The Plant Cell, Vol. 30: 1293–1308, June 2018, www.plantcell.org © 2018 ASPB.

Functional Characterization of a Glycosyltransferase from the Moss *Physcomitrella patens* Involved in the Biosynthesis of a Novel Cell Wall Arabinoglucan

Alison W. Roberts,[a,1] Jelle Lahnstein,[b] Yves S.Y. Hsieh,[b,2] Xiaohui Xing,[b,c,3] Kuok Yap,[b,4] Arielle M. Chaves,[a] Tess R. Scavuzzo-Duggan,[a,5] George Dimitroff,[b] Andrew Lonsdale,[d] Eric Roberts,[e] Vincent Bulone,[b,c] Geoffrey B. Fincher,[b] Monika S. Doblin,[d,6] Antony Bacic,[d,6] and Rachel A. Burton[b]

[a]Department of Biological Sciences, University of Rhode Island, Kingston, Rhode Island 02881
[b]ARC Centre of Excellence in Plant Cell Walls, School of Agriculture, Food, and Wine, University of Adelaide, Urrbrae, South Australia 5064, Australia
[c]Division of Glycoscience, Department of Chemistry, School of Engineering Sciences in Chemistry, Biotechnology, and Health, Royal Institute of Technology (KTH), Stockholm SE-10691, Sweden
[d]ARC Centre of Excellence in Plant Cell Walls, Plant Cell Biology Research Centre, School of BioSciences, The University of Melbourne, Victoria 3010, Australia
[e]Biology Department, Rhode Island College, Providence, Rhode Island 02908

ORCID IDs: 0000-0002-7775-5589 (A.W.R.); 0000-0002-0968-5793 (Y.S.Y.H.); 0000-0002-4567-6712 (A.M.C.); 0000-0002-9297-8123 (E.R.); 0000-0003-4703-3494 (G.B.F.); 0000-0002-8921-2725 (M.S.D.); 0000-0001-7483-8605 (A.B.)

Mixed-linkage (1,3;1,4)-β-glucan (MLG), an abundant cell wall polysaccharide in the Poaceae, has been detected in ascomycetes, algae, and seedless vascular plants, but not in eudicots. Although MLG has not been reported in bryophytes, a predicted glycosyltransferase from the moss *Physcomitrella patens* (Pp3c12_24670) is similar to a bona fide ascomycete MLG synthase. We tested whether *Pp3c12_24670* encodes an MLG synthase by expressing it in wild tobacco (*Nicotiana benthamiana*) and testing for release of diagnostic oligosaccharides from the cell walls by either lichenase or (1,4)-β-glucan endohydrolase. Lichenase, an MLG-specific endohydrolase, showed no activity against cell walls from transformed *N. benthamiana*, but (1,4)-β-glucan endohydrolase released oligosaccharides that were distinct from oligosaccharides released from MLG by this enzyme. Further analysis revealed that these oligosaccharides were derived from a novel unbranched, unsubstituted arabinoglucan (AGlc) polysaccharide. We identified sequences similar to the *P. patens* AGlc synthase from algae, bryophytes, lycophytes, and monilophytes, raising the possibility that other early divergent plants synthesize AGlc. Similarity of *P. patens* AGlc synthase to MLG synthases from ascomycetes, but not those from Poaceae, suggests that AGlc and MLG have a common evolutionary history that includes loss in seed plants, followed by a more recent independent origin of MLG within the monocots.

INTRODUCTION

The *CELLULOSE SYNTHASE* (*CESA*) gene superfamily of angiosperms consists of *CESA*s and nine families of *CESA*-like genes designated *CSLA*-*CSLJ* (excluding *CSLI*). These genes encode carbohydrate active enzyme (CAZy) family 2 glycosyltransferases (GT2s) that synthesize linear cell wall polysaccharides such as cellulose, mannan, glucomannan, and mixed-linkage (1,3;1,4)-β-glucan (MLG), as well as the backbones of branched polysaccharides such as xyloglucan and (galacto)glucomannan (Fincher, 2009). In contrast to angiosperms, the moss *Physcomitrella patens* and the lycophyte *Selaginella moellendorffii* have less diversified *CESA* superfamilies that include only members of the *CESA*, *CSLA*, *CSLC*, and *CSLD* families (Roberts and Bushoven, 2007; Harholt et al., 2012). However, genes encoding GT2s that do not cluster with either the angiosperm *CESA* or *CSL* families were also identified (Harholt et al., 2012). These GT2s are similar to a group of ascomycete sequences, one of which (*Aspergillus fumigatus* sequence; Figure 6 in Harholt et al. [2012]) encodes an MLG synthase (Samar et al., 2015). Orthologs of these *P. patens* and *S. moellendorffii* genes have not been found in seed plants, and they have low sequence similarity to the *CSLF*, *CSLH*, and *CSLJ* genes (Harholt et al., 2012) that encode MLG synthases in commelinid monocots (Burton et al., 2006; Doblin et al., 2009; Fincher, 2009).

[1]Address correspondence to aroberts@uri.edu.
[2]Current address: Division of Glycoscience, Department of Chemistry, School of Engineering Sciences in Chemistry, Biotechnology, and Health, Royal Institute of Technology (KTH), Stockholm SE-10691, Sweden.
[3]Current address: Agriculture and Agri-Food Canada, Lethbridge Research Centre, Ottawa, Ontario K1A0C5, Canada.
[4]Current address: Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia.
[5]Current address: Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720.
[6]Current address: La Trobe Institute for Agriculture and Food, Department of Animal, Plant and Soil Sciences, School of Life Sciences, La Trobe University, Bundoora, Victoria 3086, Australia.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Alison W. Roberts (aroberts@uri.edu).
www.plantcell.org/cgi/doi/10.1105/tpc.18.00082

IN A NUTSHELL

Background: The health-promoting soluble fiber in whole grains is rich in mixed-linkage glucan (MLG), so called because it consists of glucose joined through both 1,4- and 1,3-linkages. MLG is soluble because these linkages are distributed irregularly along the polysaccharide chain, so scientists want to understand how the enzymes that make MLG control linkage placement. MLG is common in cereals, but it is not found in most other flowering plants. Some seedless plants and algae make MLGs with distinct arrangements of 1,4- and 1,3-linkages. These plants do not have genes like those that encode cereal MLG synthases. So, it appears that MLG synthesis had a complex evolutionary history that may help us understand how variations in synthases affect the structure of the polysaccharides they produce.

Question: Our discovery that a gene from the moss *Physcomitrella patens* resembles a known fungal MLG synthase gene led us to ask whether the moss gene also encodes an MLG synthase. To find out, we transformed the moss gene into a wild tobacco plant that normally makes no MLG but is especially good at expressing foreign proteins. We treated the cell walls from the transformed tobacco with two enzymes that digest MLG and looked for the cell wall fragments that these enzymes usually release.

Findings: The enzyme that digests only MLG (lichenase) did not release fragments from the tobacco cell walls. The less specific enzyme (β-glucanase) released fragments, but they were not the ones we expected. They contained 1,4-linked glucose and 1,3-linked arabinose. Because the β-glucanase did not release fragments from control tobacco, we knew that the fragments had come from a polysaccharide produced by the moss enzyme. Although it contains 1,4- and 1,3-linked sugars like MLG, this polysaccharide differs from known plant cell wall polysaccharides and we named it arabinoglucan.

Next steps: We found genes that encode proteins similar to the moss arabinoglucan synthase in algae and seedless plants, including some that are known to make MLG. By testing the activities of these proteins, we can learn more about how enzyme activities evolve and how synthase variation is related to polysaccharide structure.

MLG is a major noncellulosic cell wall polysaccharide in cereals and a health-promoting dietary fiber (Fincher, 2016). Structurally, MLG is an unbranched, unsubstituted homopolymer of Glc consisting of cellobiosyl and cellotriosyl residues ($G4G4G_{Red}$ and $G4G4G4G_{Red}$, respectively, where "$_{Red}$" denotes the reducing end) connected by (1,3)-β-linkages with some longer runs of (1,4)-β-linkages (Woodward et al., 1983). In addition to its widespread occurrence in cereals and other grasses (Fincher, 2016), MLG has been detected in organisms representing widely divergent taxa (Harris and Fincher, 2009; Popper et al., 2011) ranging from red, green, and brown algae to leptosporangiate ferns and horsetails, fungi (basidiomycete and ascomycete species, including lichens), and Gram-negative bacteria (Gorin et al., 1988; Fontaine et al., 2000; Lechat et al., 2000; Honegger and Haisch, 2001; Olafsdottir and Ingólfsdottir, 2001; Pacheco-Sanchez et al., 2006; Eder et al., 2008; Fry et al., 2008; Sørensen et al., 2008, 2011; Pettolino et al., 2009; Leroux et al., 2015; Pérez-Mendoza et al., 2015; Salmeán et al., 2017). However, it has not been detected in eudicots (Burton and Fincher, 2009).

A key method for detecting MLG in cell walls is based on treating samples with lichenase enzymes. These (1,3;1,4)-β-glucanohydrolases (EC 3.2.1.73) cleave the (1,4)-β-linkages on the reducing terminal side of single (1,3)-β-linkages in MLGs, producing diagnostic tri- and tetrasaccharides (Moscatelli et al., 1961; Meikle et al., 1994), while showing no activity against other plant cell wall polysaccharides (Anderson and Stone, 1975). In addition to commelinid monocots (Fincher, 2016), diagnostic oligosaccharides have been detected in lichenase hydrolysates of cell walls from horsetails (Fry et al., 2008; Sørensen et al., 2008), green algae (Eder et al., 2008; Sørensen et al., 2011), brown algae (Salmeán et al., 2017), lichens (Gorin et al., 1988; Honegger and Haisch, 2001), a Gram-negative bacterium (Pérez-Mendoza et al., 2015), and the ascomycete fungus *Rhynchosporium secalis* (Pettolino et al., 2009). Additionally, colorimetric analysis of oligosaccharides released by lichenase (McCleary and Codd, 1991) provided evidence for the presence of MLG in *A. fumigatus* (Samar et al., 2015). In contrast, oligosaccharides released by lichenase treatment of cells walls of a leafy liverwort (*Lophocolea bidentata*) and a green alga (*Ulva lactuca*) were shown to contain Ara and Xyl, respectively, along with Glc, indicating that lichenase may not be absolutely specific for commelinid monocot-type MLGs (Popper and Fry, 2003).

Although the presence of MLG in horsetail cell walls is well established, the corresponding MLG synthases have not been identified. Based on taxonomic position, it is unlikely that horsetail genomes contain *CSLF*, *CSLH*, or *CSLJ* orthologs, suggesting multiple origins of MLG synthases in streptophytes (Sørensen et al., 2008, 2011; Burton and Fincher, 2009). The genomes of *P. patens* and *S. moellendorffii* encode GT2s that are similar to a group of proteins from ascomycetes (Harholt et al., 2012). One of these sequences from *A. fumigatus* was later shown to correspond to an MLG synthase (Af-Tft1; Samar et al., 2015), raising the possibility that the *P. patens* and *S. moellendorffii* sequences could also encode MLG synthases. As an experimental organism, *P. patens* has the advantage of well-established methods for rapid gene modification (Cove et al., 2009). The complete *P. patens* genome sequence (Rensing et al., 2008; Zimmer et al., 2013) contains only one gene that encodes an MLG synthase-like GT2 (Harholt et al., 2012).

Here, we show that this *P. patens* putative GT2 protein (Pp3c12_24670), with sequence similarity to a fungal MLG synthase, does not produce a commelinid monocot-type MLG, but instead synthesizes an unbranched, unsubstituted polysaccharide consisting of (1,4)-β-linked glucosyl residues interspersed with single (1,3)-linked arabinofuranosyl residues. This unusual arabinoglucan (AGlc) is structurally similar to MLG but is not cleaved by lichenase. Sequences similar to Pp3c12_24670 were identified in a variety of seedless plants and green algae,

suggesting that AGlc may be a relatively common component of the cell walls of these organisms.

RESULTS

Heterologous Expression of *Pp3c12_24670*

Based on the similarity of Pp3c12_24670 to both a fungal MLG synthase (Samar et al., 2015) and proteins encoded by the *S. moellendorffii* genome (Harholt et al., 2012), along with evidence that the cell walls of *S. moellendorffii* contain MLG (Harholt et al., 2012), we tested the hypothesis that *Pp3c12_24670* encodes an MLG synthase. We transiently transformed wild tobacco (*Nicotiana benthamiana*), an organism that does not normally produce MLG, with vectors driving the expression of either *Pp3c12_24670* (experimental treatment), *Hv-CSLF6* (the positive control MLG synthase gene isolated from barley [*Hordeum vulgare*]; Schreiber et al., 2014), catalytically inactive *Pp3c12_24670* (D247N/D249N mutant; negative control), or an empty pEAQ-HT-DEST1 vector (negative control). Alcohol-insoluble residue (AIR) cell wall preparations from transformed *N. benthamiana* leaves were hydrolyzed either with lichenase [(1,3;1,4)-β-glucanohydrolase from *Bacillus* sp] or with E-CELAN [endo-(1,4)-β-glucanohydrolase from *Aspergillus niger*]. Lichenase is used extensively for cereal MLG detection and releases 4-*O*-β-D-glucosyl-3-*O*-β-D-glucosyl-β-D-glucose and 4-*O*-β-cellobiosyl-3-*O*-β-D-glucosyl-β-D-glucose ($G4G3G_{Red}$ and $G4G4G3G_{Red}$) as diagnostic products (Meikle et al., 1994). E-CELAN also has high activity against barley MLG (McCleary, 2008), releasing 3-*O*-β-D-glucosyl-β-D-cellobiose and 4-*O*-β-glucosyl-3-*O*-β-D-glucosyl-β-D-cellobiose ($G3G4G_{Red}$ and $G4G3G4G_{Red}$). Using high-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD), we detected the expected MLG tri- and tetrasaccharides in lichenase hydrolysates from leaves transformed with *Hv-CSLF6*, the positive control MLG synthase gene from barley ($G4G3G_{Red}$ and $G4G4G3G_{Red}$; eluting at 12.5 and 16 min; Figure 1A). These oligosaccharides were not detected in lichenase hydrolysates from *Pp3c12_24670*-transformed leaves (Figure 1A). E-CELAN hydrolysates from leaves transformed with *Hv-CSLF6* also contained the expected MLG tri- and tetrasaccharides ($G3G4G_{Red}$ and $G4G3G4G_{Red}$; eluting at 15.9 and 19.6 min; Figure 1B). However, E-CELAN hydrolysates from leaves transformed with *Pp3c12_24670* produced unique profiles with peaks at 15.5 (doublet), 19.5, and 22.5 min (Figure 1B, asterisks), rather than the expected MLG oligosaccharide peaks. As negative controls, we tested the activity of lichenase and E-CELAN against AIR cell wall preparations from *N. benthamiana* leaves transformed with either catalytically inactive *Pp3c12_24670* or an empty vector. No peaks were detected between 10 and 25 min in the HPAEC-PAD profiles from these control hydrolysates (Figures 1A and 1B). The presence of transcripts from the *Pp3c12_24670* and catalytically inactive *Pp3c12_24670* transgenes was verified by RT-PCR (Figure 1C). Based on these results, we concluded that enzymatic activity of the transgenic Pp3c12_24670 protein generates an E-CELAN-digestible polysaccharide that is distinct from MLG. When incubated with an endo-(1,4)-β-D-glucanohydrolase from the fungus *Trichoderma longibrachiatum* (E-CELTR), AIR cell wall preparations of *N. benthamiana* leaves transformed with *Pp3c12_24670* released oligosaccharides with the same mobility as the oligosaccharides released by E-CELAN (Supplemental Figure 1). In addition, the E-CELTR hydrolysates contained numerous compounds that could be separated as individual peaks and that were also present in hydrolysates of *N. benthamiana* leaves transformed with the empty vector, including one at 14.5 min that was significantly more abundant than in the empty vector control (Supplemental Figure 1). These are likely due to the broader specificity of this enzyme (https://secure.megazyme.com/files/Booklet/E-CELTR_DATA.pdf).

The oligosaccharides released by E-CELAN from AIR cell wall preparations of *N. benthamiana* leaves transformed with wild-type *Pp3c12_24670* were separated by HPLC. The collected fractions were analyzed by HPAEC-PAD, and those containing the 19.5 and 22.5 min peaks (Supplemental Figure 2) were further analyzed. The 19.5 min peak had a monosaccharide composition of 75.6% Glc and 21.8% Ara and a mass of 637.2182 $[M+H]^+$, as expected for a tetrasaccharide composed of three hexosyl residues and one pentosyl residue (Supplemental Figure 3). This was confirmed by linkage analysis, which showed that the tetrasaccharide was linear, consisting of 50.2% 4-Glc_p, 26.1% t-Glc_p, and 17.6% 3-Ara_f (Supplemental Table 1). The compound corresponding to the 22.5 min peak had a monosaccharide content of 79.0% Glc_p and 17.0% Ara_f and a mass of 799.2705 $[M+H]^+$, as expected for a pentasaccharide composed of four hexosyl and one pentosyl residue (Supplemental Figure 3). Taken together, these results indicate that the oligosaccharides released by E-CELAN from AIR preparations of leaves of *N. benthamiana* transformed with *Pp3c12_24670* were derived either from an unbranched, unsubstituted AGlc consisting of 4-linked Glc_p and 3-linked Ara_f or a more complex polysaccharide with AGlc domains.

Polysaccharide Fractionation

To test the solubility of the AGlc-containing polysaccharide, we sequentially extracted leaf AIR cell wall preparations from *Pp3c12_24670*-transformed *N. benthamiana* with (1,2)-cyclohexylenedinitrilotetraacetic acid (CDTA) and 4 M NaOH and hydrolyzed the compounds in both extracts with E-CELAN. The hydrolysate obtained from the CDTA extract (Supplemental Figure 4) primarily contained the AGlc oligosaccharides (Figure 1B). The hydrolysate of the NaOH-soluble extract contained these oligosaccharides as well as many others, whereas no AGlc oligosaccharides were released by E-CELAN from the 4 M NaOH-insoluble residue (Supplemental Figure 4). Thus, to isolate the AGlc-containing polysaccharide from *N. benthamiana* cell walls, we started with a CDTA extract, precipitated charged polysaccharides with cetyltrimethylammonium bromide (CTAB), and fractionated the supernatant by reverse-phase HPLC (RP-HPLC). Profiles of extracts from AIR preparations of *Pp3c12_24670*-transformed leaves had a large peak between 5 and 7 min (fractions 9–13) that was absent from profiles of empty vector control extracts (Figure 2A). The fractions comprising this peak (9–13) showed the highest release of the diagnostic

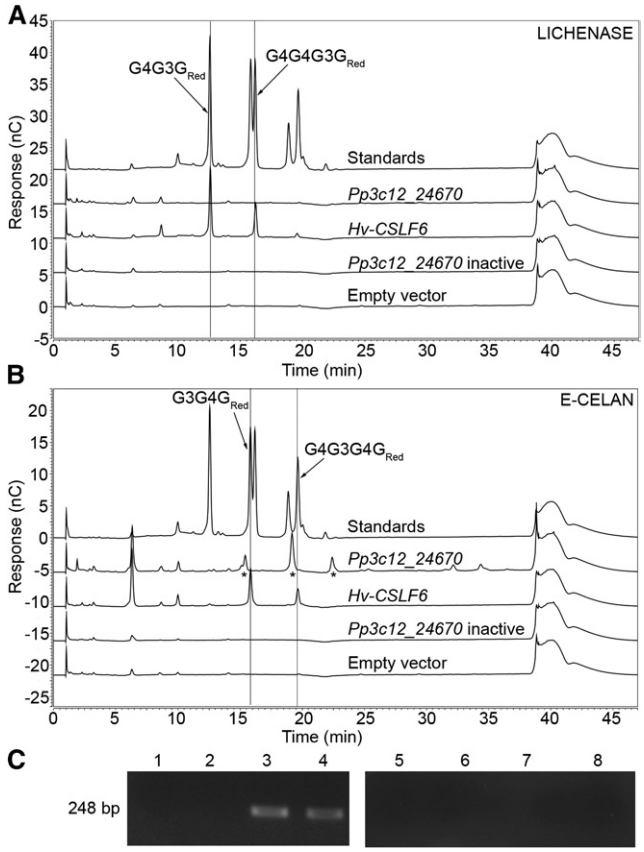

Figure 1. Endoglucanase Releases Unexpected Oligosaccharides from AIR Preparations of *N. benthamiana* Leaves Expressing *Pp3c12_24670.*

(A) and **(B)** HPAEC-PAD profiles of lichenase [(1,3;1,4)-β-D-4-glucanohydrolase from *Bacillus* sp] **(A)** and E-CELAN (endo-1,4-β-D-glucanohydrolase from *A. niger*) **(B)** hydrolysates of AIR cell wall preparations from *N. benthamiana* leaves transformed with either *Pp3c12_24670*, *Hv-CslF6* (MLG positive control), catalytically inactive *Pp3c12_24670* expression vectors, or an empty negative control vector with a mixture of G4G3G$_{Red}$, G4G4G3G$_{Red}$, G3G4G$_{Red}$, and G4G3G4G$_{Red}$ standards (Megazyme; top chromatogram in **[A]** and **[B]**).
(C) RT-PCR of *Pp3c12_24670* in *N. benthamiana* leaves transformed with either empty vector (lane 1), *Hv-CslF6* (lane 2), *Pp3c12_24670* (lane 3), or catalytically inactive *Pp3c12_24670* (lane 4) expression vectors, with no RT controls for the same samples (lanes 5–8).

tetrasaccharide when treated with E-CELTR (Figure 2B; Supplemental Figure 5) and the highest fluorescence in the presence of Fluorescent Brightener 28 (FB28, Calcofluor; Figure 2B), as expected for a β-glucan polysaccharide with a molecular mass >10 kD (Rieder et al., 2015). Size-exclusion chromatography (SEC) with postcolumn detection using FB28 showed a peak eluting after the lowest MLG standard, consistent with a molecular mass substantially lower than 33.6 kD (Supplemental Figure 6). Linkage analysis showed that 4-Glc$_p$, 3-Ara$_f$, and t-Glc$_p$ residues accounted for nearly 98% of the AGlc-enriched fractions with a 7.5:1 Glc$_p$:Ara$_f$ ratio (Table 1). The other residues detected in the linkage analysis (ranging from ~0.4 to 1.2%) likely arose from minor contaminating carbohydrates, although their covalent association to the AGlc cannot be ruled out. Lichenase treatment of fractions containing the polysaccharide did not release detectable oligosaccharides (Figure 2). The polysaccharide was also partially soluble in hot water (60°C, 4 h), with an additional fraction solubilized by extraction of the hot water residue with 6 M guanidinium-HCl at room temperature (Supplemental Figure 7).

Pp3c12_24670 Is Expressed in Older *P. patens* Cells

To localize the expression of *Pp3c12_24670*, we produced promoter-reporter fusion lines by transforming wild-type *P. patens* with a vector containing 1963 bp of sequence upstream of the *Pp3c12_24670* coding sequence in front of the *GUS* coding sequence. Six of eight stable lines from a transformation with this

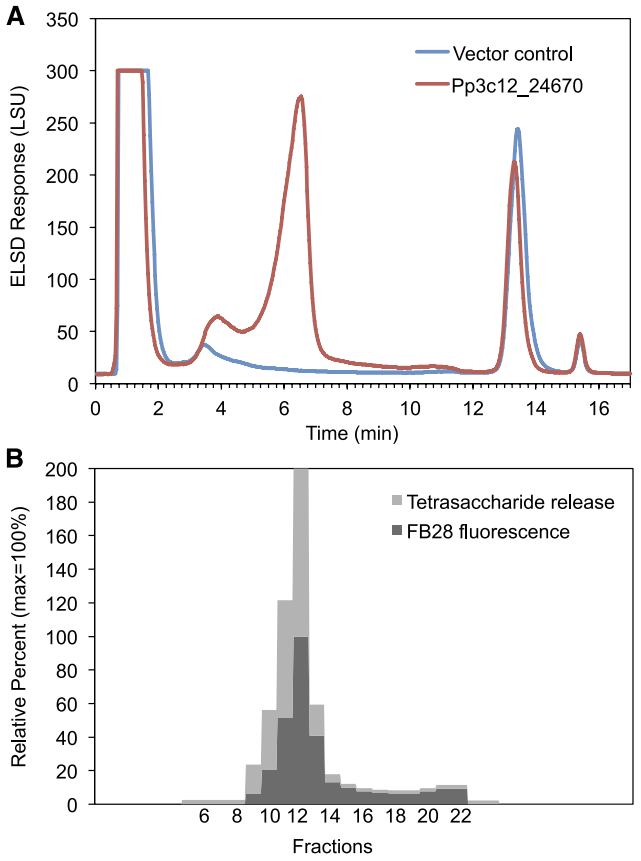

Figure 2. Fractions Enriched in Mixed-Linkage Arabinoglucan Were Identified by Oligosaccharide Release and Dye Binding.

(A) RP-HPLC of supernatants from CTAB precipitation of CDTA extracts of AIR cell wall preparations from *N. benthamiana* leaves shows material eluting from 3 to 12 min in leaves transformed with *Pp3c12_24670* (red), but not in the empty negative control vector (blue).
(B) Fractions collected at 30 s intervals between 2.5 and 13 min (fractions 6–23) were tested for fluorescence intensity in the presence of FB28 and the release of tetrasaccharide by E-CELTR, as detected by HPAEC-PAD (Supplemental Figure 5). Stacked bars report percentages of the maximum value, which is set at 100%. Fractions 11 and 12 were also digested with lichenase and no released oligosaccharides were detected.

Table 1. Linkage Analysis by Methylation of RP-HPLC Fractionated Material (Eluting between 5 and 7 min; Figure 2A) from a CDTA-Extracted AIR Cell Wall Preparation from *N. benthamiana* Transformed with *Pp3c12_24670*

Glycosyl Linkage	Molar Composition (mol %)
1,3-Ara$_f$	11.5
1,4-Glc$_p$	83.0
t-Glc$_p$	3.3
t-Xyl$_p$	0.6
1,4-Man$_p$	1.2
t-Gal$_p$	0.5

vector were positive for GUS activity. Although staining intensity varied, all positive lines showed the highest staining in the older primary filaments and subapical cells of secondary filaments and less staining in the younger apical cells (Figure 3A). Little or no staining was detected in gametophores with up to five expanded leaves (Figure 3B), while for those with more than five expanded leaves, GUS staining was detected in maturing leaves (Figure 3C). Although we did not detect staining in the rhizoids themselves, stained protonema were sometimes intermixed with the rhizoids (Figures 3B and 3C).

Pp3c12_24670 Is Not Required for Normal Growth and Development of *P. patens*

PCR genotyping of *P. patens* lines transformed with a *Pp3c12_24670* knockout vector confirmed four lines in which the vector was integrated by homologous recombination at the 5′ and 3′ ends and the target gene was undetectable (Supplemental Figure 8). We were unable to detect morphological or developmental defects in the knockout mutants grown under normal laboratory conditions.

Pp3c12_24670 Synthesizes AGlc in the Native Host

To test the activity of *Pp3c12_24670* in the native host, *P. patens* was transformed with vectors that drive the expression of hemagglutinin-tagged Pp3c12_24670 or catalytically inactive Pp3c12_24670 proteins with the *Act1-F* promoter from rice (*Oryza sativa*) (McElroy et al., 1990), which drives strong, constitutive expression in *P. patens* (Schaefer, 2002; Horstmann et al., 2004). We detected the expected AGlc-derived 15.5, 19.5, and 22.5 min oligosaccharide peaks in E-CELAN hydrolysates of CDTA extracts of 6-d-old protonema from lines transformed with *Pp3c12_24670*, but not in lines transformed with catalytically inactive *Pp3c12_24670*, the empty control vector, or *Pp3c12_24670* knockout lines (Figure 4A). Protein gel blotting confirmed expression of the tagged wild-type and catalytically inactive proteins (Figure 4B). The expected peaks were also not detected in E-CELAN hydrolysates of CDTA extracts of 6-d-old protonema from wild-type *P. patens* (Figure 4A). However, we did detect the expected peaks in E-CELAN hydrolysates of AIR cell wall preparations that were never dried (to preserve solubility) from both 8-d-old protonema and mature gametophores from wild-type *P. patens* (Figures 4C and 4D). The hydrolysate from protonema contained additional peaks of similar size that were also detected in *Pp3c12_24670*-transformed protonema when equal volumes of extract were loaded (Figure 4C). The hydrolysate from gametophores contained peaks that were larger than the expected peaks and not present in the minus-E-CELAN negative control (Figure 4D). Due to the complexity of these extracts and the low abundance of the expected peaks, we did not attempt to isolate AGlc-derived oligosaccharides from wild-type *P. patens*.

Pp3c12_24670 Resides in Cytosolic Vesicles

We examined the subcellular localization of Pp3c12_24670 in *P. patens* lines transformed with an mEGFP:Pp3c12_24670 fusion protein expression vector. Stable lines were screened for monomeric enhanced GFP (mEGFP) fluorescence using a fluorescence stereomicroscope, and three lines exhibiting strong fluorescence were selected for further analysis. The presence of the expected AGlc-derived 15.5, 19.5, and 22.5 min oligosaccharide peaks in E-CELAN hydrolysates from all three lines (Figure 4A) confirmed that the tagged protein was enzymatically active. mEGFP fluorescence was localized to cytoplasmic vesicles and was not detected in the plasma membrane in apical and subapical cells of the protonema (Figure 5B). The size and distribution of fluorescent vesicles was similar in an (1,2)-α-mannosidase Golgi marker line (Figure 5D; Furt et al., 2012). The mEGFP:Pp3c12_24670 and YFP:GmMan lines were also similar in their response to a 30 min treatment with 1 μg mL^{-1} Brefeldin A (Figures 5C to 5E), including a loss of punctate fluorescence and increase in diffuse cortical and perinuclear fluorescence.

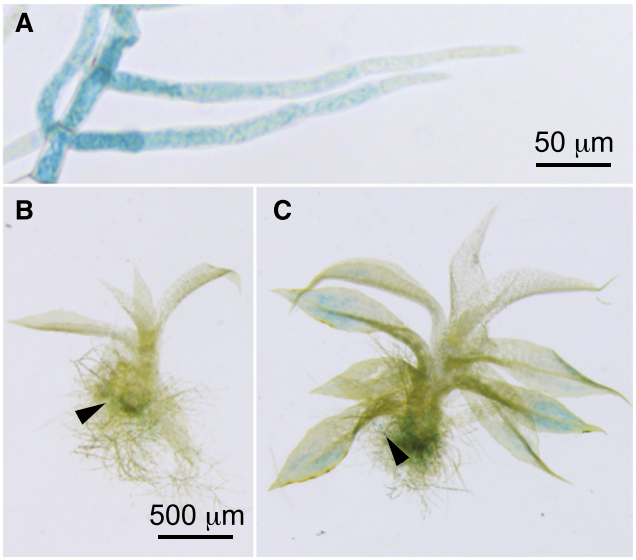

Figure 3. The *Pp3c12_24670* Promoter Is Active in Mature Tissues.

(A) In *P. patens Pp3c12_24670pro*:*GUS* lines, GUS staining was detected in subapical protonemal cells from 11-d-old cultures. Young gametophores from 14-d-old cultures showed no staining.

(B) Gametophores from 21-d-old cultures showed GUS staining in the more mature basal leaves.

Arrowheads in **(B)** and **(C)** indicate protonema. Bar in **(B)** also applies to **(C)**.

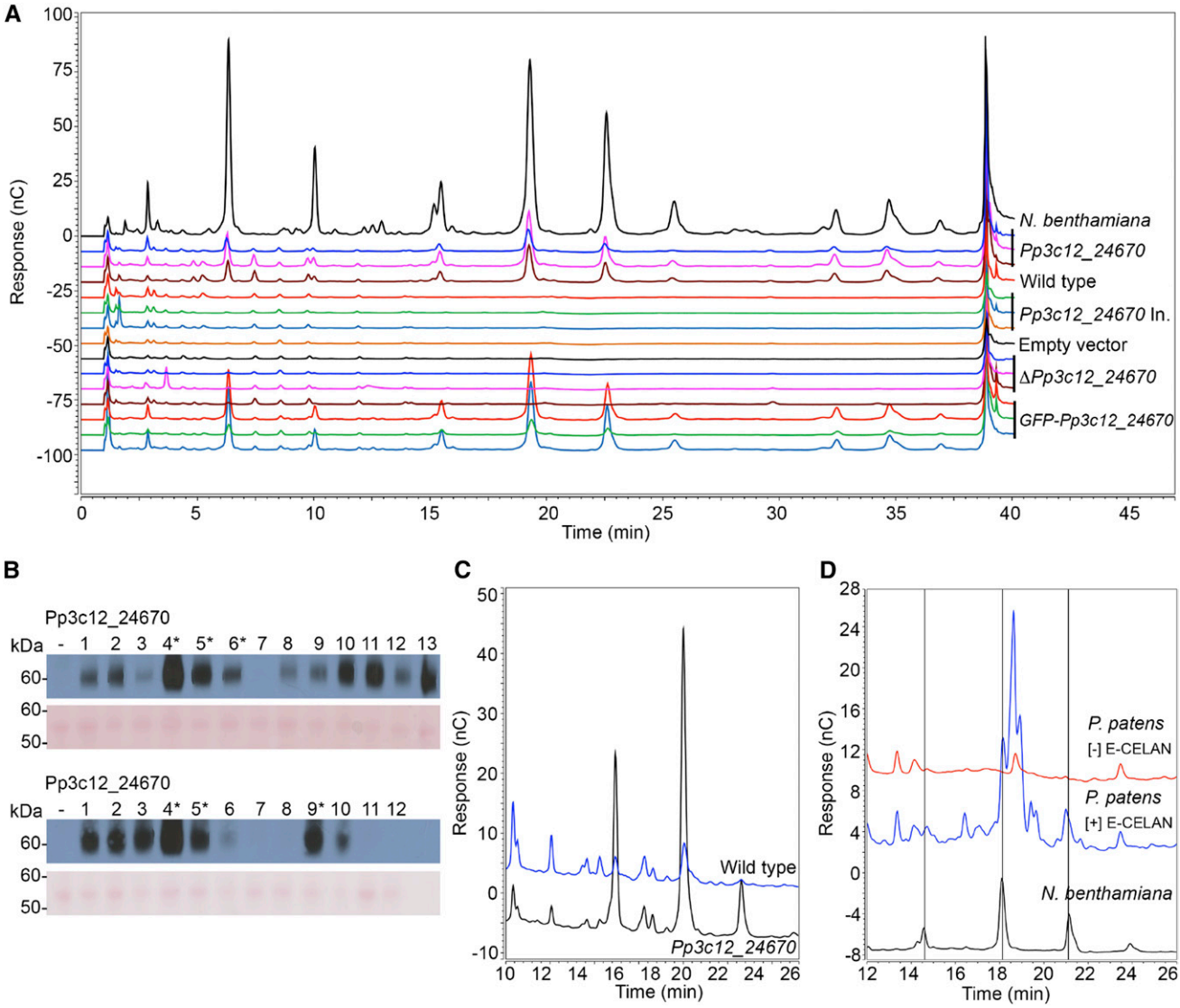

Figure 4. Pp3c12_24670 Is Active When Expressed in the Native Host.

(A) HPAEC-PAD profiles of E-CELAN hydrolysates of CDTA-soluble fractions of AIR preparations from *N. benthamiana* leaves transformed with *Pp3c12_24670* expression vector (positive control) or 6-d-old protonema from *P. patens* transformed with *Pp3c12_24670* expression vector, *P. patens* wild type, *P. patens* transformed with catalytically inactive *Pp3c12_24670* expression vector or empty negative control vector, *P. patens Pp3c12_24670* knockout lines, or *P. patens* transformed with *mEGFP*: *Pp3c12_24670* fusion protein expression vector.
(B) Protein gel blots of protein extracted from stable *P. patens* lines transformed with *Pp3c12_24670* or catalytically inactive *Pp3c12_24670* expression vectors. Blots probed with antihemagglutinin are shown above the same blot stained with Ponceau S as a loading control. Protein loading per lane was 3.3 and 5.8 μg per lane, respectively.
(C) HPAEC-PAD profiles of E-CELAN hydrolysates from equal masses of never-dried cell walls from 8-d-old protonema of wild type (blue) and *Pp3c12_24670* transformed (black) *P. patens*.
(D) HPAEC-PAD profiles of E-CELAN and no-enzyme control hydrolysates of never-dried cell walls from mature *P. patens* gametophores with positive control *N. benthamiana* transformed with *Pp3c12_24670*.

Pp3c12_24670 Is Similar to MLG Synthases

An amino acid alignment (Figure 6; Supplemental Figure 9) of Pp3c12_24670 with known MLG synthases from barley (Hv-CSLH1 and Hv-CSLF6), an ascomycete (Af-Tft1), and a prokaryote (SDW48993.1) showed that all five sequences contain the D, D, D, QxxRW motif that is characteristic of most GT2 family members (Figure 6). Only the barley sequences include segments known as the plant conserved region (amino acids 270–430) and class-specific region (amino acids 560–670) found in CESAs and some CSLs (Supplemental Figure 9). Pairwise alignments show higher identity of Pp3c12_24670 with

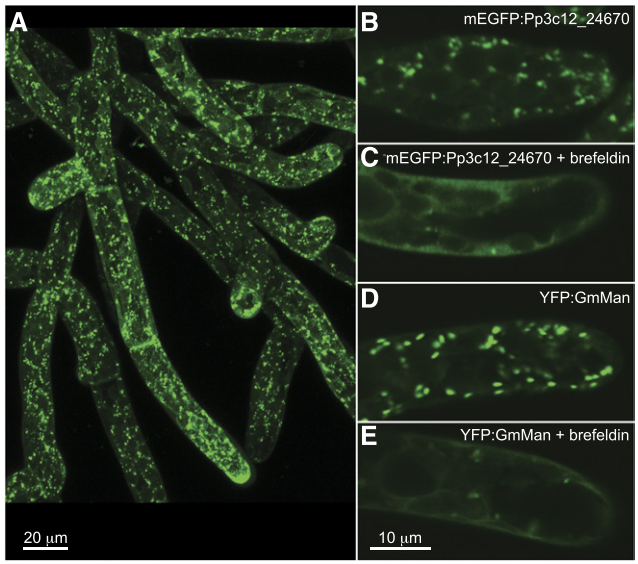

Figure 5. mEGFP:Pp3c12_24670 Fusion Protein Resides in Cytoplasmic Vesicles.

Punctate mEGFP:Pp3c12_24670 throughout the *P. patens* protonema, as displayed in a maximum projection of 16 confocal sections **(A)**. Treatment for 30 min with 1 μg mL^{-1} Brefeldin A abolishes punctate fluorescence while increasing diffuse cortical and perinuclear fluorescence in confocal sections of *P. patens* expressing fluorescent Pp3c12_24670 (mEGFP:Pp3c12_24670; compare **[B]** and **[C]**) and in *P. patens* expressing a fluorescent Golgi marker (YFP:GmMan; compare **[D]** and **[E]**).

Af-Tft1 (25.3%) compared with *Hv*-CSLH1 and *Hv*-CSLF6 (20.0% and 17.5%, respectively).

Genes with Similarity to *Pp3c12_24670* Are Found in Algae and Seedless Plants

To determine whether other species have similar genes, we used the Pp3c12_24670 peptide sequence as a query to search data generated by the 1000 Plants (1KP) Consortium (http://www.onekp.com/). These data include transcriptomes from over 1000 species representing all major lineages of the Viridiplantae (Matasci et al., 2014). We identified 473 sequences from members of all major clades of algae and seedless plants (Supplemental File 1). The list also included 20 seed plant sequences. When tested by TBLASTN analysis against GenBank sequences, all showed high identity to either CSLA (81–100% identity) or CSLC sequences (77–93% identity) or were more than 99% identical to fungal sequences (see Methods). The sequences with high identity to CSLA and CSLC sequences were judged to be members of those families and were excluded from the analysis. The sequences with near identity to fungal sequences (>99%) were judged as likely to be from fungal contaminants and were also excluded. An alignment of 243 sequences remaining after the removal of short sequences (<200 bp; see Methods) with Pp3c12_24670 and five *S. moellendorffii* sequences (Harholt et al., 2012) added (Supplemental File 2) had a length of 197 amino acids after editing (Supplemental File 3). In the unrooted maximum likelihood tree generated from the alignment of 243 sequences, many of the clades that contain sequences from major taxonomic groups are supported by high bootstrap values, whereas the order of divergence for these clades is not well resolved (Supplemental Figure 10).

To test whether any of these sequences cluster with functionally characterized GT2s, we created a new alignment that included (1) representative members of the large moss, fern, charophyte, chlorophyte, and red algal clusters (Supplemental Figure 10), (2) all sequences from the tree that are not members of these large clusters, (3) all members of the *P. patens* CESA/CSL superfamily (Roberts and Bushoven, 2007), and (4) functionally characterized MLG synthases (Supplemental Figure 9) and prokaryotic cellulose and curdlan (1,3-β-glucan) synthases (Supplemental Files 4 and 5). The tree shows several well-supported clusters radiating from a central polytomy comprising the unresolved internal nodes. The well-supported clusters (bootstrap values = 94–100%, shown in red in Supplemental

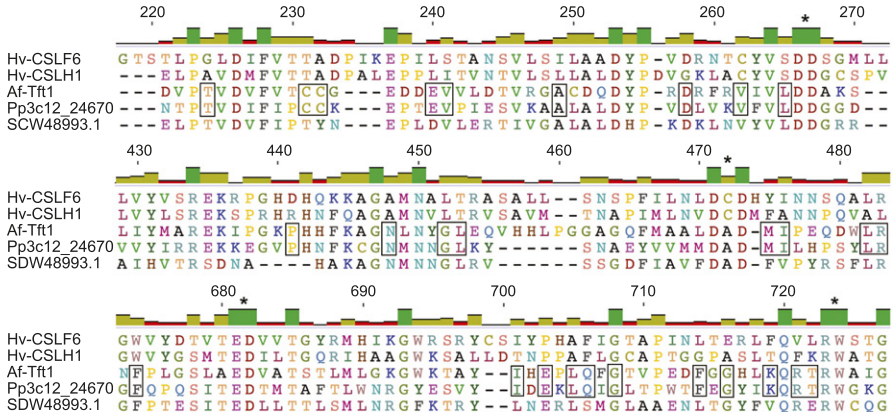

Figure 6. Pp3c12_24670 Is More Similar to Ascomycete MLG Synthase Than Barley MLG Synthases.

Segments from an MUSCLE alignment of Pp3c12_24670 with MLG synthases including Hv-CSLF6 from barley (Burton et al., 2006), Hv-CSLH1 from barley (Doblin et al., 2009), Af-Tft1 from *A. fumigatus* (Samar et al., 2015), and SDW48993.1 from *Sinorhizobium meliloti* (Pérez-Mendoza et al., 2015). Elements of the conserved D, D, D, QXXRW motif characteristic of GT2 are indicated with asterisks. Residues shared exclusively by Pp3c12_24670 and Af-Tft1 are outlined. The complete alignment is shown in Supplemental Figure 9.

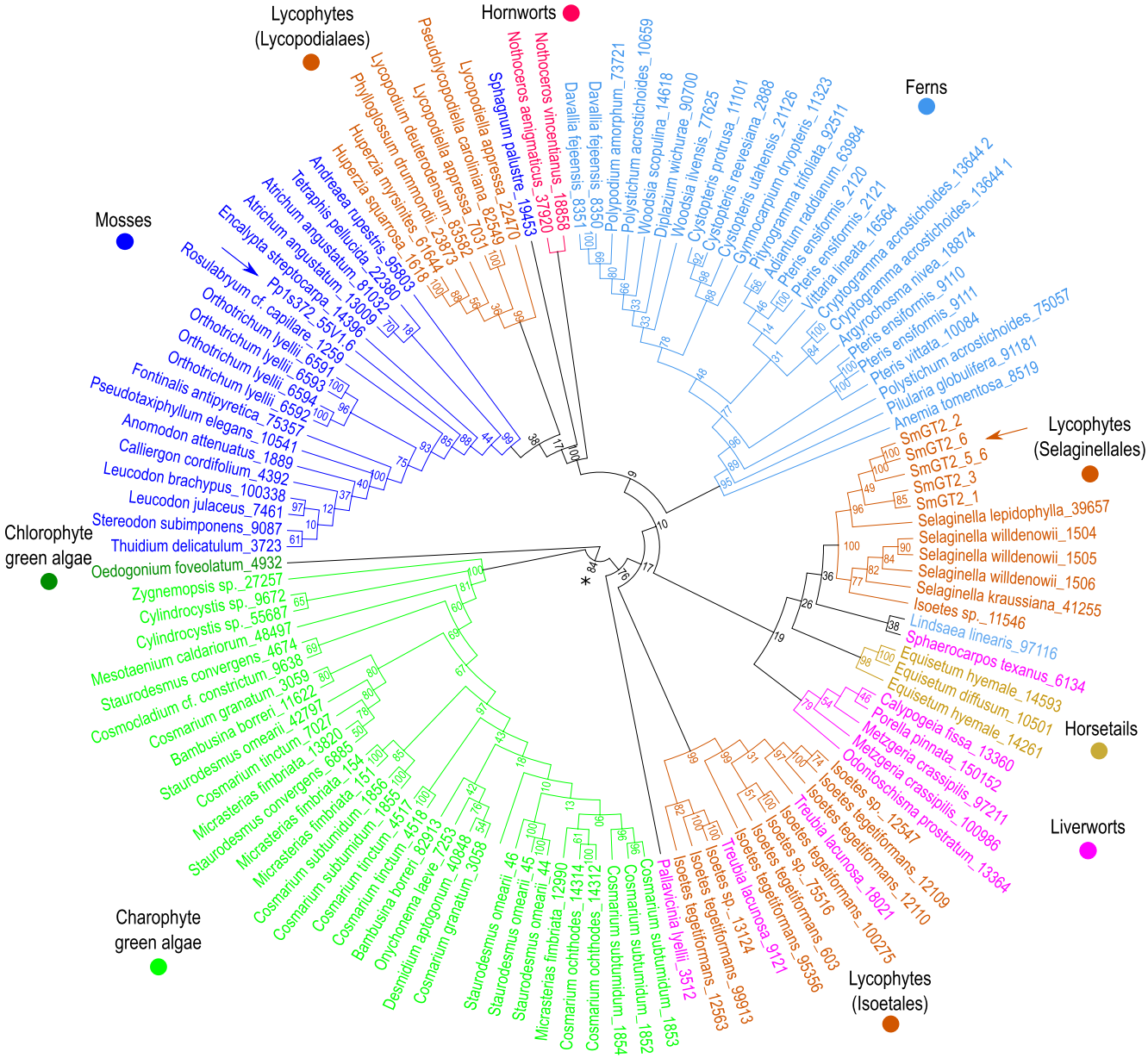

Figure 7. Sequences Similar to Pp3c12_24670 Occur in Diverse Plant Genomes.

Unrooted maximum likelihood tree (1000 bootstrap replicates) of 121 sequences representing a subtree of the cladogram shown in Supplemental Figure 10, realigned and edited before generating the tree. Sequence length of the edited alignment was 390 amino acids. A blue arrow indicates Pp3c12_24670, which is the only functionally characterized sequence in the tree, and a brown arrow indicates *S. moellendorffii* sequences. Nodes are labeled with percent bootstrap values.

Figure 11) include a group comprising CSLAs, CSLCs, and cellulose synthase from *Rhodobacter sphaeroides*; a CESA/CSLD group that also includes the barley MLG synthases; two groups of red algal sequences; four groups of green algal sequences; a group of glaucophyte sequences; a group of prokaryotic sequences; a group of brown algal sequences (Supplemental Figure 11). None of these sequences cluster with Pp3c12_24670, which occurs within the moderately well supported branch marked by an asterisk (Supplemental Figure 11; bootstrap value = 78%).

The cluster that contains Pp3c12_24670 (asterisk in Supplemental Figure 11) also includes the fungal MLG synthase Af-Tft1, other 1KP sequences from the streptophytes (land plants and charophyte green algae), and a single sequence from the chlorophyte green alga *Oedogonium foveolatum*. All other functionally characterized GT2s are excluded, supporting the previous report that Pp3c12_24670 is similar to Af-Tft1 and other fungal sequences (Harholt et al., 2012). This group of 1KP sequences is also supported by a bootstrap value of

86% in the tree of 1KP sequences (asterisk in Supplemental Figure 10) and was used to generate a new alignment (Supplemental File 6). These sequences aligned with fewer gaps and poorly resolved regions, resulting in an alignment that was 390 amino acids long after editing (Supplemental File 7). An unrooted maximum likelihood tree generated from this alignment is similar in topology to the larger tree, but the bootstrap values are higher (Figure 7), as expected for a longer alignment. Green algal sequences are separated from land plant sequences at a well-supported node (asterisk in Figure 7; bootstrap value = 84%), and moss and fern sequences are clustered in separate well-supported clades with a few outliers. The three lycophyte clades (Isoetales, Selaginellales, and Lycopodiales) are not united. Whereas the Lycopodiales sequences form a well-supported clade with the moss and hornwort sequences (100% bootstrap support), the relationships among the Selaginellales, Isoetales, fern, liverwort, and horsetail sequences are not resolved. The *Treubia lacunosa* sample has been found to be contaminated with *Isoetes durieui* DNA (https://pods.iplantcollaborative.org/wiki/display/iptol/Sample+source+and+purity), likely explaining the clustering of these sequences with Isoetales sequences. Although lack of direct functional evidence for all sequences except Pp3c12_24670 limits interpretation, this analysis provides a basis for selecting genes for further functional testing. We can speculate that the moss, hornwort, and Lycopodiales sequences may encode AGlc synthases, whereas the sequences from organisms in which MLG has been detected (horsetails, Selaginellales, and charophyte green algae) may encode MLG synthases.

DISCUSSION

Pp3c12_24670 Synthesizes a Novel Cell Wall AGlc Polysaccharide

When expressed in *N. benthamiana* or overexpressed in *P. patens*, Pp3c12_24670 synthesizes an endo-(1,4)-β-D-glucanohydrolase-sensitive linear polysaccharide composed of unbranched, unsubstituted (1,4)-Glc_p and (1,3)-Ara_f residues and can thus be named an arabinoglucan synthase (AGlcS). The presence of single, nonterminal Ara_f residues in the oligosaccharide released by E-CELAN hydrolysis of AIR cell wall preparations from *Pp3c12_24670*-transformed *N. benthamiana* indicates that (1,3)-Ara_f residues are not adjacent. Although the activity of E-CELAN against AGlc has not been characterized, the release of consistent ratios of tetra- and pentasaccharides, even after overnight incubation, indicates that (1,3)-Ara_f residues are probably separated by two or more (1,4)-β-Glc_p residues. A Glc_p:Ara_f ratio of 7.5:1 indicates that the polymer synthesized in transformed *N. benthamiana* includes longer stretches of (1,4)-β-Glc_p residues. These would be degraded by E-CELAN to cellobiose (G4G) and Glc (which is not retained on the graphitized carbon columns used to extract oligosaccharides from the digests prior to HPAEC-PAD analysis). Both cellobiose and Glc are also products of E-CELAN hydrolysis of cellulose. Thus, the available data do not define the distribution of these additional (1,4)-β-Glc_p residues within the polymer. These uncertainties notwithstanding, AGlc appears to be structurally similar to MLG, in which (1,3)-β-Glc_p residues are separated by two, three, or occasionally more (1,4)-β-Glc_p residues. In MLGs from grasses, blocks of two or three (1,4)-β-linkages are randomly arranged (Staudte et al., 1983), and longer stretches of (1,4)-β-linkages also occur (Woodward et al., 1983). The proportion and arrangement of (1,3)-β-Glc_p residues varies by species, influences solubility, and contributes to the functional versatility of MLG (Burton et al., 2010). The same could be true for AGlc.

The polymer synthesized by AGlcS in *N. benthamiana* is likely polydisperse (varying in size), with an average molecular mass less that 30 kD based on elution after the 33.6 kD standard in SEC (Supplemental Figure 6). FB28 binding is consistent with a molecular mass greater than 10 kD (Figure 2B; Rieder et al., 2015), whereas linkage analysis indicates an average degree of polymerization of ~30 and a calculated molecular mass of 4.7 kD (Table 1). We speculate that the polymer synthesized in this heterologous system differs in size and/or structure from the endogenous *P. patens* polysaccharide, which could not be extracted effectively with CDTA. It is also possible that the endogenous polymer is covalently associated with a protein or some other molecule.

When Pp3c12_24670 and similar *S. moellendorffii* sequences were first reported, similarity to cyanobacterial, red algal, and fungal sequences was noted (Harholt et al., 2012). These sequences were later referred to as either "bacterial-type CESAs" (Mikkelsen et al., 2014) or as "linear TC forming CESAs" or "linear bacterial-like cellulose synthases" based on the observation that some red algae have linear cellulose synthesis complexes, which are also known as terminal complexes (Fangel et al., 2012; Ulvskov et al., 2013). However, whether Pp3c12_24670 synthesizes cellulose or forms linear complexes has been unclear.

Although it appears that E-CELAN releases AGlc-derived oligosaccharides from never-dried AIR preparations from wild-type *P. patens* (Figures 4C and 4D), the yield and purity were not sufficient for further analysis. Thus, AGlc appears to be a minor component of wild-type *P. patens* cell walls. Knockout mutants had no obvious morphological defects. Given that the *P. patens* genome contains no paralogs of *AGlcS*, this suggests that AGlc is not required for normal vegetative development. This is consistent with the observation that AGlc is only expressed in maturing tissues in which cell division and cell expansion are complete (Figure 3). Functions in reproduction, storage, or response to environmental stress remain possible. However, further analysis of *P. patens* knockout mutants and the phylogenetic distribution of AGlc are required to understand its biological role.

AGlc polysaccharides identified previously in plants include a linear polysaccharide from *Cordia dichotoma* fruit consisting of (1,6)-Glc_p and (1,2)-Ara_f residues (Basu et al., 1984) and a linear polysaccharide from the bark of Chinaberry (*Melia azadirachta*) that contains (1,4)-α-Glc_p and (1,6)-α-Ara_f residues (Fujiwara et al., 1985). Water-soluble polysaccharides containing Ara and Glc have been isolated from *Cassia* sp seeds (Huang et al., 2012) and canola meal (Gattinger and Duvnjak, 1990) but have not been further characterized. Consistent with the absence of (1,3)-β-Glc_p residues, lichenase did not release AGlc oligosaccharides from cell walls containing AGlc (Figure 1)

or from the isolated polysaccharide. Thus, AGlc differs from the lichenase-sensitive Ara- and Glc-containing polysaccharide detected in the leafy liverwort *L. bidentata* (Popper and Fry, 2003). This indicates that the Ara- and Glc-containing oligosaccharides released by lichenase (10 units/mL) from a leafy liverwort (Popper and Fry, 2003) are from a different polysaccharide, perhaps one containing both (1,3)-β-Glc and (1,4)-β-Glc residues, and lightly substituted with Ara. These data indicate that *P. patens* AGlc is a novel plant cell wall polysaccharide.

AGlcS Shares Characteristics with MLG Synthases and Cellulose Synthases

Consistent with its role in the synthesis of a cell wall polysaccharide containing (1,4)-β-Glc, the *AGlcS* gene was originally identified in a search of the *P. patens* genome sequence using a red algal cellulose synthase as a query (Harholt et al., 2012). The encoded protein contains a Pfam glyco_transf_2 _3 domain, placing it in the GT2 family with CESAs and CSLs. Phylogenetic analysis (Supplemental Figure 11) showed that it shares the highest similarity with the fungal MLG synthase Af-Tft1 (Samar et al., 2015) among functionally characterized GT2s, confirming a previous report (Harholt et al., 2012). Like the MLG synthases CSLH1 from barley (Wilson et al., 2015) and CSLF6 from *Brachypodium distachyon* (Kim et al., 2015, 2018), AGlcS appears to reside in the Golgi based on localization of a catalytically active mEGFP:AGlcS fusion protein to intracellular compartments that are dispersed by Brefeldin A (Figure 5). CSLF6 from barley has been localized to the plasma membrane as well as the Golgi (Wilson et al., 2015); our data do not exclude the possibility that small amounts of AGlcS reach the plasma membrane. Although we cannot rule out the possibility that AGlcS works cooperatively with a synthase present in both *P. patens* and *N. benthamiana* to incorporate (1,3)-Ara$_f$ and (1,4)-β-Glc$_p$ residues, the ability to synthesize a similar polymer when expressed in two widely divergent species suggests that AGlcS can synthesize two distinct linkages and control their proportion and spacing, as has been suggested for MLG synthases (Jobling, 2015; Dimitroff et al., 2016). In addition to forming two different linkage types, AGlcS may be able to incorporate two different glycosyl units (4-Glc$_p$ and 3-Ara$_f$), as do hyaluronan synthases (4-GlcUA and 3-GlcNAc; Weigel and DeAngelis, 2007). Comparative structural analysis of the MLG, AGlc, and hyaluronan synthases could provide insight into the regulation of the fine structure of polymers synthesized by these enzymes.

Phylogenetic Distribution and Evolution of AGlcS-Like Sequences

AGlcS and related lycophyte sequences were originally identified in searches of the *S. moellendorffii* and *P. patens* genomes (Harholt et al., 2012), and similar sequences were later identified in the charophyte green alga *Nitella mirabilis* (Mikkelsen et al., 2014). Phylogenetic analysis of protein sequences retrieved from the 1KP transcriptome data collection (Matasci et al., 2014) identified a subset of sequences from green algae, bryophytes, lycophytes, and monilophytes that cluster with AGlcS (Supplemental Figure 10; Figure 7). Although *P. patens* is the only species currently known to produce AGlc, MLG and MLG-like polysaccharides have been detected in the cell walls of members of taxa represented in the tree (Figure 7). These include horsetails (Fry et al., 2008; Sørensen et al., 2008), ferns (Leroux et al., 2015), *S. moellendorffii* (Harholt et al., 2012), and charophyte green algae (Eder et al., 2008; Sørensen et al., 2011). Lichenase-sensitive MLG-like polysaccharides containing Ara and Glc have been detected in the leafy liverwort *L. bidentata* (Popper and Fry, 2003). The synthases that polymerize these MLGs and MLG-like polysaccharides have not been identified. However, the group of sequences that includes AGlcS also clusters with Af-Tft1 (Supplemental Figure 11), a bona fide ascomycete MLG synthase (Samar et al., 2015), raising the possibility that this group includes MLG synthases as well as other AGlcS sequences. The phylogenetic tree of AGlc-like sequences (Figure 7) can guide future efforts to clarify the distribution of AGlc, MLG, and MLG-like polysaccharides and identify their synthases. For example, it can be postulated that the sequences from the horsetails *Equisetum diffusum* and *E. hyemale* represent MLG synthases. The moss sequences, along with the hornwort and Lycopodiales sequences that cluster with them, may represent AGlc synthases. Synthases that produce MLG-like polysaccharides of the type found in *L. bidentata* may be represented by sequences from other members of the Jungermanniales order of liverworts.

The results reported here are consistent with the hypothesis that the capacities to synthesize AGlc and MLG have a common evolutionary history. This is supported by phylogenetic clustering of AGlcS with a bona fide fungal MLG and could be further tested through functional analysis of proteins represented by other sequences in the cluster, especially those from organisms known to produce MLG. Apparent similarities in the distribution of (1,3)- and (1,4)-linkages in MLG and AGlc provide additional support for a common origin of the synthases. A loss of the capacity to synthesize both AGlc and MLG in the angiosperm lineage is indicated by both the lack of *Pp3c12_24670* seed plant orthologs reported previously (Harholt et al., 2012; Yin et al., 2014) and supported here and by the absence of MLG in dicots. The evolution of the CSLH, CSLF, and CSLJ families through diversification and specialization of the CESA superfamily members in the Poales constitutes a later independent origin of MLG (Sørensen et al., 2008, 2011; Burton and Fincher, 2009).

METHODS

Vector Construction

Expression Vectors for Physcomitrella patens

Primers were designed based on the *Pp3c12_24670* genomic sequence downloaded from CoGe (https://genomevolution.org/coge/). The *Pp3c12_24670* coding sequence (intronless) was amplified from genomic DNA isolated as described previously (Roberts et al., 2011) using primers with flanking attB5 and attB2 sites (GT2-attB5/GT2-attB2; Supplemental Table 2) and Phusion polymerase (New England Biolabs) with a 30-s initial denaturation at 98°C; 36 cycles of 98°C for 7 s, 68°C for 20 s, and 72°C for 1.5 min; and a 4-min extension at 72°C. The

amplicon was cloned into pDONR 221 P5-P2 using BP Clonase II as described by the manufacturer (Life Technologies). A catalytically inactive version of *Pp3c12_24670* (D247N/D249N) in pDONR 221 P5-P2 was produced by PCR fusion as described previously (Scavuzzo-Duggan et al., 2015) using primers D247+249N-SF1 and D247+249N-SR1 (Table 1) to introduce the mutations. The entry clones containing wild-type or catalytically inactive *Pp3c12_24670* were sequence verified and fused with a pDONR 221 P1-P5r entry clone containing a triple hemagglutinin (3X-HA) tag in destination vector pTHAct1Gate as described previously (Scavuzzo-Duggan et al., 2015). Similarly, the *Pp3c12_24670* entry clone was fused with a pDONR 221 P1-P5r entry clone containing mEGFP (gift of Magdalena Bezanilla) in pTHAct1Gate.

Knockout Vector for P. patens

The Pp3c12_24670-KO vector was constructed using Gateway Multisite cloning (Life Technologies) as described previously (Roberts et al., 2011). Briefly, sequences flanking the coding region were amplified with primer pairs GT2KOattB1/GT2KOattB4 and GT2KOattB3/GT2KOattB2 (Table 1) and cloned into pDONR 221 P1-P4 and pDONR 221 P3-P2, respectively. The resulting entry clones were inserted into pBHSNRG. The Pp3c12_24670-KO vector was cut with *Bsr*GI for transformation into wild type *P. patens*.

Promoter:Reporter Vector for P. patens

The sequence upstream of the start of the *Pp3c12_24670* coding sequence (1963 bp) was amplified using primers with flanking attB1 and attB5r sites (GT2PROattB5r/GT2PROattB1; Supplemental Table 2) and Phusion polymerase (New England Biolabs) with a 30-s initial denaturation at 98°C; 30 cycles of 98°C for 10 s, 60°C for 45 s, and 72°C for 2 min; and a 5-min extension at 72°C. The product was cloned into pDONR 221 P1-P5r as described previously (Tran and Roberts, 2016) and sequence verified. To construct the *Pp3c12_24670pro*:*GUS* vector, entry clones containing the *Pp3c12_24670* promoter and *GUS* gene were inserted into the si3-pTH-GW destination vector as described previously (Tran and Roberts, 2016) and linearized with *Swa*I (New England Biolabs) for transformation into wild-type *P. patens*.

Expression Vectors for Nicotiana benthamiana

Wild-type and catalytically inactive *Pp3c12_24670* were amplified from coding sequences previously cloned in pDONR 221 P5-P2 (see above) using primers PpGT2_pCR8F and PpGT2_pCR8R (Supplemental Table 2) and Phusion polymerase (New England Biolabs), with PCR conditions described above for amplification of *Pp3c12_24670*, and cloned into pCR8 (Life Technologies). The expression vectors were constructed by inserting the sequence-verified entry clones into pEAQ-HT-DEST1 (GenBank GQ497235.1; Sainsbury et al., 2009) using LR Clonase II as described by the manufacturer (Life Technologies).

Transformation of *N. benthamiana*

The Pp3c12_24670 expression vectors in pEAQ-HT-Dest1, positive control Hv-CSLF6 expression vector in pEAQ-HT-Dest1, and empty pEAQ-HT-Dest1 negative control vector were transformed into *Agrobacterium tumefaciens* strain AGL1, and leaves of *N. benthamiana* were infiltrated, harvested, and dried as described previously (Schreiber et al., 2014).

Transformation of *P. patens*

Wild-type *P. patens* strain Gransden 2011 (Rensing et al., 2008) was transformed with vectors driving the expression of 3XHA:Pp3c12_24670, 3XHA:Pp3c12_24670 inactive, or mEGFP:Pp3c12_24670; a *Pp3c12_24670pro*:*GUS* vector; a *Pp3c12_24670* KO vector; or an empty negative control vector (Scavuzzo-Duggan et al., 2015) and selected for hygromycin resistance as described previously (Roberts et al., 2011). 3XHA:Pp3c12_24670 and 3XHA:Pp3c12_24670 lines were screened for heterologous protein expression by protein gel blot analysis as described previously (Scavuzzo-Duggan et al., 2015). Pp3c12_24670 KO lines were genotyped for 5′ and 3′ integration using primer pairs GT2-5′int/BHRRR and BHRRF/GT2-3′int, respectively. DNA was isolated from stably transformed lines as described previously (Roberts et al., 2011) and amplified with Paq5000 polymerase (Agilent Technologies) with a 2-min denaturation at 94°C; 30 cycles of 94°C for 45 s, 57°C for 45 s, and 72°C for 2 min; and a 5-min extension at 72°C. Deletion of the target sequence was tested using primers GT2-RTF/GT2-RTR and Taq polymerase (New England Biolabs) with a 2-min denaturation at 94°C; 30 cycles of 94°C for 45 s, 57°C for 45 s, and 72°C for 2 min; and a 5-min extension at 72°C.

mEGFP Imaging

P. patens lines stably transformed with an mEGFP:Pp3c12_24670 expression vector were screened for mEGFP fluorescence using a stereomicroscope with fluorescence optics (Leica M165 FC; Leica Microsystems), and three lines were selected for testing of transgene activity (see below). Confocal images of protonema from mEGFP:Pp3c12_24670 lines and YFP:GmMan lines (Furt et al., 2012) were captured using an Olympus Fluoview FV1000 confocal microscope with UIS2 40× oil immersion objective (numerical aperture = 1.3), diode laser (473 nm), filter detection (SDM 560, BF 490–540), and Fluoview software (file version 1.2.6.0, system version 3.1.3.3). In some cases, the Golgi apparatus of protonemal filaments was disrupted by treating cells for 30 min with 10 μg mL^{-1} Brefeldin A in BCDAT medium with 0.1% DMSO or BCDAT medium with 0.1% DMSO and mounted in the same solution for imaging (Ritzenthaler et al., 2002).

GUS Staining

For analysis of promoter activity, homogenized protonemal tissue from *Pp3c12_24670pro*:*GUS* lines was cultured on solid BCD medium overlain with cellophane as described previously (Roberts et al., 2011) and harvested after 5 and 11 d. Gametophores were cultured from tissue clumps on solid BCDAT medium without cellophane and harvested after 14 and 21 d. Harvested tissue was fixed for 15 min in 2.3% paraformaldehyde in 50 mM sodium phosphate buffer, pH 7.0, washed three times, 15 min each, in working solution [25 mM sodium phosphate buffer, pH 7.0, 0.25 mM $K_3Fe(CN)_6$, 0.25 mM $K_4Fe(CN)_6$, 0.25 mM EDTA, and 0.25% (v/v) Triton X 100], incubated for 6 h in working solution containing 0.5 mg mL^{-1} 5-bromo-4-chloro-3-indolyl-β-D-glucuronide and 5% (v/v) methanol, and stored in 70% (v/v) ethanol at 4°C. Tissue was imaged using both a stereomicroscope (Leica M65 FC; Leica Microsystems) and a compound microscope (Olympus BHS) with a Leica BFC310 FX camera (Leica Microsystems).

RT-PCR

RNA was extracted from 100 mg of frozen *N. benthamiana* leaf tissue using a PureYield RNA extraction kit according to the manufacturer's instructions (Promega). cDNA was synthesized using Superscript III reverse transcriptase according to the manufacturer's instruction (Life Technologies) with no RT controls for each sample. cDNA (1 μL) was amplified using primers GT2-RTF/GT2-RTR and Taq polymerase with an initial 3-min denaturation at 94°C; 30 cycles of 94°C for 20 s, 59°C for 30 s, and 74°C for 30 s; and a 5-min extension at 74°C.

Cell Wall Isolation and Enzyme Digestion

To prepare AIR, 5 mg of freeze dried and powdered *N. benthamiana* leaf tissue (Schreiber et al., 2014) or 1 g of 6-d-old protonemal tissue or mature gametophores from *P. patens* ground in liquid nitrogen was extracted 4 times with 2 mL 70% (v/v) ethanol for a total of 2 h at 21°C and heated to 100°C for 10 min to inactivate endogenous enzymes. The remaining AIR was pelleted, but not dried. In some cases, AIR was fractionated with CDTA and NaOH, and the CDTA fraction was cleared of charged polysaccharides by precipitation with CTAB as described previously (Fry, 1988). Lichenase (1,3;1,4)-β-D-4-glucanohydrolase (EC 3.2.1.73, from *Bacillus* sp) hydrolysates were prepared by incubating AIR for 20 min in 1 mL of 20 mM sodium phosphate buffer, pH 6.5, at 90°C; adding 40 units of enzyme (E-LICHN; Megazyme) to the cooled reaction mixture; and incubating 1.5 h at 50°C (Schreiber et al., 2014). For endoglucanase [endo-(1,4)-β-D-glucanohydrolase, EC 3.2.1.4, from *Aspergillus niger* or *Trichoderma longibrachiatum*] digests, AIR or CDTA- or NaOH-soluble fractions were incubated in 800 μL of 20 mM sodium acetate buffer, pH 4.7, containing 5 units of enzyme (E-CELAN or E-CELTR, respectively; Megazyme), 0.5 mg mL^{-1} BSA, and 0.05% (w/v) chlorobutanol for 4 to 20 h at 20°C.

Oligosaccharide Analysis and Fractionation

Oligosaccharides were extracted from the clarified digests by solid-phase extraction on graphitized carbon columns (Bond Elute; Agilent Technologies) and analyzed directly by HPAEC-PAD on a Dionex ICS-5000 (Thermo Fisher Scientific) as described previously (Ermawar et al., 2015). The separation was performed using a Dionex CarboPAC PA-20 (3 × 150 mm) column and a gradient of 1 to 13.3% B over 36 min (eluent A was 0.1 M NaOH and B 0.1 M NaOH with 1 M sodium acetate).

Oligosaccharides were fractionated on a Hypercarb column (5 μm, 100 × 4.6 mm; Thermo Fisher Scientific). The column was maintained at 20°C with a flow rate of 1.0 mL min^{-1}. Separation was performed in a gradient of 2 to 50% eluent B over 18 min, where eluent A was 1 mM ammonium hydroxide and eluent B was 90% acetonitrile. Elution was monitored by UV and evaporative light scattering detection (ELSD) using an Alltech ELSD 800 (Thermo Fisher Scientific) operated at 70°C with a nitrogen pressure of 3 bars and a gain of 8. The eluent flow was split after the UV detector with a ratio of 9 fraction collection to 1 ELSD. Fractions were analyzed by HPAEC-PAD as described above, and those containing target oligosaccharides were subjected to further analysis as follows. Monosaccharides were identified and quantified as 1-phenyl-3-methyl-5-pyrazolone derivatives after hydrolysis with 2 M trifluoroacetic acid for 3 h at 100°C (Comino et al., 2013). The molecular weights were determined using HPLC (GlycanPac AXH1 column) coupled to a Q-Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with an electrospray ionization source. The MS spectra were collected in positive-ion mode. For glycosyl linkage analysis, 40 μg samples were permethylated, depolymerized, reduced, and acetylated; the resulting partially methylated alditol acetates (PMAAs) were analyzed by gas chromatography-mass spectrometry as described previously (York et al., 1985).

Polysaccharide Isolation and Analysis

CDTA extracts of cell walls from *N. benthamiana* transformed with *Pp3c12_24670* or empty vector, which had been cleared of charged polysaccharides with CTAB, were fractionated at 20°C by reverse-phase chromatography using an Agilent 1200 HPLC fitted with a Brownlee Aquapore C18 column (30 × 4.6 mm; 7 μm). Separation was performed in gradients of water (eluent A) and 90% aqueous acetonitrile (eluent B) at a flow rate of 1 mL/min, as follows: 2 to 20% B over 10 min, followed by a washing step in 80% B for 1 min, and equilibration of the column in 2% B. Bound proteins were eluted between separating runs with 70% acetonitrile containing 0.04% trifluoroacetic acid. Elution was monitored by UV detection and ELSD as described above. The eluent flow was split after the UV detector with a 9:1 fraction collection:ELSD. Collected polysaccharide fractions were analyzed by digestion with E-CELTR and HPAEC-PAD analysis as described above. Those fractions found to contain AGlc were combined and characterized by SEC with postcolumn Fluorescent Brightener 28 detection (Mitra et al., 2017) and also hydrolyzed for monosaccharide analysis. The linkage types present in the fractions containing the polysaccharide were determined by methylation analysis according to the method of Ciucanu and Kerek (1984). The freeze-dried samples (around 100 μg for each duplicate) were dissolved in 0.5 mL of anhydrous DMSO, and the polysaccharide was methylated, hydrolyzed, reduced in the presence of $NaBD_4$, and acetylated in acetic anhydride (Ciucanu and Kerek, 1984). The resulting PMAAs were recovered by evaporating the solvent to dryness under a stream of argon, redissolving the sample in dichloromethane, and filtering the solution through a column filled with anhydrous sodium sulfate powder. The resulting filtrate was transferred to a GC vial and analyzed on Hewlett Packard 6890/5973 gas chromatography-mass spectrometry equipment fitted with a SP-2380 capillary column (30 m × 0.25 mm i.d.; Sigma-Aldrich). Helium was used as the carrier gas at a flow rate of 1 mL/min. For each sample run, the oven temperature was programmed to increase from 165°C to 213°C at 1°C/min, from 213°C to 230°C at 3°C/min, and from 230°C to 260°C at 10°C/min, followed by a plateau at 260°C for 10 min (total run time 67 min). The fragmentation electron-impact mass spectra of PMAAs were interpreted by comparison with spectra of reference derivatives and by referring to the literature (Carpita and Shea, 1989). The experiment was conducted in duplicate.

Phylogenetic Analysis

A sequence search was performed on data obtained from the 1KP initiative (http://www.onekp.com/). Predicted amino acid sequences for ~1300 plant/algal samples were used. Briefly, these sequences were created by taking scaffolds greater than 300 bp from the SOAPdenovo-Trans assembly of the 1KP transcriptomes (Xie et al., 2014). The top five hits for protein coding genes from NCBI RefSeq plant sequences (release 54, July 2012) were used to infer protein sequences using GeneWise (Birney et al., 2004) and TransPipes (Barker et al., 2010). To search for relevant proteins, a custom protein BLAST server was created with SequenceServer version 1.0.4 (Priyam et al., 2015). 1KP data include multiple samples from the same species. To reduce redundant sequences, where multiple samples were available, only the predicted proteins based on the combined transcriptome assembly of these samples were included. Otherwise, the single sample predicted proteins were included for a total of 19,567,909 protein sequences. The protein database using these sequences was created using BLAST (version 2.2.30) (Camacho et al., 2009). To aid in identifying and analyzing results, sample information (accessed from http://www.onekp.com/samples/list.php; June 2013) relating to clade, order, family, and species was added to sequence identifiers.

The Pp3c12_24670 peptide sequence was submitted as a query with an e-value cutoff of 5×10^{-12}. Among the top 500 hits, 27 lacked an organism code and the remaining 473 hits were analyzed using Geneious 8.1 (Kearse et al., 2012). The list contained 20 seed plant sequences, and each was tested by TBLASTN search against the GenBank nonredundant nucleotide collection (May 2015; repeated June 2017). All 20 sequences were found to have high identity to CSLA or CSLC sequences (77–100% identify) or were nearly identical to fungal sequences (>99% identity) and were removed from the list. Sequences shorter than 200 amino acids (115 sequences) were removed and remaining sequences were submitted

to InterProScan (Jones et al., 2014) for prediction of Pfam domains. Sequences that lacked a Pfam glyco_transf_2_3 domain of at least 200 bp (89 sequences) were removed and the remaining 243 sequences with Pp3c12_24670 and five *Selaginella moellendorffii* sequences (Harholt et al., 2012) added were aligned using ClustalW with BLOSUM cost matrix, gap open cost of 10, and gap extension cost of 0.1. Selected sequences were realigned as described in Results. The alignments were edited to remove gaps and adjacent poorly aligned segments (Baldauf, 2003), and rapid bootstrapping (1000 replicates) and a search for the best-scoring maximum likelihood tree were performed using RAxML 7.2.8, GAMMA BLOSOM62 protein model. The trees were exported from Geneious 8.1 and edited with Inkscape (https://inkscape.org/en/).

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL databases under the following accession numbers: Pp3c12_24670, PNR44324.1; AtuCrdS, AAD20440.2; RsBcsA, YP_353410.1; *Sinorhizobium meliloti* MLG synthase, SDW48993.1; Af-Tft1, XP_748682; Hv-CSLH1, ACN67534.1; Hv-CSLF6, ABZ01578.1; Pp-CESA3, PNR49373.1; Pp-CESA4, PNR48112.1; Pp-CESA5, PNR59825.1; Pp-CESA6, AAZ86086.1; Pp-CESA7, AAZ86087.1; Pp-CESA8, ABI78961.1; Pp-CESA10, PNR47708.1; Pp-CSLD1, ABI75151.1; Pp-CSLD2, ABI75152.1; Pp-CSLD3, ABI75153.1; Pp-CSLD4, ABI75154.1; Pp-CSLD5, ABI75155.1; Pp-CSLD6, ABI75156.1; Pp-CSLD7, ABI75157.1; Pp-CSLD8, ABI75158.1; Pp-CSLA1, ABD79099.1; Pp-CSLA2, ABD79100.1; Pp-CSLA3, PNR41011.1; Pp-CSLC1, ABI55233.1; Pp-CSLC2, ABI55234.1; Pp-CSLC3, ABI55235.1; Pp-CSLC4, PNR32156.1; Pp-CSLC5, PNR37849.1; Pp-CSLC6, PNR39718.1; Pp-CSLC7, PNR48367.1; SmGT2-1, XP_002991857.1; Sm-GT2-2, XP_002991856.1; Sm-GT2-3, XP_002993016.1; Sm-GT2-6, XP_002993014.1; and Sm-GT2-5-6, XP_002971505.1.

Supplemental Data

Supplemental Figure 1. HPAEC-PAD profiles of oligosaccharides released by E-CELTR from leaf AIR cell wall preparations of *N. benthamiana* expressing *Pp3c12_24670* and empty vector negative control.

Supplemental Figure 2. HPAEC-PAD profiles of an E-CELAN hydrolysate of an AIR cell wall preparation of *N. benthamiana* leaves transformed with *Pp3c12_24670* and two fractions from an HPLC separation of the digest.

Supplemental Figure 3. ESI mass spectra (positive ion mode) of the 19.5 and 22.5 min peaks from an E-CELAN hydrolysate of an AIR cell wall preparation of *N. benthamiana* leaves expressing *Pp3c12_24670*, isolated by HPLC.

Supplemental Figure 4. HPAEC-PAD profiles of E-CELAN hydrolysates of sequential CDTA and 4 M NaOH extracts, and residue from NaOH extraction of leaf AIR cell wall preparations from *N. benthamiana* transformed with *Pp3c12_24670*.

Supplemental Figure 5. HPAEC-PAD profile of an E-CELTR digest of fraction #10 from the RP-HPLC separation (shown in Figure 2A) of the CDTA extract of the AIR cell wall preparation from *N. benthamiana* leaves transformed with *Pp3c12_24670*.

Supplemental Figure 6. SEC analysis of supernatants from CTAB precipitation of CDTA extracts of AIR cell wall preparations from *N. benthamiana* leaves transformed with *Pp3c12_24670*, with post-column FB-28 and fluorescence detection.

Supplemental Figure 7. HPAEC-PAD profiles of E-CELTR hydrolysates of sequential water (black) and guanidinium HCl extracts of leaf AIR cell wall preparations from *N. benthamiana* transformed with *Pp3c12_24670*.

Supplemental Figure 8. PCR genotype analysis of *P. patens* lines stably transformed with a *Pp3c12_24670* knockout vector.

Supplemental Figure 9. MUSCLE alignment of Pp3c12_24670 with MLG synthases including Hv-CSLF6 from barley, Hv-CSLH1from barley, Af-Tft1 *Aspergillus fumigatus*, and SDW48993.1 from *Sinorhizobium meliloti*.

Supplemental Figure 10. Unrooted maximum likelihood tree of 243 sequences retrieved from the 1KP databased on similarity to Pp3c12_24670, with Pp3c12_24670 and *S. moellendorffii* sequences identified previously.

Supplemental Figure 11. Unrooted maximum likelihood tree of selected sequences from the tree in Supplemental Figure 10 with all members of the *P. patens* CESA/CSL superfamily, functionally characterized MLG synthases, and prokaryotic curdlan and cellulose synthases.

Supplemental Table 1. Linkage analysis by methylation of the 19.5 min peak isolated by HPLC (Supplemental Figure 2) from an E-CELAN hydrolysate of an AIR cell wall preparation of *N. benthamiana* leaves expressing *Pp3c12_24670*.

Supplemental Table 2. Oligonucleotide primers used in this study.

Supplemental File 1. FASTA file of 473 sequences identified in a search of the 1KP data using Pp3c12_24670 as a query.

Supplemental File 2. Unedited alignment of 249 sequences identified in a search of the 1KP data using Pp3c12_24670 as a query.

Supplemental File 3. Edited alignment of 249 sequences identified in a search of the 1KP data using Pp3c12_24670 as a query.

Supplemental File 4. Unedited alignment of representative sequences from the 1KP data with Pp3c12_24670, *P. patens* CESA/CSL sequences, and functionally characterized MLG and cellulose synthases.

Supplemental File 5. Edited alignment of representative sequences from the 1KP data with Pp3c12_24670, *P. patens* CESA/CSL sequences, and functionally characterized MLG and cellulose synthases.

Supplemental File 6. Unedited alignment of 121 sequences from the subtree containing Pp3c12_24670.

Supplemental File 7. Edited alignment of 121 sequences from the subtree containing Pp3c12_24670.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation Award IOS-1257047 and the Australian Research Council Centre for Excellence in Plant Cell Walls Grant CE1101007. DNA sequencing was conducted at a Rhode Island NSF EPSCoR research facility, the Genomics and Sequencing Center, supported in part by the National Science Foundation EPSCoR Cooperative Agreement EPS-1004057. Oligosaccharide linkage analysis was supported by the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, U.S. Department of Energy Grant DE-FG02-93ER20097 to Parastoo Azadi at the Complex Carbohydrate Research Center. We thank Luis Vidali for his gift of the YFP:GmMan Golgi marker line of *P. patens* and Magdalena Bezanilla for the gift of the mEGFP entry clone.

AUTHOR CONTRIBUTIONS

A.W.R., J.L., G.D., and V.B. designed the research. A.W.R., J.L., Y.S.Y.H., X.X., K.Y., A.M.C., T.R.S.-D., G.D., and E.R. performed the research. A.L. contributed new computational tools. A.W.R., J.L., Y.S.Y.H., X.X., A.M.C., T.R.S.-D., E.R., V.B., G.B.F., M.S.D., A.B., and R.A.B. analyzed data. A.W.R., J.L., A.L., V.B., M.S.D., and R.A.B. wrote the article.

Received February 2, 2018; revised March 27, 2018; accepted April 17, 2018; published April 19, 2018.

REFERENCES

Anderson, M.A., and Stone, B.A. (1975). A new substrate for investigating the specificity of β-glucan hydrolases. FEBS Lett. **52:** 202–207.

Baldauf, S.L. (2003). Phylogeny for the faint of heart: a tutorial. Trends Genet. **19:** 345–351.

Barker, M.S., Dlugosch, K.M., Dinh, L., Challa, R.S., Kane, N.C., King, M.G., and Rieseberg, L.H. (2010). EvoPipes.net: bioinformatic tools for ecological and evolutionary genomics. Evol. Bioinform. Online **6:** 143–149.

Basu, N.G., Ghosal, P.K., and Thakur, S. (1984). Structural studies on a polysaccharide fraction from the fruits of *Cordia dichotoma* Forst. Carbohydr. Res. **131:** 149–155.

Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and Genomewise. Genome Res. **14:** 988–995.

Burton, R.A., and Fincher, G.B. (2009). (1,3;1,4)-Beta-D-glucans in cell walls of the poaceae, lower plants, and fungi: a tale of two linkages. Mol. Plant **2:** 873–882.

Burton, R.A., Wilson, S.M., Hrmova, M., Harvey, A.J., Shirley, N.J., Medhurst, A., Stone, B.A., Newbigin, E.J., Bacic, A., and Fincher, G.B. (2006). Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)-beta-D-glucans. Science **311:** 1940–1942.

Burton, R.A., Gidley, M.J., and Fincher, G.B. (2010). Heterogeneity in the chemistry, structure and function of plant cell walls. Nat. Chem. Biol. **6:** 724–732.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics **10:** 421.

Carpita, N.C., and Shea, E.M. (1989). Linkage structure of carbohydrates by gas chromatography-mass spectrometry (GC-MS) of partially methylated alditol acetates. In Analysis of Carbohydrates by GLC and MS, C.J. Biermann and G.D. McGinnis, eds (Boca Raton, FL: CRC Press), pp. 157–216.

Ciucanu, I., and Kerek, F. (1984). A simple and rapid method for the permethylation of carbohydrates. Carbohydr. Res. **131:** 209–217.

Comino, P., Shelat, K., Collins, H., Lahnstein, J., and Gidley, M.J. (2013). Separation and purification of soluble polymers and cell wall fractions from wheat, rye and hull less barley endosperm flours for structure-nutrition studies. J. Agric. Food Chem. **61:** 12111–12122.

Cove, D.J., Perroud, P.F., Charron, A.J., McDaniel, S.F., Khandelwal, A., and Quatrano, R.S. (2009). The moss *Physcomitrella patens*. A novel model system for plant development and genomic studies. In Emerging Model Organisms: A Laboratory Manual, R.R. Behringer, A.D. Johnson, and R.E. Krumlauf, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 69–104.

Dimitroff, G., Little, A., Lahnstein, J., Schwerdt, J.G., Srivastava, V., Bulone, V., Burton, R.A., and Fincher, G.B. (2016). (1,3;1,4)-β-Glucan biosynthesis by the CSLF6 enzyme: Position and flexibility of catalytic residues influence product fine structure. Biochemistry **55:** 2054–2061.

Doblin, M.S., Pettolino, F.A., Wilson, S.M., Campbell, R., Burton, R.A., Fincher, G.B., Newbigin, E., and Bacic, A. (2009). A barley cellulose synthase-like CSLH gene mediates (1,3;1,4)-beta-D-glucan synthesis in transgenic Arabidopsis. Proc. Natl. Acad. Sci. USA **106:** 5996–6001.

Eder, M., Tenhaken, R., Driouich, A., and Lütz-Meindl, U. (2008). Occurrence and characterization of arabinogalactan-like proteins and hemicelluloses in *Micrasterias* (Streptophyta). J. Phycol. **44:** 1221–1234.

Ermawar, R.A., Collins, H.M., Byrt, C.S., Betts, N.S., Henderson, M., Shirley, N.J., Schwerdt, J., Lahnstein, J., Fincher, G.B., and Burton, R.A. (2015). Distribution, structure and biosynthetic gene families of (1,3;1,4)-β-glucan in *Sorghum bicolor*. J. Integr. Plant Biol. **57:** 429–445.

Fangel, J.U., Ulvskov, P., Knox, J.P., Mikkelsen, M.D., Harholt, J., Popper, Z.A., and Willats, W.G. (2012). Cell wall evolution and diversity. Front. Plant Sci. **3:** 152.

Fincher, G.B. (2009). Revolutionary times in our understanding of cell wall biosynthesis and remodeling in the grasses. Plant Physiol. **149:** 27–37.

Fincher, G.B. (2016). Cereals: chemistry and physicochemistry of non-starchy polysaccharides. In Encyclopedia of Food Grains, C.W. Wrigley, H. Corke, K. Seetharaman, and J. Faubion, eds (New York: Academic Press), pp. 208–223.

Fontaine, T., Simenel, C., Dubreucq, G., Adam, O., Delepierre, M., Lemoine, J., Vorgias, C.E., Diaquin, M., and Latgé, J.P. (2000). Molecular organization of the alkali-insoluble fraction of *Aspergillus fumigatus* cell wall. J. Biol. Chem. **275:** 27594–27607.

Fry, S.C. (1988). The Growing Plant Cell Wall: Chemical and Metabolic Analysis. (Essex, UK: Longman).

Fry, S.C., Nesselrode, B.H., Miller, J.G., and Mewburn, B.R. (2008). Mixed-linkage (1→3,1→4)-beta-D-glucan is a major hemicellulose of Equisetum (horsetail) cell walls. New Phytol. **179:** 104–115.

Fujiwara, T, Takeda, T, and Ogihara, Y. (1985). Synthesis of trisaccharides related to an arabinoglucan. Carbohydr. Res. **141:** 168–171.

Furt, F., Lemoi, K., Tüzel, E., and Vidali, L. (2012). Quantitative analysis of organelle distribution and dynamics in *Physcomitrella patens* protonemal cells. BMC Plant Biol. **12:** 70.

Gattinger, L.D., and Duvnjak, Z. (1990). Enzymatic saccharification of canola meal. J. Chem. Technol. Biotechnol. **49:** 155–164.

Gorin, P., Baron, M., and Iacomini, M. (1988). Storage products of lichens. In Handbook of Lichenology, M. Galun, ed (Boca Raton, FL: CRC Press), pp. 9–23.

Harholt, J., Sørensen, I., Fangel, J., Roberts, A., Willats, W.G.T., Scheller, H.V., Petersen, B.L., Banks, J.A., and Ulvskov, P. (2012). The glycosyltransferase repertoire of the spikemoss *Selaginella moellendorffii* and a comparative study of its cell wall. PLoS One **7:** e35846.

Harris, P.J., and Fincher, G.B. (2009). Distribution, fine structure and function of (1,3;1,4)-β-glucans in the grasses and other taxa. In Chemistry, Biochemistry, and Biology of (1-3)-β-Glucans and Related Polysaccharides, A. Bacic, G.B. Fincher, and B.A. Stone, eds (New York: Academic Press/Elsevier), pp. 621–654.

Honegger, R., and Haisch, A. (2001). Immunocytochemical location of the (1→3) (1→4)-β-glucan lichenin in the lichen-forming ascomycete *Cetraria islandica* (Icelandic moss). New Phytol. **150:** 739–746.

Horstmann, V., Huether, C.M., Jost, W., Reski, R., and Decker, E.L. (2004). Quantitative promoter analysis in Physcomitrella patens: a set of plant vectors activating gene expression within three orders of magnitude. BMC Biotechnol. **4:** 13.

Huang, Y.-L., Chow, C.-J., and Tsai, Y.-H. (2012). Composition, characteristics, and in-vitro physiological effects of the water-soluble polysaccharides from Cassia seed. Food Chem. **134:** 1967–1972.

Jobling, S.A. (2015). Membrane pore architecture of the CslF6 protein controls (1-3,1-4)-β-glucan structure. Sci. Adv. **1:** e1500069.

Jones, P., et al. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics **30:** 1236–1240.

Kearse, M., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28:** 1647–1649.

Kim, S.J., Zemelis, S., Keegstra, K., and Brandizzi, F. (2015). The cytoplasmic localization of the catalytic site of CSLF6 supports a

channeling model for the biosynthesis of mixed-linkage glucan. Plant J. **81:** 537–547.

Kim, S.J., Zemelis-Durfee, S., Jensen, J.K., Wilkerson, C.G., Keegstra, K., and Brandizzi, F. (2018). In the grass species *Brachypodium distachyon*, the production of mixed-linkage (1,3;1,4)-β-glucan (MLG) occurs in the Golgi apparatus. Plant J. **93:** 1062–1075.

Lechat, H., Amat, M., and Mazoyer, J. (2000). Structure and distribution of glucomannan and sulfated glucan in the cell walls of the red alga *Kappaphycus alvarezii* (Gigartinales, Rhodophyta). J. Phycol. **36:** 891–902.

Leroux, O., Sørensen, I., Marcus, S.E., Viane, R.L., Willats, W.G., and Knox, J.P. (2015). Antibody-based screening of cell wall matrix glycans in ferns reveals taxon, tissue and cell-type specific distribution patterns. BMC Plant Biol. **15:** 56.

Matasci, N., et al. (2014). Data access for the 1,000 Plants (1KP) project. Gigascience **3:** 17.

McCleary, B.V. (2008). Measurement of dietary fibre components: the importance of enzyme purity, activity and specificity. In Advanced Dietary Fibre Technology, B.V. McCleary and L. Prosky, eds (Oxford, UK: Blackwell Science), pp. 89–120.

McCleary, B.V., and Codd, R. (1991). Measurement of (1-3,1-4)-β-D-glucan in barley and oats: a streamlined enzymatic procedure. J. Sci. Food Agric. **55:** 303–312.

McElroy, D., Zhang, W., Cao, J., and Wu, R. (1990). Isolation of an efficient actin promoter for use in rice transformation. Plant Cell **2:** 163–171.

Meikle, P.J., Hoogenraad, N.J., Bonig, I., Clarke, A.E., and Stone, B.A. (1994). A (1→3,1→4)-beta-glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1→3,1→4)-beta-glucans. Plant J. **5:** 1–9.

Mikkelsen, M.D., Harholt, J., Ulvskov, P., Johansen, I.E., Fangel, J.U., Doblin, M.S., Bacic, A., and Willats, W.G. (2014). Evidence for land plant cell wall biosynthetic mechanisms in charophyte green algae. Ann. Bot. **114:** 1217–1236.

Mitra, S., Lahnstein, J., James, A.P., Fenton, H.K., Burton, R.A., Cato, L., and Solah, V. (2017). The effect of processing on (1,3;1,4)-β-glucan viscosity and molecular weight of Western Australian oat cultivars. Cereal Chem. **94:** 625–632.

Moscatelli, E.A., Ham, E.A., and Rickes, E.L. (1961). Enzymatic properties of a beta-glucanase from *Bacillus subtilis*. J. Biol. Chem. **236:**2858–2862.

Olafsdottir, E.S., and Ingólfsdottir, K. (2001). Polysaccharides from lichens: structural characteristics and biological activity. Planta Med. **67:** 199–208.

Pacheco-Sanchez, M., Boutin, Y., Angers, P., Gosselin, A., and Tweddell, R.J. (2006). A bioactive (1→3)-, (1→4)-beta-D-glucan from *Collybia dryophila* and other mushrooms. Mycologia **98:** 180–185.

Pérez-Mendoza, D., Rodríguez-Carvajal, M.A., Romero-Jiménez, L., Farias, Gde.A., Lloret, J., Gallegos, M.T., and Sanjuán, J. (2015). Novel mixed-linkage β-glucan activated by c-di-GMP in *Sinorhizobium meliloti*. Proc. Natl. Acad. Sci. USA **112:** E757–E765.

Pettolino, F., Sasaki, I., Turbic, A., Wilson, S.M., Bacic, A., Hrmova, M., and Fincher, G.B. (2009). Hyphal cell walls from the plant pathogen Rhynchosporium secalis contain (1,3/1,6)-β-D-glucans, galacto- and rhamnomannans, (1,3;1,4)-β-D-glucans and chitin. FEBS J. **276:** 3698–3709.

Popper, Z.A., and Fry, S.C. (2003). Primary cell wall composition of bryophytes and charophytes. Ann. Bot. **91:** 1–12.

Popper, Z.A., Michel, G., Hervé, C., Domozych, D.S., Willats, W.G., Tuohy, M.G., Kloareg, B., and Stengel, D.B. (2011). Evolution and diversity of plant cell walls: from algae to flowering plants. Annu. Rev. Plant Biol. **62:** 567–590.

Priyam, A., Woodcroft, B.J., Rai, V., Munagala, A., Moghul, I., Ter, F., Gibbins, M.A., Moon, H., Leonard, G., Rumpf, W., and Wurm, Y. (2015). SequenceServer: a modern graphical user interface for custom BLAST databases. bioRxiv doi/10.1101/033142.

Rensing, S.A., et al. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. Science **319:** 64–69.

Rieder, A., Knutsen, S.H., Ulset, A.-S.T., Christensen, B.E., Andersson, R., Mikkelson, A., Tuomainen, P., Maina, N., and Ballance, S. (2015). Inter-laboratory evaluation of SEC-post-column calcofluor for determination of the weight-average molar mass of cereal β-glucan. Carbohydr. Polym. **124:** 254–264.

Ritzenthaler, C., Nebenführ, A., Movafeghi, A., Stussi-Garaud, C., Behnia, L., Pimpl, P., Staehelin, L.A., and Robinson, D.G. (2002). Reevaluation of the effects of brefeldin A on plant cells using tobacco Bright Yellow 2 cells expressing Golgi-targeted green fluorescent protein and COPI antisera. Plant Cell **14:** 237–261.

Roberts, A.W., and Bushoven, J.T. (2007). The cellulose synthase (*CESA*) gene superfamily of the moss *Physcomitrella patens.* Plant Mol. Biol. **63:** 207–219.

Roberts, A.W., Dimos, C.S., Budziszek, M.J., Jr., Goss, C.A., and Lai, V. (2011). Knocking out the wall: protocols for gene targeting in *Physcomitrella patens.* Methods Mol. Biol. **715:** 273–290.

Sainsbury, F., Thuenemann, E.C., and Lomonossoff, G.P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol. J. **7:** 682–693.

Salmeán, A.A., Duffieux, D., Harholt, J., Qin, F., Michel, G., Czjzek, M., Willats, W.G.T., and Hervé, C. (2017). Insoluble (1→3), (1→4)-β-D-glucan is a component of cell walls in brown algae (Phaeophyceae) and is masked by alginates in tissues. Sci. Rep. **7:** 2880.

Samar, D., Kieler, J.B., and Klutts, J.S. (2015). Identification and deletion of Tft1, a predicted glycosyltransferase necessary for cell wall β-1,3;1,4-glucan synthesis in *Aspergillus fumigatus.* PLoS One **10:** e0117336.

Scavuzzo-Duggan, T.R., Chaves, A.M., and Roberts, A.W. (2015). A complementation assay for in vivo protein structure/function analysis in *Physcomitrella patens* (Funariaceae). Appl. Plant Sci. **3:** 1500023.

Schaefer, D.G. (2002). A new moss genetics: targeted mutagenesis in *Physcomitrella patens.* Annu. Rev. Plant Biol. **53:** 477–501.

Schreiber, M., Wright, F., MacKenzie, K., Hedley, P.E., Schwerdt, J.G., Little, A., Burton, R.A., Fincher, G.B., Marshall, D., Waugh, R., and Halpin, C. (2014). The barley genome sequence assembly reveals three additional members of the CslF (1,3;1,4)-β-glucan synthase gene family. PLoS One **9:** e90888.

Sørensen, I., Pettolino, F.A., Wilson, S.M., Doblin, M.S., Johansen, B., Bacic, A., and Willats, W.G. (2008). Mixed-linkage (1→3),(1→4)-β-D-glucan is not unique to the Poales and is an abundant component of *Equisetum arvense* cell walls. Plant J. **54:** 510–521.

Sørensen, I., Pettolino, F.A., Bacic, A., Ralph, J., Lu, F., O'Neill, M.A., Fei, Z., Rose, J.K., Domozych, D.S., and Willats, W.G. (2011). The charophycean green algae provide insights into the early origins of plant cell walls. Plant J. **68:** 201–211.

Staudte, R.G., Woodward, J.R., Fincher, G.B., and Stone, B.A. (1983). Water-soluble (1→3;1→4)-β-D-glucans from barley (*Hordeum vulgare*) endosperm. III. Distribution of cellotriosyl and cellotetraosyl residues. Carbohydr. Polym. **3:** 299–312.

Tran, M.L., and Roberts, A.W. (2016). Cellulose synthase gene expression profiling of *Physcomitrella patens*. Plant Biol (Stuttg.) **18:** 362–368.

Ulvskov, P., Paiva, D.S., Domozych, D., and Harholt, J. (2013). Classification, naming and evolutionary history of glycosyltransferases from sequenced green and red algal genomes. PLoS One **8:** e76511.

Weigel, P.H., and DeAngelis, P.L. (2007). Hyaluronan synthases: a decade-plus of novel glycosyltransferases. J. Biol. Chem. **282:** 36777–36781.

Wilson, S.M., Ho, Y.Y., Lampugnani, E.R., Van de Meene, A.M., Bain, M.P., Bacic, A., and Doblin, M.S. (2015). Determining the subcellular location of synthesis and assembly of the cell wall polysaccharide (1,3; 1,4)-β-D-glucan in grasses. Plant Cell **27:** 754–771.

Woodward, J.R., Fincher, G.B., and Stone, B. (1983). Water-soluble (1–3,1–4)-b-D-glucans from barley (*Hordeum vulgare*) endosperm. II. Fine structure. Carbohydr. Polym. **3:** 207–225.

Xie, Y., et al. (2014). SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. Bioinformatics **30:** 1660–1666.

Yin, Y., Johns, M.A., Cao, H., and Rupani, M. (2014). A survey of plant and algal genomes and transcriptomes reveals new insights into the evolution and function of the cellulose synthase superfamily. BMC Genomics **15:** 260.

York, W.S., Darvill, A.G., McNeil, M., Stevenson, T.T., and Albersheim, P. (1985). Isolation and characterization of cell walls and cell wall components. Methods Enzymol. **118:** 3–40.

Zimmer, A.D., Lang, D., Buchta, K., Rombauts, S., Nishiyama, T., Hasebe, M., Van de Peer, Y., Rensing, S.A., and Reski, R. (2013). Reannotation and extended community resources for the genome of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions. BMC Genomics **14:** 498.