



Evolution of the Cell Wall Gene Families of Grasses

Bryan W. Penning¹, Maureen C. McCann^{2,3} and Nicholas C. Carpita^{2,3,4*}

¹ Corn, Soybean and Wheat Quality Research, USDA-ARS, Wooster, OH, United States, ² Department of Biological Sciences, Purdue University, West Lafayette, IN, United States, ³ Purdue Center for Plant Biology, West Lafayette, IN, United States, ⁴ Department of Botany & Plant Pathology, Purdue University, West Lafayette, IN, United States

Grasses and related commelinid monocot species synthesize cell walls distinct in composition from other angiosperm species. With few exceptions, the genomes of all angiosperms contain the genes that encode the enzymes for synthesis of all cell-wall polysaccharide, phenylpropanoid, and protein constituents known in vascular plants. RNA-seq analysis of transcripts expressed during development of the upper and lower internodes of maize (*Zea mays*) stem captured the expression of cell-wall-related genes associated with primary or secondary wall formation. High levels of transcript abundances were not confined to genes associated with the distinct walls of grasses but also of those associated with xyloglucan and pectin synthesis. Combined with proteomics data to confirm that expressed genes are translated, we propose that the distinctive cell-wall composition of grasses results from sorting downstream from their sites of synthesis in the Golgi apparatus and hydrolysis of the uncharacteristic polysaccharides and not from differential expression of synthases of grass-specific polysaccharides.

OPEN ACCESS

Edited by:

Elisabeth Jamet, Université Toulouse III Paul Sabatier, France

Reviewed by:

Armand Seguin, Canadian Forest Service, Canada Geoff Fincher, University of Adelaide, Australia Simon Hawkins, Lille University of Science and Technology, France

*Correspondence:

Nicholas C. Carpita carpita@purdue.edu

Specialty section:

This article was submitted to Plant Physiology, a section of the journal Frontiers in Plant Science

Received: 06 May 2019 Accepted: 02 September 2019 Published: 04 October 2019

Citation:

Penning BW, McCann MC and Carpita NC (2019) Evolution of the Cell Wall Gene Families of Grasses. Front. Plant Sci. 10:1205. doi: 10.3389/fpls.2019.01205 Keywords: maize (Zea mays), cell-wall structure, cell-wall synthesis, type II cell walls, gene annotation, gene expression, stem development

INTRODUCTION

The primary walls of all angiosperms are assembled from scaffolds of cellulose microfibrils interlaced with hemicellulosic xyloglucans (XyGs), (glucuronoarabino)xylans (GAXs), and (gluco)mannans (Scheller and Ulvskov, 2010), and embedded in an independent but coextensive matrix of pectic polysaccharides (McCann and Roberts, 1991; Caffall and Mohnen, 2009). Primary cell walls of angiosperm species are classified into two types (Carpita and Gibeaut, 1993). XyGs are the major hemicellulose in the type I walls of dicots and noncommelinid monocots and contain an abundance of HG and RG-I pectic polysaccharides (Caffall and Mohnen, 2009; Scheller and Ulvskov, 2010). In contrast, type II walls of grasses and other commelinid monocots contain mostly hemicellulosic GAX of varying degrees of side branching but have little pectin (Carpita and Gibeaut, 1993). A $(1\rightarrow 3),(1\rightarrow 4)-\beta$ -D-glucan (mixedlinkage β -glucan) is found in *Poales* at certain stages of primary wall formation. Small amounts of XyG are also found with side chains terminated by galactose and no detectable Fuc (Carpita et al., 2001). Type I walls incorporate predominantly Hyp- and Gly-rich structural proteins at the end of growth to reinforce the wall into final shape, whereas this reinforcement is supplied by a phenylpropanoid network in the type II wall (Carpita, 1996). This distinction results in strong autofluorescence in the primary wall in addition to that seen in vascular tissue (Rudall and Caddick, 1994). A fraction of the phenolic material is saponifiable, yielding ferulic and p-coumaric acid and its dimers (Carpita, 1986), but a substantial amount of the aromatic material is nonsaponifiable hydroxycinnamic acids and their ethers (Scalbert et al., 1985). During vascular development and the formation of a thickened rind of the stem, cellulose, glucuronoxylans (GXs), and GAX become embedded in lignin (Carpita, 1996).

The genome sequences of *Arabidopsis (Arabidopsis thaliana*; Arabidopsis Genome Initiative, 2000), rice (*Oryza sativa*; International Rice Genome Sequencing Project, 2005), and maize (*Zea mays* cv. B73; Schnable et al., 2009) enabled comparative genomic analyses of cell-wall-related genes of two grass species with a typical eudicot, *Arabidopsis*. About 60% of *Arabidopsis* genes are annotated with predicted functions (Swarbreck et al., 2008). Based on sequence similarities, the Carbohydrate-Active enZYme database comprises families of glycosyl transferases (GTs), glycosyl hydrolases (GHs), and other carbohydratemetabolizing enzymes (Lombard et al., 2014; http://www.cazy. org/). We annotated over 750 maize homologs of these cell-wallrelated genes and assembled them into gene families predicted to function in cell-wall biogenesis (Penning et al., 2009).

We used phylogenetic comparisons of maize and rice gene families with those of Arabidopsis to characterize potential divergences that might explain the differences in composition between type I and type II walls. A further distinction in maize compared to other grasses is a recent genome duplication event (Gaut and Doebley, 1997). We find that many of these paralogous genes were retained, but loss of genes resulted in splitting of a single-gene function between paralogs (subfunctionalization), new function in a paralog gene (neofunctionalization), or a combination of both events (subneofunctionalization) to a greater extent in maize than other grasses (Penning et al., 2009). Copy-number and presence-absence variation has resulted in retention of both paralogs, only one of them, or neither of them (Springer et al., 2009; Swanson-Wagner et al., 2010). Since our previous study (Penning et al., 2009), we have developed a more robust annotation of nearly 1,200 maize genes and classified them into cell-wall-related gene families and their respective subgroups. In addition to families of substrate generation, cellulose and polysaccharide synthases, GTs, and cell-wall modifying enzymes, we broadened the inventory of cell-wall-related genes to include many new families involved in polysaccharide sidegroup construction, proteases, glycosylphosphatidylinositol (GPI)anchored proteins, glycoprotein synthesis, and signaling. We show here that, with few exceptions, the composition of gene families of any angiosperm species and levels of expression of their members are not clearly correlated with wall composition.

As the gene families had almost equal representation of all genes associated with cell-wall synthesis, we explored whether maize differential expression of these genes was correlated with the type of wall made. We established differential expression profiles of rind tissues from developing maize stem internodes representative of primary wall versus secondary wall formation to classify highly expressed members of these families. From these data, we could establish if Arabidopsis homologs closest in sequence were expressed in a similar primary or secondary wall context and, thus, could be considered functional orthologs. The number of potential orthologs with Arabidopsis genes, based on common expression during primary or secondary wall formation of the most similar sequences, was limited. However, we found robust expression of maize genes not only encoding synthases of GAX and mixed-linkage glucans but also those encoding synthases of the pectic polysaccharides, RG-I and HG, and of XyGs-polysaccharides that are only minor constituents

of the maize cell wall. Although one cannot infer that expression of a gene necessarily results in translation of their corresponding proteins, our recent analyses of the glycome and proteome of maize Golgi demonstrates unequivocally that the relevant biosynthetic enzymes are present and that substantial amounts of the pectins and XyGs accumulate in the cisternae of this organelle (Okekeogbu et al., 2019). Thus, we propose that cell-wall composition is determined by mechanisms in sorting or metabolism of polysaccharides downstream from Golgi-based synthesis.

MATERIALS AND METHODS

Sequence Alignments and Dendrogram Development

Sequence alignments and phylogenetic trees were constructed as described by Penning et al. (2009). Dendrograms were assembled from protein-coding sequences by the neighbor-joining method in ClustalW (Saitou and Nei, 1987; Chenna et al., 2003). The trees were bootstrapped 1,000 times. After the initial multiple alignment, individual clade alignments were checked using Multalin (Corpet, 1988; http://www-archbac.u-psud.fr/genomics/multalin.html). Matches to conserved regions within groups of family clades with suspect alignments were manually checked using InterProScan (Zdobnov and Apweiler, 2001; http://www.ebi.ac.uk/Tools/InterProScan/), and nonmatching members of the families were removed. Dendrograms were drawn using TreeDyn (Chevenet et al., 2006; http://www.treedyn.org/).

Protein Motif Analyses

Sequence locations and distances were verified using the Gbrowse designed for rice (http://www.jcvi.org), Arabidopsis (http://www. arabidopsis.org), and our custom maize annotation database comprising proteins from maize version 2 WGS (www.maizeGDB. org), using the largest representative transcript of each gene. We also used the Phytozome Protein domain FAMilies/Kyoto Encyclopedia of Genes and Genomes descriptions (https://phytozome.jgi.doe. gov/pz/portal.html), which were in general agreement with our maize cell-wall protein database at the Maize Genetics and Genome database, but our manual annotations based on phylogenetic trees constructed with the most similar Arabidopsis and rice sequences refined specific gene descriptions. We have augmented the public database at Maize Genetics and Genome for improved annotation of maize cell wall protein families (https://www.maizegdb.org/ gbrowse/maize_v2test?l=CellWallGenes;l=Gene_models;q= Chr1:2650000.2699999). This resource is also posted at http:// cellwall.genomics.purdue.edu.

Expression Analysis

Seeds of maize (*Z. mays* cv. B73) were obtained from the Maize Genetics Cooperation Center at Champaign, IL, and propagated and grown at the Purdue University Agricultural Center for Research and Education (West Lafayette, IN). Stem elongation began at the fifth-leaf stage and culminated with tassel formation 10 weeks postplanting. At 49-day postplanting, rind tissues of internodes 4–8 from a minimum of three plants were excised aseptically and immediately plunged in liquid N_2 and pulverized

by mortar and pestle under additional liquid N₂. Approximately 2 mg of ground tissue was incubated with 1 ml of ice-cold TRIzol reagent (Invitrogen, Life Technologies) and extracted according to the manufacturer's directions. Purified RNA was dissolved in 100 μ l of diethyl pyrocarbonate-treated nanopure water, and quality and concentration were determined spectrophotometrically. Internode 7 failed a quality control analysis for expression of a set of housekeeping genes (Sekhorn et al., 2011) and was excluded from further analysis.

Sections of each internode were frozen to -80° C in Neg 50 section medium (Richard-Allan Scientific, Kalamazoo, MI) and cross-sectioned to a thickness of 100 µm using a Microm HM550 Cryostat (Richard-Allan Scientific) at -20° C. Sections were thawed, rinsed with water, and stained using 2% w/v Wiesner's solution (phloroglucinol) in a 1:1 mixture of methanol and 50% HCl (v/v), freshly diluted to 5% in water. Images were taken using a SPOT Insight FireWire 4 Megasample Color Mosaic Camera (SPOT imaging systems, www.spotimaging.com) attached to a Nikon SMZ 1500 stereomicroscope (Nikon Corporation, Kanagawa, Japan) using a variable objective lens set to 10×. Images were captured using SPOT Advanced software version 4.1 (SPOT imaging systems).

Expression analysis was carried out as previously described (Penning et al., 2014). Briefly, pooled RNA samples from three biological replicates were sequenced using an Illumina HiSeq 2000 to process 100 bp × 100 bp libraries of ~400-bp inserts. High-quality trimmed sequences were mapped to the Maize B73 sequence V2 from Plant GDB (http://www.plantgdb.org) using Bowtie2 (Langmead et al., 2009), except in instances where the reads mapped exactly to two genes due to the high degree of gene duplication in maize. A custom Perl script was used to split these reads between the two loci (Penning et al., 2014). A separate set of Perl scripts was used to add closest Arabidopsis gene by sequence with description and expect value to the file. One read per million (20 reads per 20 M reads) or greater was used as a threshold for the detection of transcript (Li et al., 2010; Chang et al., 2012). RNA-seq data are available at https:// www.ncbi.nlm.nih.gov/sra/PRJNA522448 (datasets: SRX5387736, SRX5387731, SRX5387711, SRX5387715).

RESULTS

Annotation of Gene Function

General gene functions in encoding the enzymes of nucleotide sugar interconversion and transport, of polysaccharide synthases and glycosyl transferases, and those that encode their hydrolases and lyases have been inferred primarily from sequence similarity with bacterial genes of similar functions (Lombard et al., 2014). Bioinformatic approaches have extended predictions of gene function across eukaryotic species as increasing numbers of genomes became available. Phytozome, the Plant Comparative Genomics portal of the Department of Energy's Joint Genome Institute (https:// phytozome.jgi.doe.gosv), provides the plant science community with a large assembly of genomes from the JGI-sequencing initiative and those publicly available from other resources (Goodstein et al., 2012). Gene sets are best annotated using the Phytozome Protein domain FAMilies/Kyoto Encyclopedia of Genes and Genomes platforms and complemented with assignments by InterPro protein analysis tools that more closely predict function. Nevertheless, these remain hypothetical in the absence of biochemical or cell biological characterizations. Functional annotation has been facilitated by the identification of mutants representative of the different subgroups within the large gene families of glycosyl transferases and hydrolases. Some of these show chemical and spectroscopic signatures resulting from modified polymer fine structure that does not otherwise affect plant growth or development (Carpita and McCann, 2015).

The three angiosperms examined, Arabidopsis, rice, and maize, had members represented in every family of cell-wall-related genes. Within large families common to grasses and dicots, subgroup structure indicated putative orthology based on high sequence similarity of genes and similar patterns of expression at elongation and primary wall synthesis versus secondary wall synthesis stages of stem development. However, in other subgroups, grasses displayed a marked divergence from Arabidopsis in the number of members and degree of sequence similarity of homologs within a family subgroup, or even the presence or absence of grass-specific subgroups. Examples of all three of these characteristics are observed in the subgroup structure of the cellulose synthase (CesA)/cellulosesynthase-like (Csl) superfamily, where putative orthologs can be identified for the CesA and CslD subgroups, divergence of rice and maize CslAs from those of Arabidopsis, and grass-unique subgroups of CslF and CslH from Arabidopsis-only subgroup CslB (Figure S1).

Gene Expression During Stem Development

An estimate of potential functional orthology is obtained by comparative gene expression during the same stages of organ development. For grasses, stem elongation begins in basal internodes and progresses sequentially in upper internodes, culminating with flowering. In field-grown maize, internode development begins about the fifth-leaf stage at ~35 days postplanting and culminates with tasseling ~10 weeks. Stem elongation rates peak ~7 weeks (49 days), when lower internodes 4 and 5 have ceased elongation and the cells of the rind are more actively engaged in secondary wall formation and lignification, while internodes 6 and 8 continue elongation and transition to secondary wall formation. The rind constitutes the outer rings of vascular bundles with fibers to form distinct bands of cells ~0.5 cm thick that peel from the pith core during late development (Figure 1). To evaluate gene expression across the internodes, we used \geq 95 reads per 20 M from the four internodes as a minimal criterion of expression. Furthermore, we used \geq 500 reads per 20 M as a criterion to evaluate expression ratios from the elongation stages versus secondary wall forming stages of rind development (Table S1). From RNA-seq analysis of rind tissues from these four internodes, we found an expression ratio of 2 or higher in transcript abundance in lower internodes compared to upper internodes to be a suitable indicator of expression related to secondary wall formation. Conversely, ratios <1 indicated genes more associated with primary wall formation during internode elongation.

Although an expression ratio of ≥ 2 was consistent with association with secondary wall formation, a large proportion of the transcript ratios fell between 1 and 2, indicating more



FIGURE 1 | Cell wall thickness and lignin content increase in rind tissues of maize internodes with developmental age. Phloroglucinol staining intensity increases from faint pink to dark red in stem sections from (A) an elongation stage internode (internode 8), (B) a transitional stage internode (internode 6), and (C) a secondary wall stage internode (internode 4). Scale bar, 1 mm.

complicated expression patterns, which we termed "transitional" from primary wall formation to secondary wall formation (**Table S1**). We also summarized the expression behaviors of families and subgroups of families associated with specific polysaccharides, denoting the genes expressed during stages of primary wall formation, transitional, and secondary wall formation (**Table 1**). In general, genes of nucleotide sugar interconversion and transport are balanced across all three stages. More genes of XyG, glucomannan, mixed-linked β -glucan, and pectin synthesis are expressed during primary wall synthesis stages, but those of GAX are expressed across all stages. By

contrast, genes of monolignol and lignin biosynthesis are highly expressed during secondary wall synthesis (**Table 1**).

As similar discriminations of gene expression related to primary and secondary wall formation were established in *Arabidopsis* (Brown et al., 2005), putative orthology could be established by function in the same developmental context rather than the homologs most similar in sequence. Using these criteria, we found that only one quarter of maize cell-wall-related genes expressed in stems during secondary wall development were putatively orthologous with those of *Arabidopsis*. A complete index of over 1,200 maize cell-wall-related genes, ratio of TABLE 1 | Expression profiles of maize genes in family subgroups of cell-wall-related functions.

Cell Wall function	Number of Genes (Number expresssed) ¹	Elongation	Transitional	Secondary
Sucrose Synthases	8(8)	4	3	1
Nucleotide sugar interconversion	46(39)	12	8	12
Nucleotide sugar transporters	65(58)	20	15	15
Cellulose synthases (CesAs)	20(19)	6	5	6
GAX synthesis				
GT8A (GUX)	7(6)	2	2	0
GT8C (GATL)	10(7)	3	2	1
GT43	16(14)	0	6	6
GT47E	11(10)	4	1	4
GT61	33(22)	7	3	6
XyG synthesis				
CsIC	8(6)	6	0	0
GT34 (XXT)	18(6)	3	0	0
GT37 (FUT)	17(4)	1	0	1
GT47A	23(5)	3	1	0
(Gluco)mannan synthesis				
CsIA	10(10)	8	1	0
GT106B (MSR)	6(6)	6	0	0
Mixed-linkage β-glucan synthesis				
CsIF	9(7)	4	0	1
AGP/N-Glyc/HRGP-like	- ()			
AGP/Fasciclin	10(7)	3	0	2
GT31	40(34)	18	6	3
GT77	23(12)	4	1	1
ER/Golgi resident	41(37)	27	4	1
Pectin Synthesis	× /			
GT8D (GAUT)	23(23)	16	3	2
GT106A(RRT), GT106C(PAGR), and D	16(13)	6	3	1
GT47B	9(8)	6	2	0
Polygalacturonases	47(19)	8	3	1
Acetyl-transferases (TBL/BAHD)	77(59)	20	13	6
GPI-anchored proteins (COB/SKU)	22(15)	9	1	1
Expansins/XTHs	87(41)	23	4	4
Monolignol Synthesis	100(71)	17	7	30
Peroxidases/Laccases	148(67)	26	7	15

¹Total expression \geq 95 reads per 20 M; expression ratios calculated for those with \geq 500 reads per 20 M.

Ratio of transcripts from rind tissue of internodes 4–5 (secondary): internodes 6 and 8 (elongation) for genes expressed at >500 reads per 20 M were classified as elongation

 (≤ 1.04) , transitional (1.05–1.94), and secondary (≥ 1.95) expression. A complete collection of the genes we annotated, including potential Arabidopsis orthologs is given in Table S1.

secondary/primary wall expression, and putative *Arabidopsis* orthologs is presented (Table S1).

Cellulose Synthase/Cellulose-Synthase-Like Superfamily

The CesA/Csl superfamily comprises up to 10 distinct subfamilies that display divergent membership between grasses and *Arabidopsis*, with the CslB group not represented in the rice or maize genome and the CslF and CslH groups found only in grasses (**Figure S1**). CesA gene families of rice, maize, and *Arabidopsis* show similar subgroup structure (**Figure 2A**), and association within the same subgroup is indicative of a role in primary or secondary wall formation (Tanaka et al., 2003; Appenzeller et al., 2004; Brown et al., 2005). We found that maize CesA1 through CesA9 and their paralogs were associated with elongation and primary wall synthesis stages of development, and CesA10 through CesA12 and their paralogs were associated with secondary wall formation (**Figure 3A**; **Table S1**). From their subgroup membership, those expressed exhibited a ratio of relative secondary to primary wall expression that indicated orthology with *Arabidopsis* for all but one of the 17 CesAs expressed at \geq 500 reads per 20 M. Five CesAs whose expression was regarded as "transitional" were homologous with *Arabidopsis* genes of similar sequence that were associated with primary wall formation.

Members of the CslA family in *Arabidopsis* and poplar encode proteins with (gluco)mannan synthase from GDP-Man and GDP-Glc substrates (Dhugga et al., 2004; Liepman et al., 2007). Apart from a few exceptions, maize and rice CslA genes diverged markedly in sequence similarity from *Arabidopsis* genes (**Figure 2B**; **Figure 3B**). All 10 maize CslA genes were expressed at \geq 95 reads per 20 M, but expression was either constitutive or higher during primary wall formation (**Figure 3B**; **Table S1**). However, because of the divergence from *Arabidopsis* of the genes in this family, only a single CslA2 gene could be considered potentially orthologous.

Among angiosperm species, the mixed-linkage β -glucans were found only in abundance in the cell walls of *Poales* species





(Smith and Harris, 1999), but more recent studies reported immunocytochemical evidence for small amounts of these β -glucans in species outside the Order Poales (Trethewey et al., 2005). The CsIF and CsIH genes that encode the synthases of the mixed-linkage β -glucans are found only in grass species (Burton et al., 2006; Doblin et al., 2009; Little et al., 2018). Subgroups CsIJ and CsIM have also been defined in the grasses but are represented outside the *Poales* (Little et al., 2018). Although CsIF

genes are numerous in maize and rice, CslH is not found in the B73 genome, and the single CslG gene cannot be distinguished as a separate subgroup from CslJ or CslM (**Figure S1**).

Seven of the nine maize CslF genes were expressed at \geq 95 reads per 20 M, but five of them were expressed at levels to calculated expression ratios. Four of the five showed primary wall association, but a CslF5 was more strongly expressed during secondary wall formation (**Figure 2C**; **Figure 3C**). The mixed-linkage

 β -glucans are synthesized during primary wall formation in growing coleoptiles and are largely degraded at the end of elongation (Carpita et al., 2001), but these glucans can continue to be synthesized and persist during secondary wall formation (Vega-Sanchez et al., 2013).

Genes of Substrate Generation

All eight sucrose synthase genes were expressed, with only one associated with secondary wall formation (Table 1). Ten families of genes encode enzymes of nucleotide sugar interconversion pathways responsible for synthesis de novo of the major neutral and acidic monosaccharides for polysaccharide synthesis (Reiter and Vanzin, 2001; Yin et al., 2011). Apart from duplications in several maize families, genes are generally orthologous based on close phylogenetic relationships and common constitutive or primary wall stage expression patterns (Figure S2; Table S1). However, at least one member of each family is more highly expressed during secondary wall formation. Although UDP-Ara is synthesized in the pyranose form, a substantial portion of the Ara in cell wall polysaccharides is in the furanose form. Konishi et al. (2007) characterized reversibly glycosylated proteins (RGPs) as UDP-Ara mutases (UAMs) that interconvert UDP-Arap and UDP-Araf. Downregulation of UAM expression results in arabinose deficiencies in both rice (Konishi et al., 2011) and Arabidopsis (Rautengarten et al., 2011). Nine maize homologs show high sequence similarity with four of the five Arabidopsis RGP (UAM) genes; three of these were highly expressed during secondary wall formation (Figure S2F; Table S1).

The large multigene family of nucleotide sugar transporters (NSTs) of plant species comprises six subgroups (Handford et al., 2004; Orellana et al., 2016). GONST1 of subgroup III has been defined as a GDP-Man transporter (Baldwin et al., 2001), although GDP-Fuc and UDP-Gal/UDP-Glc are transported by other members of this subgroup (Rautengarten et al., 2014); GONST1 channels substrate to sphingolipid glycosylation and not glucomannan synthesis (Mortimer et al., 2013). Subgroup II includes several UDP-Gal/UDP-Glc transporters (Norambuena et al., 2002), and the UDP-Gal transporter members of subgroup I, or the NST-KT clade, were shown in heterologous systems to have UDP-Rha/UDP-Gal transport activity (Rautengarten et al., 2014). More recently, subgroup V has been shown to contain UDP-GlcA/GalA transporters (Saez-Aguayo et al., 2017) and subgroup VI to contain UDP-Araf transporters (Rautengarten et al., 2017). The maize genome has several NSTs in all six subgroups, including 18 UDP-Gal transporters and UDP-Xyl transporters from subgroup I, five UTRs from subgroup II, and 20 GONST-, NST-, and UTR-like genes in subgroup III (Figure S3; Table S1). The majority of these were expressed constitutively or predominantly during primary wall formation, but, with the exception of group II, at least one member was highly expressed during secondary wall formation (Figure S3; Table S1).

Glucuronoarabinoxylan Synthesis

GAXs are the major noncellulosic glycans in the type II primary walls of grasses. The GAX polysaccharides comprise $(1\rightarrow 4)$ - β -D-xylan backbones with side groups of α -GlcA and α -4-O-Me-GlcA linked at the xylosyl O-2 position, and Araf residues linked either

at the O-2 or O-3 position depending on species. The grasses show a predominance of substitution of Araf at the O-3 xylosyl residue, whereas in noncommelinid species, the Araf residues are at the O-2 position (Carpita and Gibeaut, 1993). However, the GAXs of all angiosperm species are capable of adding Araf residues to either or both of the O-2 and O-3 xylosyl units of the backbone (Carpita, 1996), indicating that genes encoding these transferase activities are expressed in grasses and dicots.

Synthesis of xylan backbones require two distinct members of family GT43, which encode the irregular xylem9 (IRX9) and IRX14 inverting-type xylosyl transferases (Brown et al., 2007; Wu et al., 2009; Smith et al., 2017). Xylan chain synthesis also requires participation of IRX10 and IRX10-like xylan xylosyl transferases of family GT47 subgroup E (Brown et al., 2009; Wu et al., 2010; Zeng et al., 2016). Within this same subgroup is FRA8, which is reported to encode a putative xylan-decorating glucuronosyl transferase (Zhong et al., 2005). These two distinct activities within the same GT47 subgroup underscore the need for a more thorough determination of specific function of members of the six subgroups of the GT47 family. Thus far, all are inverting-type glycosyl transferases, i.e., those that convert the α -D- or β -L-nucleotide sugar moiety into a β -D- or α -L-linkage in a polysaccharide, but members of different subgroups use different nucleotide sugars or polymer substrates in transferase reactions (Penning et al., 2009). Arabinosyl and xylosyl side groups are attached by family GT61 inverting-type transferases (Anders et al., 2012; Chiniquy et al., 2012).

Of the 16 maize IRX9 and IRX14 genes, 12 were highly expressed during stem development, 6 of which had expression ratios ≥ 2 (Figure 4A; Table S1). Similarly, of the 11 *IRX10* genes in maize, 9 of them were expressed during stem development, and 4 IRX10-1 genes more highly expressed during secondary wall formation (Figure 4B; Table S1). The maize family GT61 comprises 33 genes, 6 of which are expressed during secondary cell wall formation (Figure 4C; Table S1). Glucuronosyl residues are ubiquitous side groups of xylans, the sole sugar substituent of secondary wall xylans. These α -linked GlcA side groups are attached by GT family 8 subgroup A retaining-type glucuronosyl transferases (GUX) by members that display selectivity with respect to the periodicity of GlcA attachment along the xylan backbone (Mortimer et al., 2010). Six of seven maize GUX genes were expressed at \geq 95 reads per 20 M, with four of them expressed constitutively (Figure 4D; Table S1). Members of GT8 subgroup C of galacturonosyl transferase-like (GATL) genes function to initiate GX synthesis through participation in the synthesis of a tetrasaccharide primer sequence (Lee et al., 2007). Maize GATL2 is homologous with the Arabidopsis PARVUS gene, which was constitutively expressed, whereas GATL7b showed high secondary wall expression (Figure 4E; Table S1).

In addition to the sugar side groups, xylans can have high degrees of acetylation, and the large family of *trichomebirefringence-like* (*TBL*) and *TBL-like* genes encode acetyl transferase enzymes, at least some of which are involved specifically in acetylation of xylans (Gille and Pauly, 2012; Gao et al., 2017). The maize *TBL* family comprises 56 genes in five subgroups, with 7 additional genes related to *reduced wall O-acetylation* (*RWA*) and two homologs to *Arabidopsis xyloglucan*



acetyl transferase9 (AXY9); a majority of these acetyl transferases were expressed constitutively or higher during primary wall formation (**Table S1**), but several exhibited particularly high expression during secondary wall formation (**Figure S4**). One of the hallmarks of the grass cell wall is the autofluorescence of the primary wall from a phenylpropanoid network integrated with xylans. Ferulic acid and *p*-coumaric acid are known to be extended from arabinosyl residues of GAX, and the BAHD family contains CoA-dependent transferases thought to participate in addition of these hydroxycinnamic acids (Rautengarten et al., 2012; Molinari et al., 2013). The maize BAHD family numbers 12 genes, of which 3 are highly expressed during secondary wall formation (**Figure S5; Table S1**).

Xyloglucan Synthesis

The principal cross-linking glycans of dicots and noncommelinid monocots are XyGs, whose fundamental structure is a $(1\rightarrow 4)$ - β -D-glucan backbone branched at the O-6 by α -D-Xyl residues (Carpita and Gibeaut, 1993). In most angiosperms, including *Arabidopsis*, three consecutive glucosyl residues of every four in the backbone are subtended by Xyl residues, and the two Xyl residues closer to the reducing end of the backbone can be substituted further at the O-2 of one or both positions with β -D-Gal residues (Scheller and Ulvskov, 2010). If a Gal residue subtends the Xyl residue closest to the reducing end, an α -L-Fuc is likely to be added to the Gal O-2 position. The glucan backbones of XyGs of grasses are irregularly branched with one or two Xyl residues, and these contain an occasional Gal residue (Carpita, 1996).

From heterologous expression experiments, genes of the CslC subgroup likely encode the XyG glucan backbone synthases (Cocuron et al., 2007), and genes of the GT34 family encode the XyG xylosyl transferases (XXTs) (Cavalier et al., 2008; Zabotina et al., 2012). The *mur3* mutation was traced to a gene encoding a xyloglucan galactosyl transferase (GalT) in family GT47 subgroup A, whose product decorates the xylosyl residue closest to the reducing end of the oligomer unit (Madson et al., 2003), and the subgroup A homolog GalT decorates the "middle" xylosyl residue (Kong et al., 2015). When a Gal is added to the first Xyl residue, it becomes a possible substrate for a family GT43 transferase that fucosylates it (Vanzin et al., 2002).

Despite small amounts of the truncated form of XyGs that accumulate in the walls of maize and other grasses, these species have the capacity to synthesize a fucosylated XyG. The maize genome contains eight CslC genes and expressed six of them (Figure 5A; Table S1). Expression of all six was associated with primary wall formation. Only one of these genes, CslC5a, had an expression profile and sufficient sequence similarity with Arabidopsis to be considered a potential ortholog. The maize genomes possess 17 Family GT34 XXT and XXT-like genes (Figure 5B; Table S1). The more highly expressed members of the family had expression ratios that indicated association with primary wall XyG synthesis. The maize genome also has 22 members of family GT47 subgroup A that encodes the XyGspecific Gal transferases, four of which were highly expressed at early stages of stem development. Three of them are putatively orthologous to MUR3 (Figure 5C; Table S1). Despite the lack of detectable fucosylation of XyGs of grasses, maize also has 17 GT37 *fucosyl transferases* (*FUTs*); four of these were expressed at \geq 500 reads per 20 M, one (*FUT2*) during primary cell wall growth stages and one (*FUT11*) with strong expression during secondary wall formation (**Figure 5D**; **Table S1**).

Pectin Synthesis

Although the type II cell walls of commelinid monocots generally have limited amounts of pectic polysaccharides, they are found transiently in higher abundance in some developmental contexts, such as endosperm development (Chateigner-Boutin et al., 2014; Zhang et al., 2016). The pectin matrix consists primarily of two kinds of pectic polysaccharide backbones of $(1\rightarrow 4)$ - α -Dhomogalacturonan (HG) and repeating units of the *O*-2- α -D-Rha- $(1\rightarrow 4)$ - α -D-Gal disaccharide in rhamnogalacturonan-I (RG-I) (Caffall and Mohnen, 2009). Some HGs are branched with Xyl residues to form Xyl-HGs or possess clusters of four complex oligosaccharides to form RG-II, a polysaccharide that forms boron di-diester crosslinks. Branched $(1\rightarrow 5)$ - α -L-arabinans and type I $(1\rightarrow 4)$ - α -D-galactans with appendant Ara residues are attached typically to the Rha *O*-4 position of RG-I.

The maize genome contains numerous genes associated with the synthesis of pectins. The GT8 subgroup D retaining-type galacturonosyl transferase (GAUT) family comprises 23 genes, and all of them were expressed at \geq 95 reads per 20 M throughout stem development (Figure S6A; Table S1). The vast majority of them were expressed at primary wall or transitional stages, but two nonorthologous genes, GAUT11b and GAUT11c, displayed higher relative expression during secondary wall synthesis. Pectin RG-I synthesis requires a rhamnosyl transferase (RRT1) from family GT106 (Takenaka et al., 2018). From gene expression related to seed mucilage RG-I synthesis, the Arabidopsis GAUT11 is the prime candidate for the GalA transferase associated with RG-I synthesis (Voiniciuc et al., 2018), with a close homolog also found in maize (Table S1). The maize genome also contains a large family of GT106 genes, including three of four RRTs that are expressed mostly constitutively, but with RRT1b more highly expressed during secondary wall formation (Figure S6B). A related subgroup of GT106 are the pectin arabinogalactan synthesis-related (PAGR; Stonebloom et al., 2016) genes, and mannan synthesis-related transferase (MSR; Wang et al., 2013) genes. Seven of the nine maize PAGR genes were expressed mostly during primary wall formation, but PAGR-L1 showed higher expression during transitional stages of development (Figure S6C). All six MSR genes were constitutively expressed (Table S1).

The maize genome contains a vast number of genes that encode enzymes of pectin depolymerization. *Polygalacturonase* (*PGase*) genes comprise six subgroups with high expression of members in all of them except subgroups E and F (**Figure S7**; **Table S1**). None is expressed specifically during secondary wall formation. Maize also has four *RG-I lyases* (*RGILs*), with only *RGIL1* significantly expressed (**Figure S7**F).

AGPs and Other GPI-Anchored Proteins

GPI-anchored peptidoglycans, such as type II arabinogalactan proteins (AGPs), with their highly branched $(1\rightarrow 3)$ -, $(1\rightarrow 6)$ -, and





 $(1\rightarrow3, 1\rightarrow6)$ - β -D-galactan chains and Ara side groups, are found in small amounts in primary cell walls (Showalter, 1993; Borner et al., 2003; Johnson et al., 2003). Two, *FLA2c* and *FLA11*, of seven AGP/ fasciclin genes showed relatively higher secondary wall expression (**Figure S8A**; **Table S1**). Family GT31 represents a large family of six subgroups and includes GalTs that are predicted to form the $(1\rightarrow3)$ - β - and $(1\rightarrow6)$ - β -linked galactan chains of type II AGPs. Most of the GT31 genes are expressed during primary wall formation, but one member of GT31 subgroup A, *GT31A3*, and two members of subgroup F, *GT31F4* and *GT31F5*, exhibited predominantly secondary wall expression (**Figure S8B**; **Table S1**). Two notable families of GPI-anchored proteins discovered in *Arabidopsis* are *skewed growth* (*SKU*) proteins, mutations in which result in abnormal growth symmetry (Rutherford and Masson, 1996; Sedbrook et al., 2002), and COBRA proteins involved in determining the direction of wall expansion and patterning of cellulose in primary walls (Schindelman et al., 2001) and cellulose content and tensile strength of the floral stem (Brown et al., 2005). Maize *SKU* genes number 13, with seven expressed predominantly during primary wall formation (**Figure S9**; **Table S1**). Nine *COBRA* genes comprise the maize family, with three expressed more or less constitutively and COBL4b expressed during secondary wall formation (Figure S9; Table S1).

Expansins and Endotransglucosylase/ Hydrolases

The cell wall is residence to hundreds of enzymes, such as expansins and transglucosylases, polysaccharide hydrolases and lyases, oxidoreductases, and proteases, which function in wall remodeling and metabolism (Boudart et al., 2005; Hervé et al., 2016), and maize expresses many of these during both primary and secondary wall stages of growth (Table S1). Expansins and the GH16 family of xyloglucan endotransglucosylase/hydrolases (XTHs) are implicated in cellulose microfibril separation during growth and the rejoining of XyGs to maintain tensile strength, respectively (Cosgrove, 2000; Rose et al., 2002). The maize genome contains over 50 α -Expansin $(\alpha$ -*Exp*), α -*Expansin-like* (α -*Exp*-*like*), and β -*Expansin* (β -*Exp*) genes, with most expressed during early growth except for an α -Exp8a and an α -ExpL2c more highly expressed during secondary wall formation (Figure S10; Table S1). Similarly, 30 maize XTH genes form three subgroups, with the majority of them expressed during elongation and primary wall stages of growth, but three XTHB genes and one XTHC gene were more highly expressed during secondary wall formation (Figure S11; Table S1).

Monolignol and Lignin Synthesis

A major distinction of the type II primary cell walls of grasses is the presence of a phenylpropanoid network (Carpita, 1996). Several members in each family of genes that encode the enzymes of monolignol synthesis were expressed during primary wall synthesis, even though most of the genes were upregulated in older internodes. All but one of the phenylalanine/tyrosine ammonia lyase (PAL) genes were expressed in secondary cell-wall-forming tissues (Figure 6A; Table S1). In contrast, expression of all other genes of monolignol formation, such as 4-coumarate CoA ligase, cinnamyl alcohol dehydrogenase, hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase, cinnamyl CoA reductase (CCR), and caffeoyl-coenzyme A 3-O-methyltransferase (CCoAOMT), were expressed during primary wall formation and secondary wall formation in roughly equal numbers (Figures 6B-D; Table S1). Peroxidases are encoded by a huge family of 86 genes in seven subgroups and Laccases numbered 24 genes (Table S1). Although expression ratios showed genes with strong elongation-specific expression or secondary wall expression, ratios were widely clustered between 2 and 12, indicating more complex patterns of expression.

Other Wall-Related Proteins

Apart from extensin domains within some of the AGP proteoglycans, the classic Ser-Hyp₄-type extensin proteins are in low abundance in the maize genome. Nevertheless, we identified several genes related to their synthesis and glycosylation. For example, all but 1 of the 10 prolyl-4-hydroxylases that synthesize Hyp were expressed, and all of these during primary wall formation. Of the 23 genes that encode family GT77 arabinosyl transferases, which attach the Ara side groups of extensin, 12 were expressed at \geq 95 reads per 20 M, mostly during primary wall stages, but 1, *reduced residual arabinose1b* (*RRA1b*; Egelund et al., 2007), was more strongly expressed during secondary wall formation (**Table S1**).

The progenitor species of commelinid monocots diverged from dicotyledonous and other monocotyledonous species about 120 million years ago (Hertweck et al., 2015) to species that made cell walls with mostly GAX as the cross-linking glycan, and with much less XyG and pectin than their common eudicot ancestors (Smith and Harris, 1999). The feruloylation of the Ara residues in the commelinid species and the initiation of phenylpropanoid networks in the primary walls of all commelinids gave characteristic strong autofluorescence in UV light (Rudall and Caddick, 1994). The Order Poales split from its closest commelinid relatives about 65 million years ago (Hertweck et al., 2015), with expansion of CsIF and CsIH subgroups that encode the synthases of the mixed-linkage β -glucans (Burton et al., 2006; Doblin et al., 2009).

Despite the evolution of commelinid species with cell walls distinct from dicots and noncommelinid monocots, we show here that their genomes retained the capacity to make the cell wall polysaccharides of all angiosperm species. In addition, families of genes for synthesis of polysaccharides uncharacteristic of the two wall types are expressed in the expected developmental context of primary or secondary wall. Key questions are if the transcripts expressed become translated into protein, and if so, are these uncharacteristic polysaccharides actually made? Our recent work in comparative glycome and proteome analysis indicated that Arabidopsis and maize, representative species with type I and type II cell walls, respectively, synthesize and accumulate, in the Golgi, polysaccharides that are not found in abundance in their cell walls (Okekeogbu et al., 2019). Maize Golgi had a higher proportion of XyG to GAX, and Arabidopsis had a higher proportion of GAX to XyG in their respective Golgi. Three possible explanations were considered for these findings: (1) that polysaccharides that accumulate to higher abundances in Golgi might result from lower trafficking rates compared to higher rates for material trafficked to the wall, (2) that the lower trafficking rates might reflect diversion of a subset of polysaccharides to lytic compartments instead of the cell wall, or (3) that all polysaccharides are trafficked to the wall, but polysaccharides uncharacteristic of wall type are digested extracellularly and, therefore, fail to accumulate. As the trans-Golgi network/early endosome compartment plays a central role in post-Golgi synthesis, sorting, and packaging of polysaccharide and protein cargoes destined for the wall (Kang et al., 2011; Rosquete et al., 2018), we suggest that this compartment might be the site of discrimination.

The broader question of evolutionary significance is why a species would place a large metabolic investment in polymers that never accumulate in the wall. However, retaining the capacity to make alternative polysaccharides if mutations occur could be a selective advantage. If mutations occur in synthases of the characteristic polysaccharide, viability is not jeopardized if an alternative polysaccharide can be made. Such plasticity in wall composition is illustrated by the mutations that introduce severe alterations in polysaccharide structure or abundance yet plants maintain near normal growth and development. Among the starkest examples are the *Arabidopsis xxt1/xxt2* double mutant that completely lacks XyG in the wall, but despite a slight lowering





of tensile strength, plant structure is remarkably unchanged from wild type (Cavalier et al., 2008; Zabotina et al., 2012), and the survival of cells in liquid culture in the near absence of cellulose induced by a potent cellulose synthesis inhibitor regardless of wall type (Shedletzky et al., 1990; Shedletzky et al., 1992).

CONCLUSIONS

While it was tempting to hypothesize that the evolution of a completely distinct type of cell wall resulted in a drastic change in the gene families that encode its synthesis machinery, expression of specific gene family members is not the basis for the difference.

Although a few examples exist of the divergence of certain subgroups of families unique to grass species, the vast majority are populated with members in all angiosperms. Divergence within a subgroup gives evidence of neofunctionalization during speciation, but it is equally evident that grass species express large numbers of genes of these subfamilies that encode polysaccharides that accumulate in the Golgi but never traffic to or assemble in the wall. Thus, the type of cell wall made is established by post-Golgi mechanisms that remain to be determined. The capacity to make in any species the entire repertory of cell wall polysaccharides widens the spectrum of design properties for cell walls as materials with emergent properties.

DATA AVAILABILITY STATEMENT

An improved annotation of maize cell wall protein families is searchable at Maize GDB (https://www.maizegdb.org/ gbrowse/maize_v2test?l = CellWallGenes;l = Gene_models; q = Chr1:2650000.2699999). This resource is also downloadable at (http://cellwall.genomics.purdue.edu). RNA-seq data are available at https://www.ncbi.nlm.nih.gov/sra/PRJNA522448 (datasets: SRX5387736, SRX5387731, SRX5387711, SRX5387715).

AUTHOR CONTRIBUTIONS

BP, MM, and NC designed the research. BP carried out the experiments with assistance from MM and NC, and BP, MM, and NC analyzed the results and wrote the manuscript.

FUNDING

The work on cell-wall gene annotations was supported by the Center for Direct Catalytic Conversion of Biomass to Biofuels, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Office of

REFERENCES

- Anders, N., Wilkinson, M. D., Lovegrove, A., Freeman, J., Tryfona, T., Pellny, T. K., et al. (2012). Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proc. Natl. Acad. Sci. USA* 109, 989–993. doi: 10.1073/pnas.1115858109
- Appenzeller, L., Doblin, M., Barreiro, R., Wang, H., Niu, X., Kollipara, K., et al. (2004). Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (*CesA*) gene. *Cellulose* 11, 287–299. doi: 10.1023/B:C ELL.0000046417.84715.27
- Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815. doi: 10.1038/35048692
- Baldwin, T. C., Handford, M. G., Yuseff, M. I., Orellana, A., and Dupree, P. (2001). Identification and characterization of GONST1, a Golgi-localized GDPmannose transporter in Arabidopsis. *Plant Cell* 13, 2283–2295. doi: 10.1105/ tpc.010247
- Borner, G. H. H., Sherrier D. J., Stevens, T. J., Arkin, I. T., and Dupree, P. (2003). Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. *Plant Physiol.* 132, 568–577. doi: 10.1104/pp.103.021170
- Brown, D. M., Zeef, L. A. H., Ellis, J., Goodacre, R., and Turner, S. R. (2005). Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* 17, 2281– 2295. doi: 10.1105/tpc.105.031542
- Brown, D. M., Goubet, F., Vicky, W.W.A., Goodacre, R., Stephens, E., Dupree, P., et al. (2007). Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J.* 52, 1154–1168. doi: 10.1111/j.1365-313X.2007.03307.x
- Brown, D.M., Zhang, Z., Stephens, E., Dupree, P, and Turner, S.R. (2009). Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *Plant J.* 57, 732–746. doi: 10.1111/j.1365-313X.2008.03729.x
- Boudart, G., Jamet, E., Rossignol, M., Lafitte, C., Borderies, G., Jauneau, A., et al. (2005). Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: identification by mass spectrometry and bioinformatics. *Proteomics* 5, 212– 221. doi: 10.1002/pmic.200400882

Basic Energy Sciences (grant no. DE-SC0000997), and the expression analyses were supported by the US Department of Energy Feedstock Genomics Program, Office of Biological and Environmental Research, Office of Science (grant no. DE-FOA-0000598). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. USDA is an equal opportunity provider and employer.

ACKNOWLEDGMENTS

We thank Jyothi Thimmapuram and Ketaki Bhide (Bioinformatics Core, Purdue) for their help in analysis and submission of the RNA-seq data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01205/ full#supplementary-material

- Burton, R.A., Wilson, S.M., Harvey, A.J., Shirley, N.J., Medhurst, A., Stone, B.A., et al. (2006). Cellulose synthase-like *CslF* genes mediate the synthesis of cell wall (1,3;1,4)-β-D-glucans. *Science* 311, 1940–1942. doi: 10.1126/ science.1122975
- Caffall, K.H., and Mohnen, D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr. Res.* 344, 1879–1900. doi: 10.1016/j.carres.2009.05.021
- Carpita, N. C. (1986). Incorporation of proline and aromatic amino acids into cell walls of maize coleoptiles. *Plant Physiol.* 80, 660–666. doi: 10.1104/pp.80.3.660
- Carpita, N. C. (1996). Structure and biogenesis of the cell walls of grasses. Annu. Rev. Plant Physiol. Plant Mol. Biol. 47, 445–476. doi: 10.1146/annurev. arplant.47.1.445
- Carpita, N. C., and Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30. doi: 10.1111/j.1365-313X.1993.tb00007.x
- Carpita, N. C., Defernez, M., Findlay, K., Wells, B., Shoue, D. A., Catchpole, G., et al. (2001). Cell wall architecture of the elongating maize coleoptile. *Plant Physiol.* 127, 551–565. doi: 10.1104/pp.010146
- Carpita, N. C., and McCann, M. C. (2015). Characterizing visible and invisible cell wall mutant phenotypes. J. Exp. Bot. 66, 4145–4163. doi: 10.1093/jxb/erv090
- Cavalier, D. M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., et al. (2008). Disrupting two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell* 20, 1519–1537. doi: 10.1105/ tpc.108.059873
- Chang, Y. M., Liu, W. Y., Shih, A. C. C., Shen, M. N., Lu, C. H., Lu, M. Y. J., Yang, H. W., et al. (2012). Characterizing regulatory and functional differentiation between maize mesophyll and bundle sheath cells by transcriptomic analysis. *Plant Physiol.* 160, 165–177. doi: 10.1104/pp.112.203810
- Chateigner-Boutin, A-L., Bouchet, B., Alvarado, C., Bakan, B., and Guillon, F. (2014). The wheat grain contains pectic domains exhibiting specific spatial and development-associated distribution. PLoS ONE 9, e89620. doi: 10.1371/ journal.pone.0089620
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., and Higgins, D. G., et al. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31, 3497–3500. doi: 10.1093/nar/gkg500

- Chevenet, F., Brun, C., Banuls, A.L., Jacq, B., and Christen, R. (2006). TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics 7, 439. doi: 10.1186/1471-2105-7-439
- Chiniquy, D., Sharma, V., Schultink, A., Baidoo, E.E., Rautengarten, C., Cheng, K., et al. (2012). XAX1 from glycosyltransferase family 61 mediates xylosyl transfer to rice xylan. *Proc. Natl. Acad. Sci.*, USA 109, 17117–17122. doi: 10.1073/pnas.1202079109
- Cocuron, J. C., Lerouxel, O., Drakakaki, G., Alonso, A. P., Liepman, A. H., Keegstra, K., et al. (2007). A gene from the cellulose synthase-like C family encodes a b-1,4 glucan synthase. *Proc. Natl. Acad. Sci. USA* 104, 8550–8555. doi: 10.1073/pnas.0703133104
- Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 16, 10881–10890. doi: 10.1093/nar/16.22.10881
- Cosgrove, D. J. (2000). Loosening of plant cell walls by expansins. *Nature* 407, 321–326. doi: 10.1038/35030000
- Dhugga, K. S., Barreiro, R., Whitten, B., Stecca, K., Hazebroek, J., Randhawa, G. S., Dolan, M., Kinney, A. J., Tomes, D., Nichols, S., and Anderson, P. (2004) Guar seed β -mannan synthase is a member of the cellulose synthase super gene family. *Science* 303, 363–366. doi: 10.1126/science.1090908
- Doblin, M.S., Pettolino, F.A., Wilson, S.M., Campbell, R., Burton, R.A., Fincher, G.B., et al. (2009). A barley cellulose synthase-like CSLH gene mediates (1,3;1,4)-b-D-glucan synthesis in transgenic Arabidopsis. Proc. Natl. Acad. Sci. USA 106, 5996–6001. doi: 10.1073/pnas.0902019106
- Egelund, J., Obel, N., Ulvskov, P., Geshi, N., Pauly, M., Bacic, A., et al. (2007). Molecular characterization of two *Arabidopsis thaliana* glycosyltransferase mutants, *rra1* and *rra2*, which have a reduced residual arabinose content in a polymer tightly associated with the cellulosic wall residue. *Plant Mol. Biol.* 64, 439–451. doi: 10.1007/s11103-007-9162-y
- Gaut, B. S., and Doebley, J. F. (1997). DNA sequence evidence for the segmental allotetraploid origin of maize. *Proc. Natl. Acad. Sci. USA* 94, 6809–6814. doi: 10.1073/pnas.94.13.6809
- Gille, S., and Pauly, M. (2012). O-Acetylation of plant cell wall polysaccharides. Frontiers Plant Sci. 3, 12. doi: 10.3389/fpls.2012.00012
- Gao, Y., He, C., Zhang, D., Liu, X., Xu, Z., Gao, Y., et al. (2017). Two trichome birefringence-like proteins mediate xylan acetylation, which is essential for leaf blight resistance in rice. *Plant Physiol.* 173, 470–481. doi: 10.1104/pp.16.01618
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944
- Handford, M. G., Sicilia, F., Brandizzi, F., Chung, J. H., and Dupree, P. (2004). *Arabidopsis thaliana* expresses multiple Golgi-localised nucleotide-sugar transporters related to GONST1. *Mol. Genet. Genom.* 272, 397–410. doi: 10.1007/s00438-004-1071-z
- Hertweck, K.L., Kinney, M.S., Stuart, S.A., Maurin, O., Mathews, S., Chase, M.W., et al. (2015). Phylogenetics, divergence times and diversification from three genomic partitions in monocots. *Bot. J. Linn. Soc.* 178, 375–393. doi: 10.1111/ boj.12260
- Hervé, C., Duruflé, H., San Clemente, H., Albenne, C., Balliau, T., Zivy, M., et al. (2016). An enlarged cell wall proteome of *Arabidopsis thaliana* rosettes. *Proteomics* 16, 3183–3187. doi: 10.1002/pmic.201600290
- International Rice Genome Sequencing Project. (2005). The map-based sequence of the rice genome. *Nature* 436, 793–800. doi: 10.1038/nature03895
- Johnson, K. L., Jones, B. J., and Schultz, C. J. (2003). The fasciclin-like arabinogalactan proteins of *Arabidopsis*. A multigene of putative cell adhesion molecules. *Plant Physiol*. 133, 1911–1925. doi: 10.1104/pp.103.031237
- Kang, B-H., Nielsen E., Preuss, M. L., Mastronarde, D., and Staehelin, L. A. (2011). Electron tomography of RabA4b and PI-4Kb1-labeled *Trans* Golgi network compartments in *Arabidopsis. Traffic* 12, 313–329. doi: 10.1111/j.1600-0854.2010.01146.x
- Kong, Y., Peña, M.J., Renna, L., Avci, U., Pattathil, S., Tuomivaara, S.T., et al. (2015). Galactose-depleted xyloglucan is dysfunctional and leads to dwarfism in *Arabidopsis. Plant Physiol.* 167, 1296–1306. doi: 10.1104/ pp.114.255943
- Konishi, T., Takeda, T., Miyazaki, Y., Ohnishi-Kameyama, M., Hayashi, T., O'Neill, M.A., et al. (2007). A plant mutase that interconverts UDParabinofuranose and UDP-arabinopyranose. *Glycobiology* 17, 345–354. doi: 10.1093/glycob/cwl081

- Konishi, T., Aohara, T., Igasaki, T., Hayashi, N., Miyazaki, Y., Takahashi, A., et al. (2011). Down-regulation of UDP-arabinopyranose mutase reduces the proportion of arabinofuranose present in rice cell walls. *Phytochemistry* 72, 1962–1968. doi: 10.1016/j.phytochem.2011.07.012
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25. doi: 10.1186/gb-2009-10-3-r25
- Lee, C., Zhong, R., Richardson, E. A., Himmelsbach, D. S., McPhail, B. T., Ye, Z-H. (2007). The PARVUS gene is expressed in cells undergoing secondary wall thickening and is essential for glucuronoxylan biosynthesis. *Plant Cell Physiol.* 48, 1659–1672. doi: 10.1093/pcp/pcm155
- Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S.L., Kebrom, T.H., Provart, N., et al. (2010). The developmental dynamics of the maize leaf transcriptome. *Nature Genet.* 42, 1060–1067. doi: 10.1038/ng.703
- Liepman, A.H, Nairn, C.J., Willats, W.G.T., Sørensen, I., Roberts, A.W., and Keegstra, K. (2007). Functional genomic analysis supports conservation of function among cellulose synthase-likeA gene family members and suggests diverse roles of mannans in plants. *Plant Physiol.* 143, 1881–1893. doi: 10.1104/ pp.106.093989
- Little, A., Schwerdt, J. G., Shirley, N. J., Khor, S. F., Neumann, K., O'Donovan, L. A., et al. (2018) Revised phylogeny of the *Cellulose Synthase* gene superfamily: insights into cell wall evolution. *Plant Physiol.* 177, 1124–1141. doi: 10.1104/ pp.17.01718
- Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178
- Madson, M., Dunand, C., Li, X., Verma, R., Vanzin, G. F., Caplan, J., et al. (2003). The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. *Plant Cell* 15, 1662–1670. doi: 10.1105/tpc.009837
- McCann, M. C., and Roberts, K. (1991). Architecture of the primary cell wall. In CW Lloyd, ed, Cytoskeletal basis of plant growth and form. Academic Press, New York, pp 109–129.
- Molinari, H. B. C., Pellny, T. K., Freeman, J., Shewry, P. R., and Mitchell, R. A. C. (2013). Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in *Brachypodium distachyon. Frontiers Plant Sci.* 4, 50. doi: 10.3389/fpls.2013.00050
- Mortimer, J. C., Miles, G. P., Brown, D. M., Zhang, Z., Segura, M. P., Weimar, T., Yu, X., et al. (2010). Absence of branches from xylan in *Arabidopsis gux* mutants reveals potential for simplification of lignocellulosic biomass. *Proc. Natl. Acad. Sci. USA* 107, 17409–17414. doi: 10.1073/pnas.1005456107
- Mortimer, J. C., Yu, X., Albrecht, S., Sicilia, F., Huichalaf, M., Ampuero, D., et al. (2013). Abnormal glycosphingolipid mannosylation triggers salicylic acidmediated responses in Arabidopsis. *Plant Cell* 25, 1881–1894. doi: 10.1105/ tpc.113.111500
- Norambuena, L., Marchant, L., Berninsone P., Hirschberg, C. B., Silva, H., and Orellana, A. (2002). Transport of UDP-galactose in plants. Identification and functional characterization of AtUTr1, an *Arabidopsis thaliana* UDP-galactose/ UDP-glucose transporter. *J. Biol. Chem.*277, 32923–32929. doi: 10.1074/jbc. M204081200
- Okekeogbu, I. O., Pattathil, S., Fernández-Niño, S. M. G., Penning, B. W., Lao, J., Heazlewood, J. L., et al. (2019). Glycome and proteomic components of Golgi membranes are common between two angiosperms with distinct cell wall structures. *Plant Cell* 31, 1094–1112. doi: 10.1105/tpc.18.00755
- Orellana, A., Moraga, C., Araya, M., and Moreno, A. (2016). Overview of nucleotide sugar transporter gene functions across multiple species. J. Mol. Biol. 428, 3150–3165. doi: 10.1016/j.jmb.2016.05.021
- Penning, B., Hunter, C. T., Tayengwa, R., Eveland, E., Dugard, C. K., Olek, A. et al. (2009). Genetic resources for maize cell wall biology. *Plant Physiol.* 151, 1703– 1728. doi: 10.1104/pp.109.136804
- Penning, B. W., Sykes, R. W., Babcock, N. C., Dugard, C. K., Held, M. A., Klimek, J. F., et al. (2014). Genetic determinants for enzymatic digestion of lignocellulosic biomass are independent of those for lignin abundance in a maize recombinant inbred population. *Plant Physiol.* 165, 1475–1487. doi: 10.1104/pp.114.242446
- Rautengarten, C., Ebert, B., Herter, T., Petzold, C.J., Ishii, T., Mukhopadhyay, A., et al. (2011). The interconversion of UDP-arabinopyranose and UDP-arabinofuranose

is indispensable for plant development in Arabidopsis. Plant Cell 23, 1373-1390. doi: 10.1105/tpc.111.083931

- Rautengarten, C., Ebert, B., Ouellet, M., Nafisi, M., Baidoo, E. E. K., Benke, P., et al. (2012). Arabidopsis Deficient in Cutin Ferulate encodes a transferase required for feruloylation of w-hydroxy fatty acids in cutin polyester. Plant Physiol. 158, 654–665. doi: 10.1104/pp.111.187187
- Rautengarten, C., Ebert, B., Moreno, I., Temple, H., Herter, T., Link, B., et al. (2014). The Golgi localized bifunctional UDP-rhamnose/UDP-galactose transporter family of *Arabidopsis. Proc. Natl. Acad. Sci. USA* 111, 11563–11568. doi: 10.1073/pnas.1406073111
- Rautengarten, C., Birdseye, D., Pattathil, S., McFarlane, H.E., Saez-Aguayo, S., Orellana, A., et al. (2017). The elaborate route for UDP-arabinose delivery into the Golgi of plants. *Proc. Natl. Acad. Sci. USA* 114: 4261–4266. doi: 10.1073/ pnas.1701894114
- Reiter, W. D., and Vanzin, G. F. (2001). Molecular genetics of nucleotide sugar interconversion pathways in plants. *Plant. Mol. Biol.* 47, 95–113. doi: 10.1023/A:1010671129803
- Rose, J. K. C., Braam, J., Fry, S. C., and Nishitani, K. (2002). The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. *Plant Cell Physiol.* 43, 1421–1435. doi: 10.1093/pcp/pcf171
- Rosquete, M. R., Davis, D. J., and Drakakaki, G. (2018). The plant *Trans*-Golgi network: not just a matter of distinction. *Plant Physiol.* 176, 187–198. doi: 10.1104/pp.17.01239
- Rudall, P. J., and Caddick, L. R. (1994). Investigation of the presence of phenolic compounds in monocotyledonous cell walls, using UV fluorescence microscopy. Ann Bot. 74, 483–491. doi: 10.1006/anbo.1994.1145
- Rutherford, R., and Masson, P. H. (1996). Arabidopsis thaliana sku mutant seedlings show exaggerated surface-dependent alteration in root growth vector. Plant Physiol. 111, 987–998. doi: 10.1104/pp.111.4.987
- Saez-Aguayo, S., Rautengarten, C., Temple, H., Sanhueza, D., Ejsmentewicz, T., Sandoval-Ibañez, O., et al. (2017). UUAT1 is a Golgi-localized UDP-uronic acid transporter that modulates the polysaccharide composition of *Arabidopsis* seed mucilage. *Plant Cell* 29, 129–143 doi: 10.1105/tpc.16.00465
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/ oxfordjournals.molbev.a040454
- Scalbert, A., Monties, B., Lallemand, J-Y., Guittet, E., and Rolando, C. (1985) Ether linkage between phenolic acids and lignin fractions from wheat straw. *Phytochemistry* 24, 1359–1362. doi: 10.1016/S0031-9422(00)81133-4
- Scheller, H.V., and Ulvskov, P. (2010). Hemicelluloses. Annu. Rev. Plant Biol. 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315
- Schindelman, G., Morikami, A., Jung, J., Baskin, T. I., Carpita, N. C., McCann, M. C., et al. (2001). The COBRA gene encodes a putative glycosylphosphatidylinositol anchored protein, which is polarly localized and necessary for oriented cell expansion in Arabidopsis. Genes Dev. 15, 1115–1127. doi: 10.1101/ gad.879101
- Schnable, P. S., Ware, D., Fulton, R.S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115. doi: 10.1126/science.1178534
- Sedbrook, J. C., Carroll, K. L., Hung, K. F., Masson, P. H., Somerville, C.R. (2002). The Arabidopsis SKU5 gene encodes an extracellular glycosyl phosphatidylinositol-anchored glycoprotein involved in directional root growth. Plant Cell 14, 1635–1648. doi: 10.1105/tpc.002360
- Sekhorn, R. S., Lin, H., Childs, K. L., Hansey, C. N., Buell, C. R., de Leon, N., et al. (2011). Genome-wide atlas of transcription during maize development. *Plant J.* 66, 553–63. doi: 10.1111/j.1365-313X.2011.04527.x
- Shedletzky, E., Shmuel, M., Delmer, D.P., and Lamport, D.T.A. (1990). Adaptation and growth of tomato cells on the herbicide 2,6-dichlorobenzonitrile leads to production of unique cell-walls virtually lacking a cellulose-xyloglucan network. *Plant Physiol.* 94, 980–987. doi: 10.1104/pp.94.3.980
- Shedletzky, E., Shmuel, M., Trainin, T., Kalman, S., and Delmer, D. (1992). Cell wall structure in cells adapted to growth on the cellulose synthesis inhibitor 2,6-dichlorobenzonitrile: a comparison between two dicotyledonous plants and a gramineous monocot. *Plant Physiol.* 100, 120–130. doi: 10.1104/pp.100.1.120
- Showalter, A. M. (1993). Structure and function of plant-cell wall proteins. *Plant Cell* 5, 9–23. doi: 10.1105/tpc.5.1.9

- Smith, B. G., and Harris, P. J. (1999). The polysaccharide composition of *Poales* cell walls: Poaceae cell walls are not unique. *Biochem. System. Ecol.* 27, 33–53. doi: 10.1016/S0305-1978(98)00068-4
- Smith, P. J., Wang, H.-T., York, W. S., Peña, M. J., and Urbanowicz, B. R. (2017). Designer biomass for next-generation biorefineries: leveraging recent insights into xylan structure and biosynthesis. *Biotechnol. Biofuels* 10, 286. doi: 10.1186/ s13068-017-0973-z
- Springer, N. M., Ying, K., Fu, Y., Ji, T., Yeh, C-T., Jia, Y., et al. (2009). Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. *PLoS Genet.* 5, e1000734. doi: 10.1371/ journal.pgen.1000734
- Stonebloom, S., Ebert, B., Xiong, G., Pattathil, S., Birdseye, D., Lao, J., et al. (2016). A DUF-246 family glycosyltransferase-like gene affects male fertility and the biosynthesis of pectic arabinogalactans. *BMC Plant Biol.* 16, 90. doi: 10.1186/ s12870-016-0780-x
- Swanson-Wagner, R. A., Eichten, S. R., Kumari, S., Tiffin, P., Stein, J. C., Ware, D., et al. (2010). Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. *Genom Res.* 20, 1689–1699. doi: 10.1101/gr.109165.110
- Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T.Z., Garcia-Hernandez, M., Foerster, H., et al. (2008). The *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. *Nucleic Acids Res.* 36: D1009–D1014. doi: 10.1093/nar/gkm965
- Takenaka, Y., Kato, K., Ogawa-Ohnishi, M., Tsuruhama, K., Kajiura, H., Yagyu, K., et al. (2018). Pectin RG-I rhamnosyl-transferases represent a novel plantspecific glycosyltransferase family. *Nat. Plants* 4, 669–676. doi: 10.1038/ s41477-018-0217-7
- Tanaka, K., Murata, K., Yamazaki, M., Onosato, K., Miyao, A., and Hirochika, H. (2003). Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. *Plant Physiol.* 133, 73–83. doi: 10.1104/pp.103.022442
- Trethewey, J. A., Campbell, L. M., and Harris, P. J. (2005). (1→3),(1→4)-β-D-Glucans in the cell walls of the *Poales (sensu lato)*: an immunogold labeling study using a monoclonal antibody. *Am J Bot.* 92, 1660–1674. doi: 10.3732/ ajb.92.10.1660
- Vanzin, G. F., Madson, M., Carpita, N. C., Raikhel, N. V., Keegstra, K., and Reiter W. D. (2002). The *mur2* mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. *Proc Natl Acad Sci* USA 99: 3340–3345. doi: 10.1073/pnas.052450699
- Vega-Sanchez, M., Verhertbruggen, Y., Scheller, H. V., and Ronald, P. (2013). Abundance of mixed linkage glucan in mature tissues and secondary cell walls of grasses. *Plant Signal. Behav.* 8, e23143. doi: 10.4161/psb.23143
- Voiniciuc, C., Engle, K. A., Günl, M., Dieluweit, S., Schmidt, M. H., Yang, J. Y., et al. (2018). Identification of key enzymes for pectin synthesis in seed mucilage. *Plant Physiol.* 178, 1045–1064. doi: 10.1104/pp.18.00584
- Wang, Y., Mortimer, J. C., Jonathan, D., Dupree, P., Keegstra, K. (2013). Identification of an additional protein involved in mannan biosynthesis. *Plant J.* 73, 105–117. doi: 10.1111/tpj.12019
- Wu, A.M., Rihouey, C., Seveno, M., Hörnblad, E., Singh, S.K., Matsunaga, T., et al. (2009). The *Arabidopsis* IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *Plant J.* 57, 718–731. doi: 10.1111/j.1365-313X.2008.03724.x
- Wu, A-M., Hörnblad, E., Voxeur, A., Gerber, L., Rihouey, C., Lerouge, P., and Marchant, A. (2010). Analysis of the *Arabidopsis* IRX9/IRX9-L and IRX14/ IRX14-L pairs of glycosyltransferase genes reveals critical contributions to biosynthesis of the hemicellulose glucuronoxylan. *Plant Physiol.* 153, 542–554. doi: 10.1104/pp.110.154971
- Yin, Y., Huang, J., Gu, X., Bar-Peled, M., and Xu, Y. (2011). Evolution of plant nucleotide-sugar interconversion enzymes. *PLoS One* 6, e27995. doi: 10.1371/ journal.pone.0027995
- Zabotina, O. A., Avci, U., Cavalier, D., Pattathil, S., Chou, Y-H., Eberhard, S., et al. (2012). Mutations in multiple XXT genes of *Arabidopsis* reveal the complexity of xyloglucan biosynthesis. *Plant Physiol.* 159, 1367–1384. doi: 10.1104/ pp.112.198119
- Zdobnov, E.M., and Apweiler, R. (2001). InterProScan: an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17, 847–848. doi: 10.1093/bioinformatics/17.9.847

- Zeng, W., Lampugnani, E. R., Picard, K. L., Song, L. L., Wu, A. M. Farion, I. M., et al. (2016). Asparagus IRX9, IRX10, and IRX14A are components of an active xylan backbone synthase complex that forms in the Golgi apparatus. *Plant Physiol.* 171, 93–109. doi: 10.1104/pp.15.01919
- Zhang, R., Tucker, M. R., Burton, R. A., Shirley, N. J., Little, A., Morris, J., et al. (2016). The dynamics of transcript abundance during cellularization of developing barley endosperm. *Plant Physiol.* 170, 1549–1565. doi: 10.1104/pp.15.01690
- Zhong, R., Peña, M. J., Zhou, G. K., Nairn, C. J., Wood-Jones, A., Richardson, E. A., et al. (2005). Arabidopsis fragile fiber8, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. *Plant Cell* 17, 3390–3408. doi: 10.1105/tpc.105.035501

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Penning, McCann and Carpita. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.