hawes, C. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat. Protoc.* **1**, 2019–2025 (2006).
73. Guillaumie, F., Sterling, J.D., Jensen, K.J., Thomas, O.R. & Mohnen, D. Solid-supported enzymatic synthesis of pectic oligogalacturonides and their analysis by MALDI-TOF mass spectrometry. *Carbohydr. Res.* **338**, 1951–1960 (2003).
74. Doong, R.L. & Mohnen, D. Solubilization and characterization of a galacturonosyltransferase that synthesizes the pectic polysaccharide homogalacturonan. *Plant J.* **13**, 363–374 (1998).
75. Sterling, J.D., Lemons, J.A., Forkner, I.F. & Mohnen, D. Development of a filter assay for measuring homogalacturonan: α-(1,4)-Galacturonosyltransferase activity. *Anal. Biochem.* **343**, 231–236 (2005).
76. York, W.S., Darvill, A.G., McNeil, M., Stevenson, T.T. & Albersheim, P. Isolation and characterization of plant cell walls and cell wall components. *Methods Enzymol.* **118**, 3–40 (1986).
77. Heiss, C., Klutts, J.S., Wang, Z., Doering, T.L. & Azadi, P. The structure of *Cryptococcus neoformans* galactoxylomannan contains beta-D-glucuronic acid. *Carbohydr. Res.* **344**, 915–920 (2009).
78. Peña, M.J., Tuomivaara, S.T., Urbanowicz, B.R., O'Neill, M.A. & York, W.S. Methods for structural characterization of the products of cellulose- and xyloglucan-hydrolyzing enzymes. *Methods Enzymol.* **510**, 121–139 (2012).
79. Kumar, R., Hu, F., Hubbell, C.A., Ragauskas, A.J. & Wyman, C.E. Comparison of laboratory delignification methods, their selectivity, and impacts on physiochemical characteristics of cellulosic biomass. *Bioresour. Technol.* **130**, 372–381 (2013).
80. Foston, M.B., Hubbell, C.A. & Ragauskas, A.J. Cellulose isolation methodology for NMR analysis of cellulose ultrastructure. *Materials (Basel)* **4**, 1985–2002 (2011).
81. Davis, M.W. A rapid modified method for compositional carbohydrate analysis of lignocellulosics by high pH anion-exchange chromatography with pulsed amperometric detection (HPAEC/PAD). *J. Wood Chem. Technol.* **18**, 235–252 (1998).
82. Yan, J., Hu, Z., Pu, Y., Brummer, E.C. & Ragauskas, A.J. Chemical compositions of four switchgrass populations. *Biomass Bioenergy* **34**, 48–53 (2010).
83. Meng, X., Wells, T., Sun, Q., Huang, F. & Ragauskas, A. Insights into the effect of dilute acid, hot water or alkaline pretreatment on the cellulose accessible surface area and the overall porosity of *Populus*. *Green Chem.* **17**, 4239–4246 (2015).
84. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

doi:10.1038/nbt.4067

nature research

Corresponding author(s): Dr. Debra Mohnen

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   The sample size (n) of each experiment is provided in the figure/table legends in the main manuscript and supplementary information files. For laboratory and greenhouse experiments: in our laboratory, we regularly use 3 biological replicates with 3 technical replicates (with exceptions for the enzymatic activity data and the cellulose characterization data) which results in reproducible data and the detection of significant changes that support meaningful conclusions. For the switchgrass field trial: there were 12 separate plants (ramets) for each transgenic event arranged in 3 plots with 4 pseudoreplicates per plot. In our experience over the past 7 years of growing transgenic switchgrass in the field, this amount of replication is adequate to detect meaningful (biologically and practically) phenotypic changes rendered by single gene modifications made in switchgrass.

2. Data exclusions

   Describe any data exclusions.

   No data were excluded from the analysis.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   All attempts at replication were successful.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   For the switchgrass field trials: 12 separate plants (ramets) for each event were arranged in 3 plots with 4 pseudoreplicates per plot, which were arranged in a completely randomized design.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   Not applicable.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

Nature Biotechnology: doi:10.1038/nbt.4067

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	Statistica 5.0 and SAS version 9.4 were used for statistical analysis of the data.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	All unique materials used are readily available from the authors.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	For the western blots, the anti-His mouse antibody was from Bio-Rad (cat. # MCA1396GA, clone AD1.1.10, batch # 0215) and was used as per manufacturer's instruction. For ELISA and immunolabeling, the 2F4 monoclonal antibody was a kind gift from Dr. Paul Knox (PlantProbes, www.plantprobes.net), and the JIM5 and JIM7 antibodies were provided by Dr. Michael Hahn (http://www.wallmabdb.net). For the glycome profiling: all the antibodies used have been described previously in Pattathil et al., 2010 (Plant Physiol. 153:514-525) and Pattahil et al., 2012 (Methods Mol. Biol. 908:61-72) . Further information on these antibodies is available at http://www.wallmabdb.net.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	No eukaryotic cell lines were used.
b. Describe the method of cell line authentication used.	No eukaryotic cell lines were used.
c. Report whether the cell lines were tested for mycoplasma contamination.	No eukaryotic cell lines were used.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	Not applicable.

Nature Biotechnology: doi:10.1038/nbt.4067

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	The study did not involve human research participants.

Nature Biotechnology: doi:10.1038/nbt.4067