diversity

MDPI

Article

The Cell Wall-Related Gene Families of Wheat (*Triticum aestivum*)

Bryan W. Penning

Corn, Soybean, and Wheat Quality Research Unit, Agricultural Research Services, United States Department of Agriculture, 1680 Madison Ave., Wooster, OH 44691, USA; bryan.penning@usda.gov

Abstract: Wheat crops provide 20% of calories worldwide. Cell walls function in plant growth, are part of biotic and abiotic stress resistance, and provide plant mechanical strength and adaptability. These functions factor into the productivity of wheat. The genes that produce and maintain the plant cell wall are up to 10% of the genome in many varied families. Previously, curated cell wall gene families have been published for maize and rice, two other important crop grasses. Here, 81 cell wall-related wheat gene families curated via sequence similarity to maize and rice and unique family protein motif searches are presented. A total of 4086 wheat, 1118 maize, 1036 rice, and 955 Arabidopsis genes were aligned and placed into gene family trees to present homologs for all four species. Due to hexaploidy, many wheat cell wall gene families show expected triplication of genes per family over maize, rice, and Arabidopsis. However, several families contained more wheat genes than expected. The utility of this research is demonstrated with an example from a pre-harvest sprouting study to identify specific gene families rather than the less descriptive identification available with standard bioinformatic searches.

Keywords: cell wall genes; wheat; maize; rice; Arabidopsis

Citation: Penning, B.W. The Cell Wall-Related Gene Families of Wheat (*Triticum aestivum*). *Diversity* **2023**, *15*, 1135. https://doi.org/10.3390/d15111135

Academic Editor: Mario A. Pagnotta

Received: 25 September 2023
Revised: 3 November 2023
Accepted: 6 November 2023
Published: 7 November 2023

1. Introduction

Wheat accounts for twenty percent of worldwide calories [1]. Seven hundred and ninety-three million metric tons of wheat were produced worldwide in 2023. The top five producers were China, the European Union, India, Russia, and the United States [2]. In addition to human consumption, wheat grain that does not meet human use quality standards can be consumed by animals [3]. One factor leading to the loss of wheat grain quality is pre-harvest sprouting, which has become more prevalent with climate change [4]. Plant cell wall genes comprise about ten percent of a plant genome and are important for growth and development, pathogen resistance, abiotic stress, food quality, conversion of plant material for fuels, and ruminant animal digestion [5–8]. Plant cell walls may affect flour quality and pre-harvest sprouting in wheat [9,10]. Cell wall-related genes representing almost ten percent of differentially expressed genes between a treated and untreated pre-harvest sprouting-resistant wheat variety were reported for the first time. However, some of their descriptions from bioinformatic searches were too vague for assignment to gene families and thus function [10]. Cell wall genes have been curated for maize, rice, and Arabidopsis into families with unique protein motifs to identify them. However, a curated accounting for all cell wall-related wheat genes and family affiliation with insight into function is lacking [11–15].

Plant cell walls are built from a multitude of carbohydrate molecules assembled into complex structures with over 80 gene families encoding >1000 genes to produce their components and maintain their structure and function throughout the plant's life. All plant cell walls have in common cellulose microfibrils forming the bulk of the cell wall structure linked together by other carbohydrate molecules. Primary cell walls vary based on species, with major architectural differences between commelinids such as grasses with Type II

primary cell walls and dicots/noncommelinids with Type I primary cell walls. Type II primary cell walls contain a large portion of glucoarabinoxylans (GAXs) linking cellulose structures reinforced with phenylpropanoids. Type I primary cell walls contain a large portion of xyloglucans (XyG) and pectic polysaccharides such as homogalaturonan (HG) and rhamnogalacturonan (RG) I, linking the cellulose structure reinforced with Hyp- and Gly-rich proteins [16]. In addition to a primary cell wall, plants often have a secondary one. The secondary cell walls of plants are thickened and contain high amounts of lignin to provide rigidity in structures such as the stem and make up the bulk of the biomass used in bioenergy [17]. There are many groups of glycosyl transferases (GT) and glycosyl hydrolases (GH) and a few groups of polysaccharide lyases (PL) and carbohydrate esterases (CEs) with roles in the production and maintenance of plant cell walls. The families are described on the CAZY website [11,18].

There are many conserved functions within the cell wall of species such as maize, rice, and Arabidopsis, including the production of cellulose, hemicelluloses, and lignin, covered in detail in earlier publications but much less in wheat. Many important family groups carry out specific purposes or encode biochemical pathways. The nucleotide sugar transporters form a large six-group gene family in plants that transfer sugars to proteins, lipids, and polysaccharides in the Golgi [19]. The phenylpropanoid pathway (PP) encodes the biosynthetic pathway to produce monolignols, the building blocks for lignin, and includes ten gene families: 4-coumarate CoA ligase (4CL), *p*-coumaroyl shikimate 3′-hydroxylase (C3′H), cinnamate 4-hydroxylase (C4H), ferulate (coniferaldehyde) 5-hydroxylase (F5H), (hydroxy)cinnamyl alcohol dehydrogenase (CAD), caffeoyl CoA O-methyltransferase (CCoAOMT), caffeic acid (5-hydroxy-coniferaldehyde) O-methyltransferase (COMT), (hydroxy)cinnamoyl CoA reductase (CCR), hydroxycinnamoyl CoA:shikimate hydroxycinnamoyl transferase (HCT), and phenylalanine ammonia lyase (PAL) [20].

The GT families form a large set of cell wall-related genes. The sucrose synthases (GT4) and nucleotide–sugar interconversion pathway (GT1) form the basic sugar building blocks, later linked to many cell wall carbohydrates. The nucleotide–sugar interconversion pathway families include: membrane-anchored UDP-D-glucuronate decarboxylase (AUD) and soluble UDP-D-glucuronate decarboxylase (SUD), UDP-D-glucose 4-epimerase (AXS), UDP-D-glucuronate 4-epimerase (GAE), GDP-4-keto-6-deoxy-D-mannose 3,5-epimerase-4-reductase (GER), GDP-D-mannose-4,6-dehydratase (GMD), GDP-D-mannose 3,5-epimerase (GME), UDP-L-rhamnose synthase (RHM), UDP-4-keto-6-deoxy-D-glucose 3,5-epimerase-4-reductase (UER), UDP-D-glucose dehydrogenase (UGD), UDP-D-glucose 4-epimerase (UGE), and UDP-D-xylose 4-epimerase (UXE) genes [21,22]. The GT2 family has members that produce many carbohydrate backbones, including the cellulose synthase family (CesA), which produces cellulose everywhere, and the cellulose synthase-like (Csl) D family, which produces cellulose in root hairs and pollen tubes [23,24]. The CslA family synthesizes the (1,4)-β-D-mannan backbone of galactomannan. The β-1,4 glucan backbone of XyG is produced by the CslC family. The CslF family synthesizes a grass cell wall unique mixed linkage (1,3;1,4)-β-D-glucan [25–27]. Products of the Csl E, G, and H/J families are still unknown [28]. The GT48 family encodes callose synthases, which synthesize callose from UDP-glucose. Callose production is important for the development of mature cell walls as well as responses to biotic and abiotic stresses [29]. Many other GT families either have few members or few of their members have been characterized. The GT14 family is not well characterized, but one member is a glucuronosyl transferase involved in the cell wall elongation of seedlings in Arabidopsis. The GT61 family are arabinosyl, arabinofuranosyl, and xylosyl transferases involved in producing arabinoxylan and GAX from a xylan backbone [30–32]. The GT77 family is a group of arabinosyl transferases [33]. The GT8 family is a large group of genes subdivided into five groups; some are galacturonosyl transferases linked to pectin biosynthesis [34]. The gene At3g61130_GAUT1 from group D of GT8 is thought to synthesize HG [35], while the *PARVUS* gene, At1g19300_GATL1, is necessary for glucuronoxylan synthesis [36]. The GT31 family is a six-group family with one member having a galactosyl transferase activity and other members potentially involved in building

the galactan backbone of arabinogalactan proteins (AGPs) [37,38]. Two genes from the GT34 family, At3g62720_XXT1 and At4g02500_XXT2, encode xylosyltransferases specific for XyG biosynthesis [39]. The GT37 family encodes fucosyltransferases in Arabidopsis, including At2g03220.1_FUT1 [40,41]. The GT47 family is a large set of genes in five groups whose members can have galactosyl—(At2g20370.1_MUR3), arabinosyl—(At2g35100.1_ARAD1), xylosyl—(At5g33290.1_XGD1), or glucuronosyl—(At5g22940.1 and *fra8*) transferase activity [42–46]. The genes At1g27440_IRX10 and At5g61840_IRX10-like from GT47 are needed for xylan backbone synthesis [47].

Glycosyl hydrolases in the cell wall provide several functions in plant growth, including defense and reproduction. The GH16 family, xyloglucan endotransglucosylase/hydrolase (XTH), and expansin families both function in the loosening of cell walls during growth [48,49]. The GH17 family encodes β-(1,3)-glucanases that appear involved in fugal defense and seed germination [50,51]. The GH18 family are classified as chitinases that degrade fungal cell walls but may also function in cell wall loosening [52–54]. The GH9 family has been characterized in bacteria where they function as endo-β-(1,4)-glucanases [11,55]. The GH28 family encodes polygalacturonases (PGases) that are highly expressed during seed pod dehiscence in *Arabidopsis thaliana* and *Brassica napus* [56]. The GH35 family encodes β-galactosidases. A mutation in the At5g63810.1_BGAL10 gene leads to shorter siliques and an altered XyG in Arabidopsis [57].

Some CE and PL cell wall-related genes have been observed to alter the composition or molecular bonding for the growth and development of cell walls in different tissues. The gene family CE13 (pectin acetylesterases, PAEs) regulates the degree of pectin acetylation by cleaving acetylester bonds from pectin [58,59]. The CE8 family (pectin methylesterases, PMEs) demethylesterifies pectin, which can increase cell wall strength but also act with PGases during fruit ripening to degrade the pectin matrix, softening the fruit [60]. A pectate lyase gene in the PL1 family was observed to be highly upregulated during induced differentiation of cultured Zinnia mesophyll cells. Gene expression was associated with vascular bundles, shoot primordia, and pectate lyase activity found in elongating and differentiating cells in vitro [61]. The PL4 family encodes RG lyase activity, cleaving the bonds between rhamnose and galacturonic residues of the plant cell wall pectin RG I. In tomatoes, a PL4 gene was observed to affect fruit firmness during ripening and pollen tube elongation [62].

There are gene families involved in modifying the cell wall functioning to strengthen cell walls or prepare them for different stages of growth and development through cell wall remodeling. While the biological role of proteases is not yet well understood, they are important to degrade other proteins as cells undergo development, such as seed germination or response to the environment through abiotic or biotic stress [63]. Laccases polymerize monolignols into larger subunits in the early stages of lignification. Peroxidases are involved in lignification with the presence of hydrogen peroxide [64]. The COBRA gene family alters cell wall mechanical strength. Mutants in At5g15630.1_COBL4 lead to normal-looking plants with reduced stem strength, while mutations in the closest rice homolog, Os03g30250.1_BC1, lead to reduced cell wall thickness and brittleness of plant organs [65,66]. The prolyl-4-hydroxylase (P4H) gene family appears to be involved in proper cell wall organization and architecture by establishing sites for O-glycolylation in root hairs [67]. The Arabidopsis skewed-growth proteins (SKU) are involved in directional root growth through cell wall expansion and are not gravity-dependent [68,69].

The hexaploid wheat genome consists of three complete seven-chromosome subgenomes (A, B, and D) with a high sequence conservation between homologous genes that have arisen through several rounds of genome hybridization [70]. In addition to differences in ploidy level, genes can be duplicated in a genome and, over time, develop new functions through the mutation of one gene copy (neofunctionalization). In some cases, both gene copies mutate to each perform part of the function of the original gene (subfunctionalization). This has been illustrated in the Arabidopsis and rice monosaccharide transporter gene families [71]. In addition to partitioning multiple functions from one gene

into two, neofunctionalized expression or subfunctionalized expression can occur through expression differences brought about by changes in the promoter regions of duplicated genes instead of the protein-coding region [72].

The objectives of this research were to provide a curated set of wheat cell wall-related genes complementing those previously published for maize and determine how the wheat homologs are related by sequence to the cell wall genes of other species. Arabidopsis, maize, and rice often have genes of defined function in cell wall-related families. However, the function is not as often defined in wheat. A potential function for wheat homologs or at least the correct family classification would improve the understanding of how those wheat genes might function when discovered through bioinformatics analyses. A recent bioinformatic study on pre-harvest sprouting in wheat was used to illustrate the utility of this research to improve gene classification and how they might function. While preparing the families of cell wall-related wheat genes, it was discovered that more wheat genes were present than expected due to hexaploidy, and many tandem duplications of genes occurred in these wheat gene families. They are discussed in greater detail.

2. Materials and Methods

Searching for cell wall-related wheat genes and building trees based on sequence similarity for each gene family was performed as previously described [13]. The following methods were used without any updates to the process. Previously published cell wall-related maize sequences were used to search version 2.1 of the reference wheat genome from the International Wheat Genome Sequencing Consortium [73]. The search was performed by downloading and formatting the wheat V2.1 high-confidence and low-confidence (LC) protein fasta files as databases with formatdb and performing protein–protein BLASTs using blastall-blastp from NCBI in a custom batch file under Windows PowerShell [74]. A Perl script version of the original C++ program was used to retrieve wheat gene names considered significant via a protein alignment score of 200 or more to the maize query genes and place wheat gene names in a text file. The NCBI fastacmd under Windows PowerShell was used to retrieve the full-length wheat protein sequences from the same wheat high-confidence and LC protein databases. At this point, all gene models were kept as indicated by ".#" where different protein sequence models are numbered sequentially after a period and the gene name. To recover the names of V2.1 wheat cell wall genes from V1.0, a similar strategy was used. Version 1 gene sequences were recovered via fastacmd from its database and then used in blastall -blastp from NCBI to find V2.1 genes. The top hit was taken and was always on the same chromosome, with expected scores near zero and protein alignment scores above 1000.

Previously published rice (release 7), maize (version 2.59a), and Arabidopsis (TAIR10) sequences were added to the wheat sequences for each gene family. The websites to download each of the protein sequence files are wheat [75], maize [76], rice [77], and Arabidopsis [78]. Gene names were kept from the files, except LOC_, which was removed from the front of rice genes, and abbreviations of known genes or a group designation were added to the end of the gene behind an underscore (_). Trees were made via slow/accurate sequence alignment using the neighbor-joining method followed by bootstrapping 1000 times with bootstrap numbers positioned at the nodes in clustalW [79–81]. Individual trees were visualized in TreeDyn with a leaf length scaled to substitutions per site and exported to postscript files to be converted to PDFs using Adobe Acrobat [82]. In some cases, large gene families were broken apart by groups previously published [12,13,17] to make their trees easier to read. In cases of ambiguity as to whether to include a gene in the family, the proper protein motifs were verified using the EBI protein motif viewer [14,83] or PROSITE [15,84] as compared to the closest rice, maize, or Arabidopsis gene. To determine the best gene model or test for full length and correct family placement, multiple wheat genes were aligned using Clustal Omega [85] with the closest rice or maize gene by the initial tree. The protein identity of at least 60% was used as a cutoff along with the protein motif analysis, which in all accepted cases identified the family-specific motifs [86]. After confirming

the proper inclusion in the gene family and the best protein model, the final trees were produced using clustalW and TreeDyn as above.

Tandem duplications of wheat genes were chosen from Supplemental File S1 based on their assigned number after the "G" in the name, which lists the genes in numerical order along the chromosome and their closeness on the tree. Physical locations were found by downloading the GFF3 file from the International Wheat Genome Sequencing Consortium for version 2.1 of the wheat sequence, high confidence or LC, as appropriate, and searching for the gene name in Microsoft WordPad [73]. The start and stop locations in the basepairs (bp) of the gene were used. The distance between the genes was calculated by subtracting the stop location of the previous gene from the start location of the next gene numerically on the list and then dividing the value by 1000 to convert to kilobases (kb). Distances were reported between gene 1 and gene 2, then gene 2 and gene 3, etc., in order until the end of the repeated genes along the chromosome location. Figures were produced by colorizing potential tandem duplicated genes in TreeDyn, exporting them to PDFs, loading the PDF tree files into Adobe Photoshop, adding any arrows to indicate potential tandem duplications, and exporting them as TIFF or JPEG images.

3. Results and Discussion

A total of 4086 wheat genes across 81 cell wall-related families were annotated and performed by searching Chinese spring wheat genome protein sequences with maize cell wall family genes followed by manual curation using sequence alignment and protein motif calling. Wheat has a hexaploid genome comprising three sub-genomes (A, B, and D). Thus, three homologous wheat genes (Traes) were expected for every maize (GRMZM or AC), rice (Os), and Arabidopsis (At) gene. The number of wheat genes was several hundred more than would be predicted based on the 1118, 1036, and 955 total genes from maize, rice, and Arabidopsis, respectively (Table 1). The specific versions of the maize, rice, and Arabidopsis genes were chosen because they had previously been curated as cell wall-related [12,13,17]. Many of the increased wheat genes were from families PP, GT2 (Csl), GT14, GT48 (callose synthase), GT61, GT77, expansin, laccase, peroxidase, protease, and most GH families. Other gene families remain close to an expected 3:1 ratio of wheat compared to the diploid species (Table 1). The trees of all cell wall-related gene families for the four species were bootstrapped 1000 times to indicate likely groupings from clustalW and are available as Supplemental File S1 in a PDF format to find species homologs. Examples of gene family differences are highlighted in the figures and described below. All gene names, families, and species were listed in an Excel file in Supplemental File S2 to aid gene-of-interest searches. High-confidence wheat genes, defined by the IWGSC, have known gene models from other plants and expression in wheat tissues. Low-confidence wheat genes with LC following their names contain the appropriate gene family protein motifs but unknown expression from the IWGSC [73].

Table 1. Total numbers of wheat, maize, rice, and Arabidopsis genes for each cell wall family.

	Number of Genes per Family			
Cell Wall Family	**Wheat**	**Maize**	**Rice**	**Arabidopsis**
Nucleotide Sugar Transferases	133	63	50	51
PP 4CL	45	10	14	13
PP C3H_C4H_F5H	50	8	10	6
PP CAD	48	4	12	9
PP CCoAOMT	21	6	6	7
PP COMT	33	3	7	16
PP CCR	121	18	24	7
PP HCT	149	38	5	1
PP PAL	47	10	9	4

Table 1. *Cont.*

	Number of Genes per Family			
Cell Wall Family	**Wheat**	**Maize**	**Rice**	**Arabidopsis**
GT1–NSI–AUD_SUD	15	9	6	6
GT1–NSI–AXS	3	1	1	2
GT1–NSI–GAE	15	9	4	6
GT1–NSI–GER	3	1	1	2
GT1–NSI–GMD	6	1	1	2
GT1–NSI–GME	6	2	2	1
GT1–NSI–RHM_UER	8	4	3	4
GT1–NSI–UGD	12	3	5	4
GT1–NSI–UGE	9	3	4	5
GT1–NSI–UXE	9	4	3	4
GT2–CesA	28	20	10	10
GT2–CSL A	38	10	11	9
GT2–CSL C	15	8	6	5
GT2–CSL D	15	5	5	5
GT2–CSL E	12	2	3	1
GT2–CSL F	18	7	8	0
GT2–CSL G & H	11	1	3	3
GT4–Sucrose Synthases	21	5	7	6
GT8–A	15	7	4	5
GT8–B	13	2	2	8
GT8–C	21	10	8	10
GT8–D	52	23	22	15
GT8–E	12	5	5	3
GT10	6	3	3	3
GT13	6	1	1	1
GT14	44	10	11	9
GT16	3	1	1	1
GT17	11	4	4	6
GT21	3	1	1	1
GT22	9	6	4	3
GT24	3	2	1	1
GT29	12	8	5	3
GT30	3	1	1	1
GT31	107	40	39	33
GT32	6	2	2	6
GT33	6	2	2	1
GT34	24	17	7	6
GT37	53	17	18	10
GT41	6	2	3	2
GT43	27	16	10	4
GT47	126	50	38	39
GT48–Callose Synthases	35	7	7	12
GT50	2	1	1	1
GT57	3	1	1	1
GT58	3	1	1	1
GT59	3	1	1	1
GT61	117	33	24	8
GT64	6	3	3	3
GT66	9	4	2	2
GT68	3	1	1	3
GT75–RGP	12	9	3	5
GT76	3	2	1	1
GT77	74	23	16	18
GT92	9	3	3	3

Table 1. *Cont.*

	Number of Genes per Family			
Cell Wall Family	**Wheat**	**Maize**	**Rice**	**Arabidopsis**
GT96	3	2	1	1
GH9	82	22	25	25
GH16–XTHs	142	31	28	33
GH17	209	51	60	47
GH18–Yieldins	60	10	9	10
GH28–Pgases–A	29	12	11	13
GH28–Pgases–B	18	0	1	10
GH28–Pgases–C	11	8	8	9
GH28–Pgases–D	34	7	9	6
GH28–Pgases–E	15	3	3	8
GH28–Pgases–F	14	16	6	10
GH28–Pgases–G	13	1	4	8
GH35–Beta Galactosidases	48	16	10	17
Expansins	248	55	61	34
CE8–PMEs	120	34	41	66
CE13–PAEs	53	11	10	11
PL1–Pectin Lyases	31	12	12	26
PL4–RG Lyases	15	4	3	7
P4H	28	10	12	11
Laccases	113	24	27	17
Peroxidases	643	124	135	73
SKU	34	12	4	3
COBRA-like	38	9	11	11
GDPlike	15	7	6	7
Proteases	296	49	50	18
HIP-like	11	3	3	3
AGPs	22	9	8	39
HRGPs	3	2	2	13
Total	**4086**	**1118**	**1036**	**955**

The monolignol biochemical pathway is encoded by the PP family in both monocots and dicots. PAL is the first committed step in the process, while CAD is the last step, producing either sinapyl alcohol (syringyl lignin subunit) or coniferyl alcohol (guaiacyl lignin subunit) [20]. The CAD gene family includes 9 Arabidopsis, 12 rice, and 4 maize genes. However, wheat had 48, which was four times that of rice and five times that of Arabidopsis genes. At the top of Figure 1, fifteen wheat genes are distributed among the A, B, and D sub-genomes on chromosomes 6 and 7, closely related to two rice and one maize gene: Os08g16910.1, Os10g29470.1, and GRMZM2G118610_P01. The wheat gene naming convention indicates genes close to each other on the physical map based on the numbers after the "G" being or nearly sequential. The numbers can be nearly sequential if gene models are later removed [73]. Rice and Arabidopsis genes follow the same convention. At least two wheat genes close on the physical map were also near each other on the gene family tree, indicating similar protein sequences for wheat chromosomes 2A, 2B, 3A, 3B, 3D, 6A, 6B, and 6D, an indication of tandem duplications [71]. For example, the physical locations for TraesCS6D03G0816500, TraesCS6D03G0816600, and TraesCS6D03G0817100 were ~26 kb away from each other on the IWGSC 2.1 reference sequence GFF3 map (Table 2) [73]. Three CAD genes were on the 6A and 6B chromosomes, ranging from 17.6 to 43.3 kb apart (Figure 1, Table 2). Gene duplications may be present in the middle of Figure 1 with nine wheat genes on chromosomes 2A, 2B, and 2D for one rice gene, Os04g15920.1. TraesCS2A03G0131200 and TraesCS2A03G0131300 were 19.9 kb apart, while TraesCS2B03G0191100 and TraesCS2B03G0191300 were 33.7 kb apart. Potential duplications of CAD genes were found on chromosomes 3A, 3B, and 3D from 34 to 321 kb apart (Figure 1, Table 2). As the last enzyme in the production of monolignol precursors for lignin, increased total lignin or more tissue-selective lignin deposition may occur in wheat.

The PAL family had approximately five times the wheat genes of maize and rice and twelve times that of Arabidopsis. There were multiple tandem duplications on the A, B, and D sub-genomes of chromosomes 1 and 2, with 1.4 to 268 kb between each duplicated gene (Table 2). Increased numbers of wheat genes with many tandem duplications compared to the other three species were also observed in HCT, CCR, COMT, C3′H, C4H, F5H, and 4CL.

Table 2. Location of select tandem duplicated wheat genes.

Gene Family	Wheat Gene	Location (bp) Start	Stop	Distance Apart (kb)
CAD	TraesCS2A03G0131200	34,450,307	34,451,760	19.9
	TraesCS2A03G0131300	34,471,637	34,473,153	
CAD	TraesCS2B03G0178200	50,262,749	50,264,076	78.9
	TraesCS2B03G0178300	50,342,947	50,344,421	
CAD	TraesCS2B03G0191100	54,229,490	54,230,578	33.8
	TraesCS2B03G0191300	54,264,332	54,266,016	
CAD	TraesCS3A03G1041000	688,689,431	688,692,028	34.1
	TraesCS3A03G1041600	688,726,119	688,729,557	
CAD	TraesCS3B03G1199000	745,631,611	745,634,419	321.7
	TraesCS3B03G1200700	745,956,098	745,958,925	
CAD	TraesCS3D03G0968100	552,033,322	552,035,463	116.5
	TraesCS3D03G0968500	552,151,988	552,155,191	
CAD	TraesCS6A03G0939400	597,869,237	597,875,251	17.6
	TraesCS6A03G0939600LC	597,892,856	597,894,201	21.3
	TraesCS6A03G0939900	597,915,465	597,921,915	
CAD	TraesCS6B03G1146700	690,063,231	690,067,770	26.8
	TraesCS6B03G1147100	690,094,598	690,099,654	43.3
	TraesCS6B03G1147200	690,142,991	690,147,748	
CAD	TraesCS6D03G0816500	470,593,540	470,597,691	26.6
	TraesCS6D03G0816600	470,624,314	470,629,321	26.0
	TraesCS6D03G0817100	470,655,366	470,661,613	
PAL	TraesCS1A03G0089500	22,809,328	22,817,055	2.8
	TraesCS1A03G0089700	22,819,864	22,822,522	
PAL	TraesCS1B03G0108100	30,156,990	30,159,840	13.2
	TraesCS1B03G0108300	30,173,031	30,175,842	48.2
	TraesCS1B03G0108400	30,224,080	30,226,937	67.0
	TraesCS1B03G0108600	30,293,907	30,296,751	37.8
	TraesCS1B03G0108700	30,334,559	30,337,342	
PAL	TraesCS1D03G0078900	20,552,689	20,561,537	1.4
	TraesCS1D03G0079200	20,562,959	20,565,612	
PAL	TraesCS2A03G0922600	628,113,634	628,116,620	32.0
	TraesCS2A03G0922700	628,148,609	628,151,525	49.4
	TraesCS2A03G0922800	628,200,915	628,203,659	8.4
	TraesCS2A03G0922900	628,212,066	628,214,916	
PAL	TraesCS2B03G1014000	572,920,722	572,923,789	145.0
	TraesCS2B03G1015000	573,068,831	573,071,748	140.2
	TraesCS2B03G1015300	573,211,969	573,214,748	
PAL	TraesCS2D03G0862600	483,823,970	483,826,687	84.6
	TraesCS2D03G0863100	483,911,277	483,914,370	268.0
	TraesCS2D03G0863200	484,182,377	484,185,488	23.8
	TraesCS2D03G0863300	484,209,252	484,211,934	
CslA	TraesCS7A03G0953700	576,059,990	576,064,129	49.7
	TraesCS7A03G0953800	576,113,787	576,117,206	

Table 2. *Cont.*

Gene Family	Wheat Gene	Location (bp) Start	Stop	Distance Apart (kb)
CslA	TraesCS7B03G0796700	537,109,006	537,112,692	3.1
	TraesCS7B03G0796800	537,115,750	537,132,129	
CslA	TraesCS7D03G0919100	506,334,490	506,337,868	9.5
	TraesCS7D03G0919200	506,347,416	506,352,132	
CslE	TraesCS5A03G0631600	469,521,391	469,527,274	2.4
	TraesCS5A03G0631700	469,529,676	469,535,899	
CslE	TraesCS5B03G0653700	437,270,468	437,275,686	2.7
	TraesCS5B03G0653800	437,278,393	437,284,101	
CslE	TraesCS5D03G0599500	370,151,008	370,156,445	4.7
	TraesCS5D03G0599600	370,161,140	370,165,637	
GH18	TraesCS3A03G0882000	623,487,813	623,487,813	6.4
	TraesCS3A03G0882100	623,494,204	623,495,097	2.5
	TraesCS3A03G0882200	623,497,609	623,498,502	137.5
	TraesCS3A03G0882500	623,636,002	623,636,895	13.2
	TraesCS3A03G0882600	623,650,108	623,651,001	6.9
	TraesCS3A03G0882700	623,657,857	623,658,750	26.5
	TraesCS3A03G0882800	623,685,242	623,686,135	214.1
	TraesCS3A03G0883100	623,900,253	623,901,602	
GH18	TraesCS3B03G1008300	655,681,956	655,682,987	4.1
	TraesCS3B03G1008400	655,687,134	655,688,027	27.7
	TraesCS3B03G1008600	655,715,751	655,716,644	6.3
	TraesCS3B03G1008700	655,722,943	655,723,836	6.7
	TraesCS3B03G1009100	655,730,571	655,731,464	122.3
	TraesCS3B03G1009200	655,853,730	655,854,731	
GH18	TraesCS3D03G0810000	480,870,250	480,871,131	3.9
	TraesCS3D03G0810100	480,875,009	480,875,902	31.7
	TraesCS3D03G0810300LC	480,907,624	480,908,268	199.2
	TraesCS3D03G0810400	481,107,472	481,108,365	16.6
	TraesCS3D03G0810500	481,124,934	481,125,827	457.3
	TraesCS3D03G0810900	481,583,081	481,584,377	
PAE	TraesCS3A03G1263600	748,276,487	748,280,357	53.1
	TraesCS3A03G1264000	748,333,460	748,335,357	20.7
	TraesCS3A03G1264100	748,356,041	748,358,835	3.3
	TraesCS3A03G1264200	748,362,146	748,365,282	8.5
	TraesCS3A03G1264400	748,373,743	748,377,054	
PAE	TraesCS3B03G1508100	844,439,077	844,442,856	27.3
	TraesCS3B03G1508200	844,470,116	844,472,978	28.1
	TraesCS3B03G1508300	844,501,107	844,509,112	13.6
	TraesCS3B03G1508500	844,522,685	844,524,017	22.3
	TraesCS3B03G1508700	844,546,309	844,549,502	
PAE	TraesCS3D03G1188000	615,245,943	615,249,493	29.9
	TraesCS3D03G1188400	615,279,382	615,282,166	9.9
	TraesCS3D03G1188500	615,292,106	615,295,886	2.7
	TraesCS3D03G1188600	615,298,614	615,301,615	31.6
	TraesCS3D03G1188700	615,333,190	615,336,310	40.9
	TraesCS3D03G1188900	615,377,258	615,380,183	
PAE	TraesCS5A03G1140500	658,736,919	658,740,227	72.9
	TraesCS5A03G1140600	658,813,140	658,817,069	
PAE	TraesCS5B03G1211700	666,788,941	666,792,805	127.0
	TraesCS5B03G1212500	666,919,777	666,924,347	
PAE	TraesCS5D03G1094700	532,276,790	532,280,796	29.9
	TraesCS5D03G1094800	532,310,668	532,314,929	

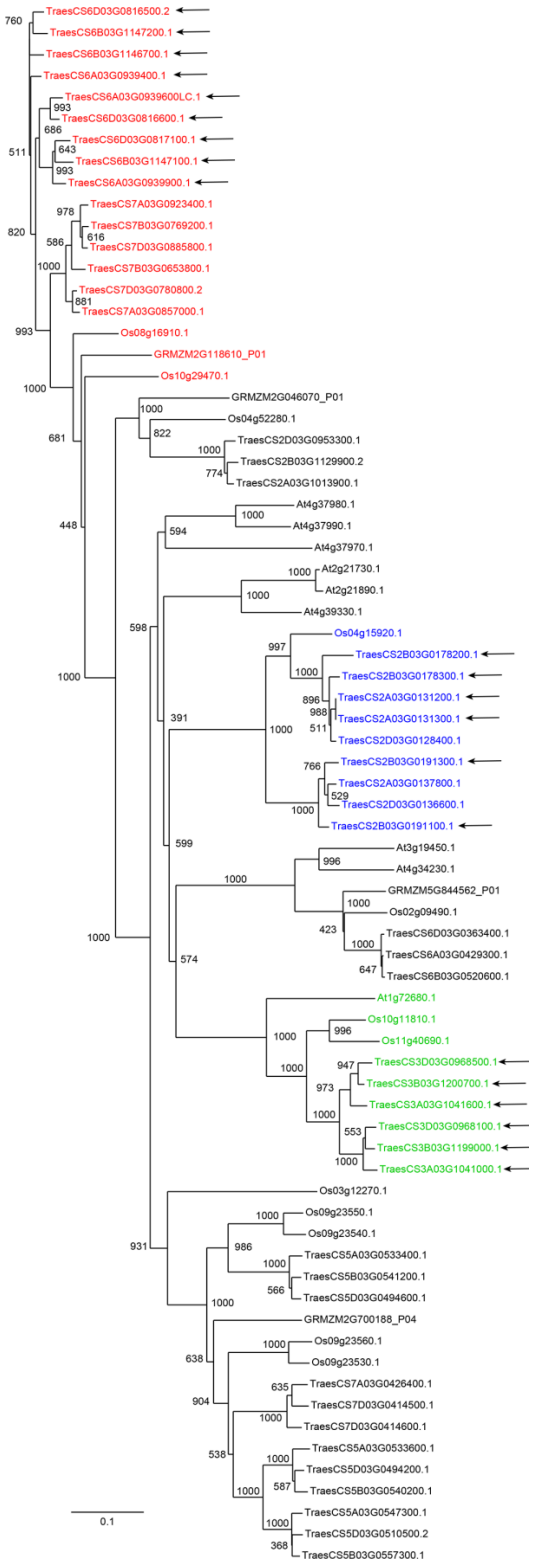

Figure 1. The CAD gene family. Wheat CAD genes are more numerous than expected. Fifteen wheat genes are closely associated with just two rice and one maize gene (red). Nine wheat genes are associated with just one rice gene (blue). Six wheat genes are associated with one Arabidopsis and two rice genes (green). Arrows indicate twenty-one genes that appear to be tandem duplications along the wheat chromosomes. The scale is the number of substitutions per site. The numbers are branching based on 1000 bootstraps.

The GT2 gene families produce many carbohydrate molecules, such as cellulose, via the CesA family. Maize CesA genes were previously reported as duplicated and expressed in the stem, numbering 20 as opposed to 10 for other diploid species [13,17]. The wheat CesA gene family did not show duplication with 28 members, 2 less than expected from hexaploidy (Figure 2). However, Csl A, E, G, and H have more members than expected. The CslA family synthesizes the galactomannan (1,4)-β-D-mannan backbone (Figure 3) [27]. A duplication of three wheat genes on 7A, 7B, and 7D closest in sequence to maize GRMZM2G443715_P01 with distances of 3.1, 9.5, and 49.7 kb apart from one another was observed (Table 2). Two sets of wheat genes on different chromosomes are near GRMZM2G107754_P02. Both maize and wheat show duplicated sequences closest in sequence to Os08g33740.1 (Figure 3). On the other hand, Csl E (Figure 4A), Csls G, and H (Figure 4B), whose functions are unknown, showed wheat sequence duplications at the periphery of the other family members [28]. The CslE 5A, 5B, and 5D distances between genes are 2.4, 2.7, and 4.7 kb, respectively (Table 2).

Thirty-five GT48 callose synthase genes were observed in wheat, over three times those in maize and rice. However, it was about three times the number of callose synthases in Arabidopsis. Two clades in Figure 5 had two Arabidopsis genes but only one maize and rice gene and at least two sets of wheat genes. Arabidopsis genes At4g04970.1_AtGSL1 and At4g03550.1_AtGSL 5 were both on chromosome 4 with two sets of wheat genes (3A, 3B, and 3D) but one rice and maize gene. Arabidopsis genes At1g06490.1_AtGSL7 and At3g59100.1_AtGSL 11 were close to wheat genes in 2A, 2D, 3A, 3B, 3D, 5B, and 5D with only one maize and rice gene. Callose coats the surface of cell plates as cell walls mature and can have many effects. Callose deposited in plasmodesmata helped determine the size exclusion limit and plasmodesmata permeability affecting the stomata and sieve development and transportation of phloem. Callose deposition can alter cell wall permeability during stress events such as pathogen attack, chilling stress, water stress, heat, or induction of heavy metals to reduce plant damage [29].

Commelinid monocot cell walls have lower amounts of pectic polysaccharides than dicots and were observed to have fewer genes that modify them [13,16]. The PL1 (pectin and pectate lyases), PL4 (RG I lyases), and CE8 (PMEs), while larger in wheat, have a ratio of ~2:1 compared to Arabidopsis, indicating that the hexaploidy of wheat did not lead to the complexity of genes found in the model dicot (Table 1). One notable exception was the gene family CE13 (PAEs). Wheat had 53 genes, five times the number of maize, rice, and Arabidopsis genes. The PAE family functions to cleave the acetyl ester bond from pectin but not xylan, thus regulating the degree of pectin acetylation [58,59]. Overexpression of a PtPAE1 in tobacco led to frequent wilting of apical buds reduced the lengths of floral styles and filaments, and had leaves with curled edges and wavy surfaces. The levels of pectin acetylation were lower in overexpressed lines, while xylose acetylation was unaffected, and pectin methylester content was potentially increased to compensate for the altered pectin. The lengths of floral epidermal cells were reduced, indicating severe impairment in cell elongation [58]. Other research indicated that one clade of Arabidopsis PAEs had high expression under water and salt stress and was involved in abiotic stress response [59]. Pre-harvest sprouting is caused by too much water at the maturation of the seed on the spike. A total of 3 maize and rice genes and 19 wheat genes, 16 on chromosomes 3A, 3B, and 3D, formed another clade away from Arabidopsis. The wheat genes were between 2.7 and 53.1 kb away from each other, indicating tandem duplications. Within the Arabidopsis clade, PAE wheat genes were duplicated once on chromosomes 5A, 5B, and 5D, although they were between 20 and 127 kb apart (Table 2; Supplemental File S1).

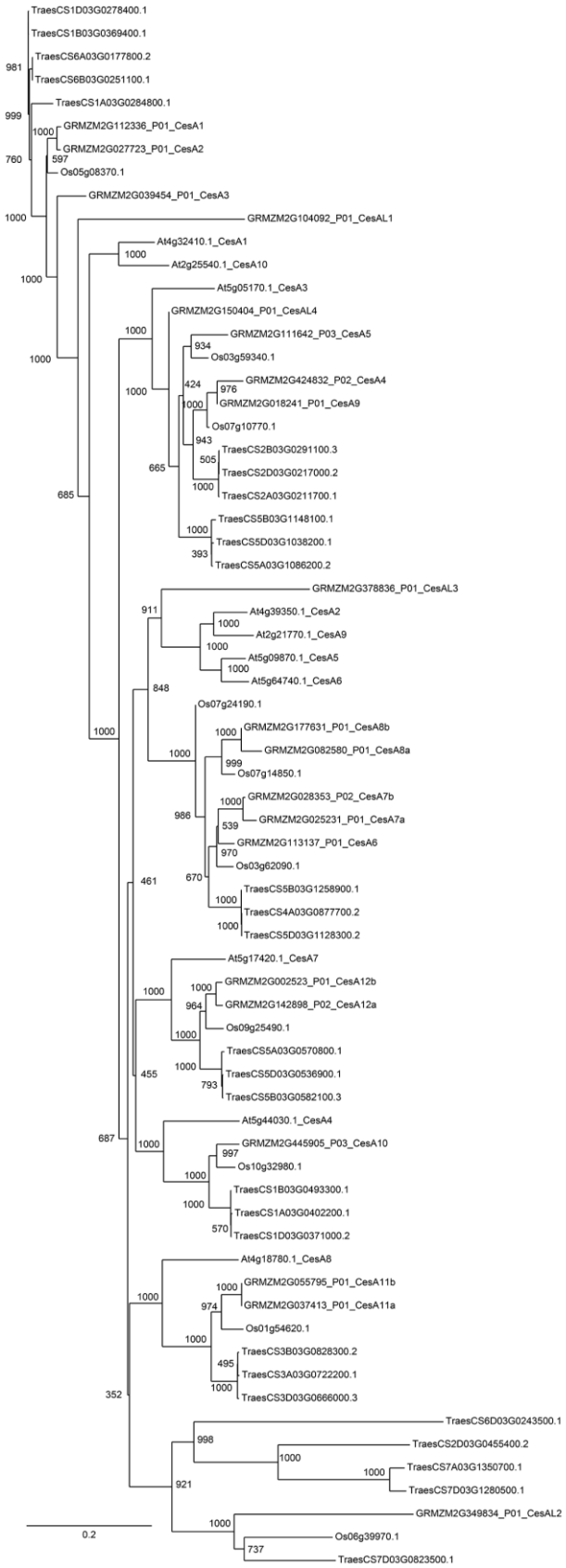

Figure 2. The CesA gene family. CesAs in wheat are as expected, while maize has a duplication of the entire gene family, twenty rather than ten genes. The scale is the number of substitutions per site. The numbers are branching based on 1000 bootstraps.

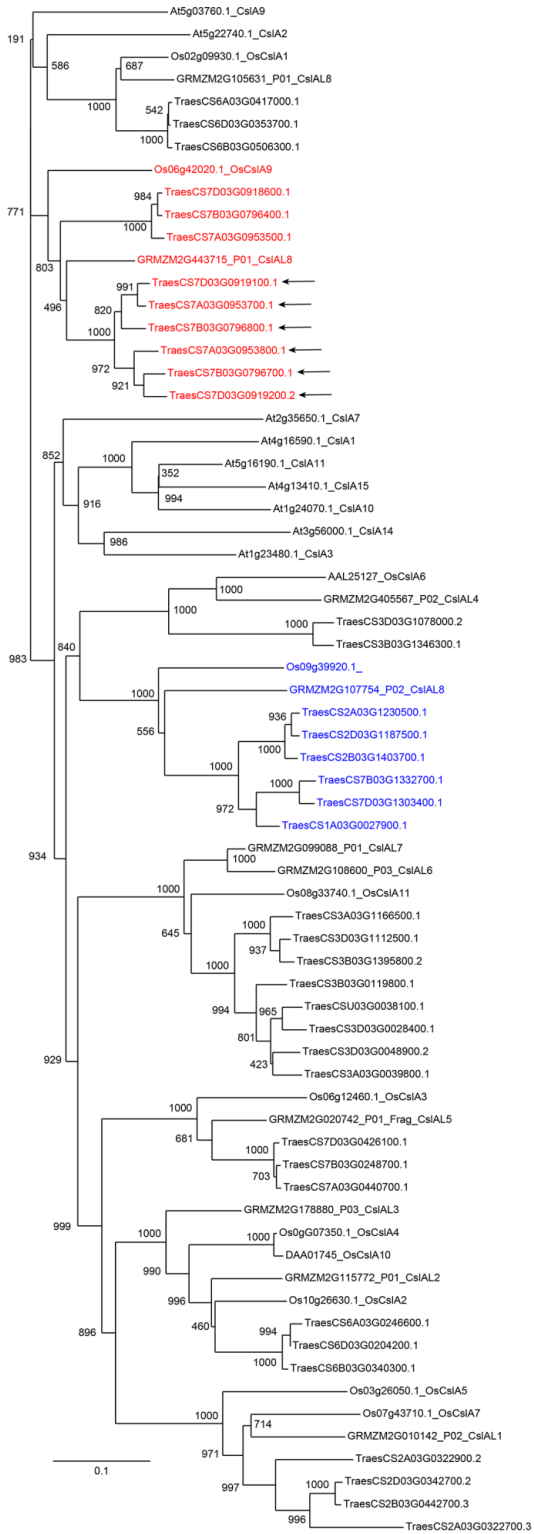

Figure 3. The CslA gene family. Wheat CslA genes are more numerous than expected. Nine wheat genes are associated with one maize and rice gene (red). Six wheat genes from multiple chromosomes are closely related to one maize and rice gene (blue). Arrows indicate six genes that appear to be tandem duplications along the wheat chromosomes. The scale is the number of substitutions per site. The numbers are branching based on 1000 bootstraps.

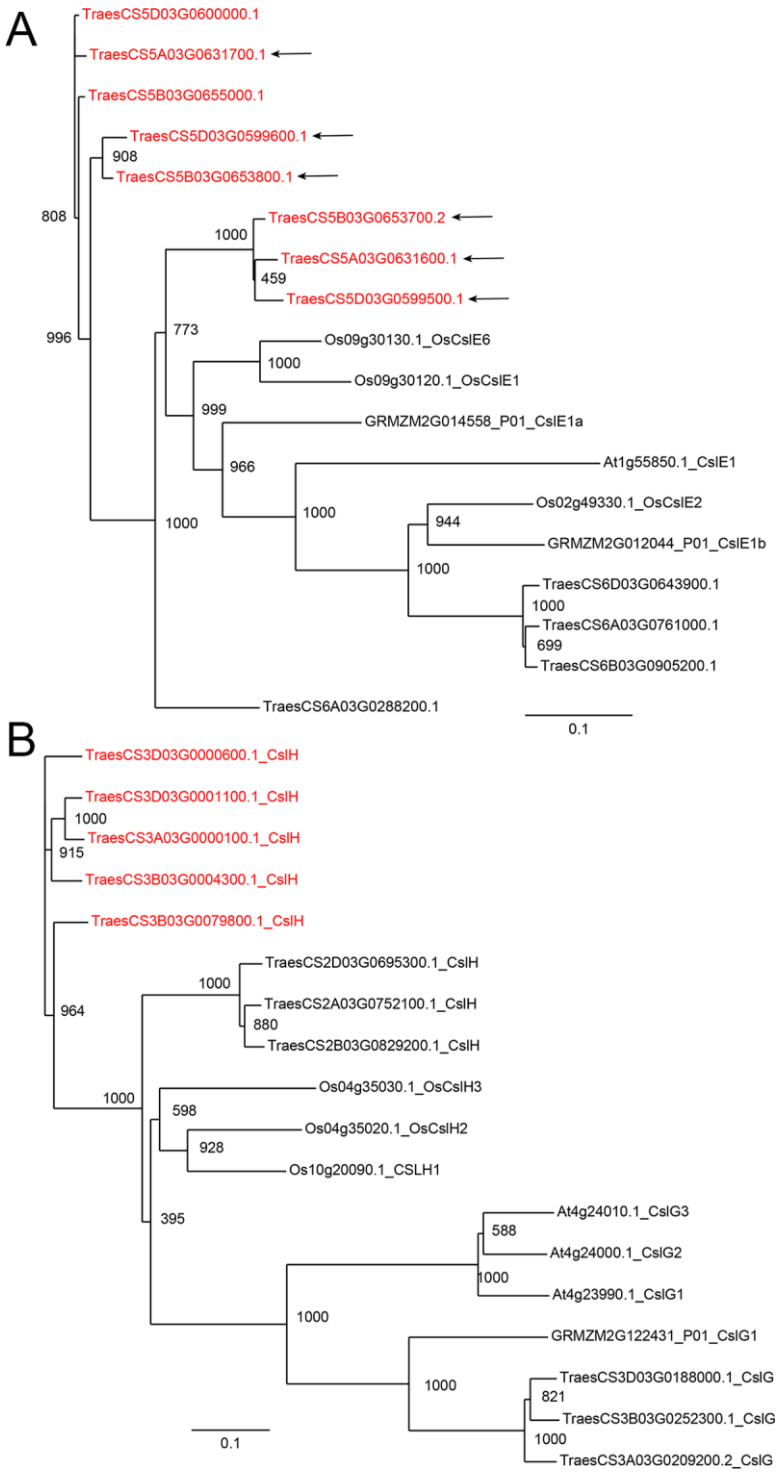

Figure 4. Wheat Csl E, G, and H genes are more numerous than expected. (**A**) CslE genes. Eight wheat genes on chromosomes 5A, 5B, and 5D are associated with two rice and one maize gene (red). Arrows indicate six genes that appear to be tandem duplications along the wheat chromosomes. (**B**) CslG and H genes. Wheat genes are more numerous than expected. Five wheat genes on chromosomes 3A, 3B, and 3D are shown in red. The scale is the number of substitutions per site. The numbers are branching based on 1000 bootstraps.

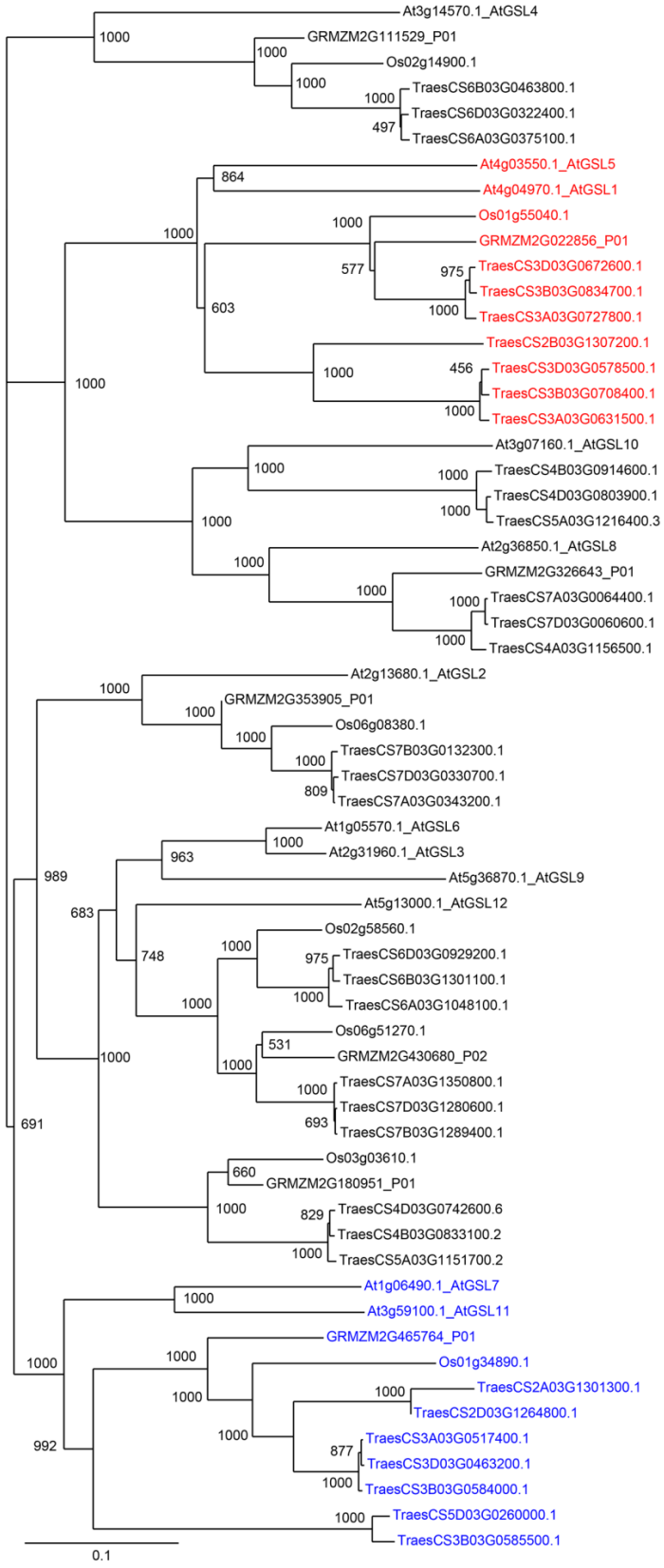

Figure 5. Family GT48. Wheat genes are more numerous than expected for the callose synthase gene family. Seven wheat genes are associated with two Arabidopsis, one rice, and one maize gene (red). Seven wheat genes are associated with two Arabidopsis and one maize and rice gene (blue). The scale is the number of substitutions per site. The numbers are branching based on 1000 bootstraps.

The GT61 and GT77 families also had a large expansion in wheat genes compared to maize, rice, and Arabidopsis, 191 in total. The rice gene, Os02g22380, a GT61 family member, was required for a β-(1,2) xylosyl substitution of α-(1,3) arabinosyls along the xylan chain, unique to grasses [31]. There were two sets of wheat genes closely associated with this rice gene on chromosomes 2 and 3 (Supplemental File S1). The wheat gene TaXAT1 (TraesCS6A03G0809100.1) of the GT61 family carried out most of the α-(1,3) arabinofuranose substitution on starchy endosperm arabinoxylans. The wheat gene TaXAT2 (TraesCS1B03G1010700.2) of the GT61 family, when transformed into Arabidopsis, converted some of the dicot glucuronoxylans to arabinoxylans. The function of GT61 was hypothesized to be the synthesis of grass-specific GAXs. In addition to adding arabinose to arabinoxylans, a rice mutant XAX1, and GT61 family member was observed to have reduced hydroxycinnamic acid-modified arabinosyl-substituted xylose [32]. With GT61 family members TaXAT1 and TaXAT2 having abrabinofuranosyltransferase activity in wheat, it is possible that the role of arabinoxylan feruloylation may partly explain the expansion of the GT61 family [30,32]. Wheat grain endosperm that is milled to flour contains arabinoxylan, a source of fiber. The alteration of TaXAT1 and XAX1 changed the amount of arabinoxylan feruloylation, which changed the fraction of soluble to insoluble fiber in wheat flour. The feruloylation of arabinoxylan increases the cross-linking of molecules, leading to more insoluble fiber. However, wheat grain is an important source of soluble fiber for human health [30,32,87]. One Arabidopsis mutant, At2g35610_XEG113, classified as a GT77 family member, had elongated hypocotyls in liquid culture with the presence of xyloglucanase. It also exhibited reduced arabinose, increased mannose, and increased glucose content in the extensin-rich fraction of the cell wall. Some GT77 family members add arabinose to extensins in the plant cell wall [88]. The GT14 family was slightly expanded to 44 members in wheat. The Arabidopsis gene At5g39990.1_AtGlcAT14A was characterized as a β-glucuronosyltransferase involved in Type II arabinogalactan biosynthesis and cell elongation [89].

The GH18 family, yieldins, had a large expansion of wheat genes over maize, rice, and Arabidopsis (Figure 6). The wheat genome showed signs of several tandem duplications. Wheat chromosomes 3A, 3B, and 3D had six to eight genes, each located between 2.5 and 457 kb from each other (Table 2). A smaller set of duplications in a second clade was also present in 2A, 2B, 2D, 7A, 7B, and 7D but included three maize genes (Figure 6). The GH18 family members were originally named chitinases for their ability to hydrolyze fungal cell walls composed of chitin and β-1,3, -1,6-glucan and provide resistance to fungal diseases [52,54]. More recent studies have suggested that the GH18 family operated in cell wall loosening, but their function(s) in plants are still not well understood [53]. The GH16 (XTH) and expansin families were greatly increased in wheat, accounting for 390 genes (Table 1). These families are also involved in cell wall loosening [53]. Expansins are related to the GH45 family but do not have glycolytic activity and may perform cell wall loosening through hydrogen bond disruption at the cell wall and aid glycoside hydrolase activity on cellulose [48,53,90]. Members of the XTH family have been hypothesized to tether or re-tether xyloglucans to cellulose during growth [49]. However, XTHs have been found to be highly upregulated and involved in remodeling the cell wall during biotic and abiotic stresses, indicating other roles in grasses [91]. The GH17 family encoding β-(1,3) glucanases showed an increase in wheat with 209 wheat genes across four groups of clades (Supplemental File S1). A rice gene, OsGLN2 (Os01g71670.1), was expressed in flowers and germinating seeds, strongly induced in the presence of gibberellin A3, and greatly decreased in the presence of abscisic acid [50]. This gene and wheat homologs may play a role in pre-harvest sprouting, as it has been suggested that many genes under the control of abscisic acid in germinating seeds are involved in pre-harvest sprouting [92]. A wheat gene in the GH17 family, TaGluD (TraesCS3A03G1134100.1), was characterized as a β-1,3-glucanase involved in fungal defense [51]. The expansion of genes in the family could be due to an increased role in pathogen and abiotic stress defense.

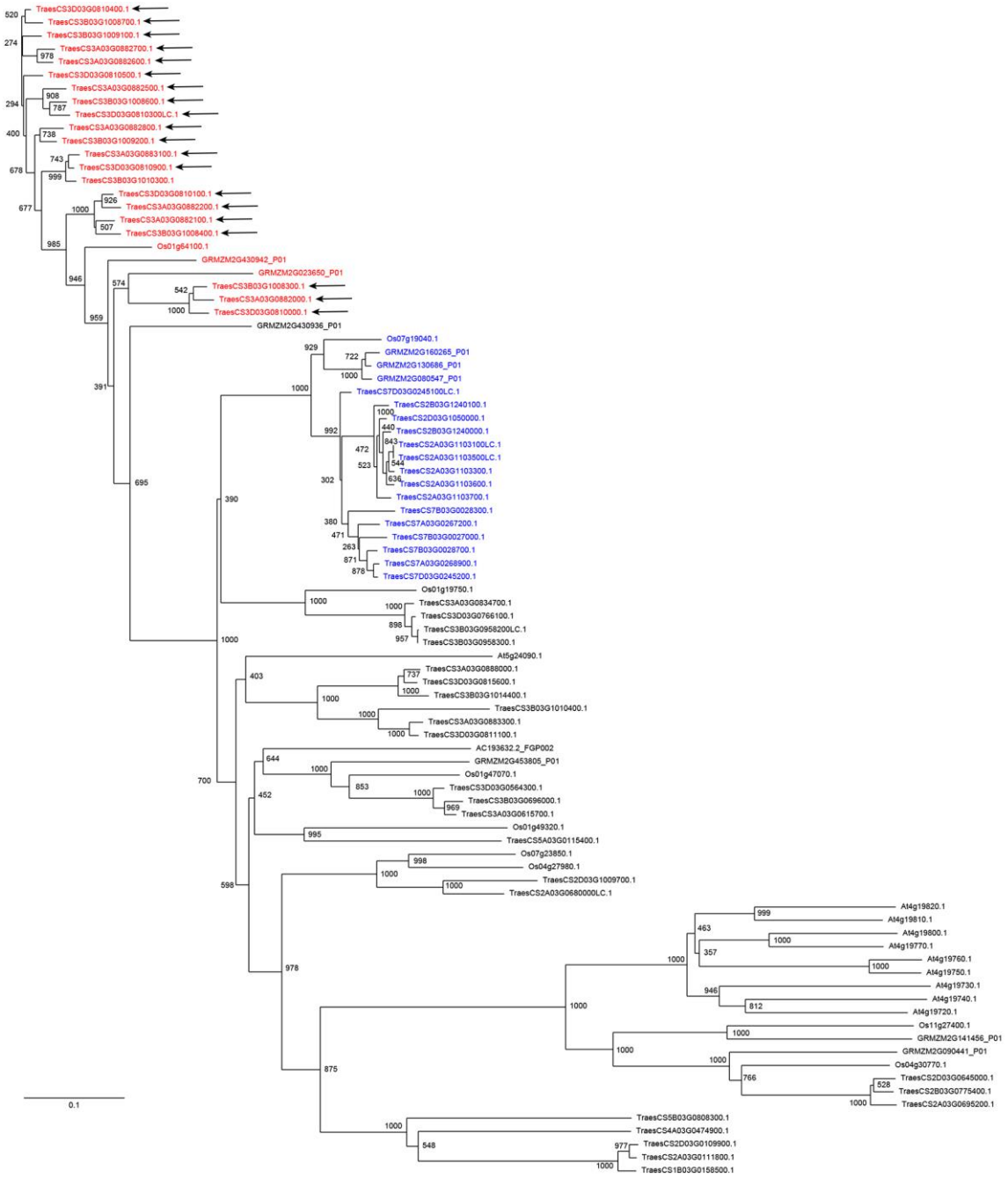

Figure 6. The GH18 gene family. Wheat yieldins are more numerous than expected. Twenty-one wheat genes on chromosomes 3A, 3B, and 3D are associated with two maize and one rice gene (red). Fifteen wheat genes on multiple chromosomes are associated with three maize and one rice gene (blue). Arrows indicate twenty genes that appear to be tandem duplications along the wheat chromosomes. The scale is the number of substitutions per site. The numbers are branching based on 1000 bootstraps.

Other cell wall-modifying enzymes with more numerous genes were present in wheat. Laccases had four to six times the number of wheat genes as maize, rice, and Arabidopsis, while peroxidases had five to eight times as many (Table 1). Laccases were observed to polymerize monolignols into larger lignin subunits during the early stages of lignification, while peroxidases were observed to perform lignification in the presence of hydrogen peroxide later in the development of dark-grown cell-suspension cultures of sycamore maple xylem cells [64]. A cotton laccase overexpressed in poplar showed dramatically

increased laccase activity and up to 19.6% increased lignin [93]. Peroxidases are a large gene family in plants that function in cell wall lignification and elongation, pathogen and abiotic stress defense, and seed germination. Peroxidases were proposed to function in cell wall stiffening through the suberization or cross-linking of extensins and degradation of anthocyanins [94,95]. The protease family of nearly 300 genes was greatly expanded in wheat, about 6 times those in maize and rice, and 16 times in Arabidopsis (Table 1). Proteases degrade other proteins [63]. Proteases have many roles in the plant but were very prevalent in grass seeds such as wheat, barley, rye, and oat, where they formed a complex mixture of grain storage proteins with the function of hydrolysis of other seed proteins during germination [96]. The cysteine (CYS) proteases were the most enlarged protease group (Supplemental File S1). This could be expected as wheat seeds have many proteins, including gluten proteins associated with celiac disorders that CYS proteases act upon, but are not as prevalent in maize and rice [96].

The results of this research are trees for all 81 cell wall gene families of wheat, maize, rice, and Arabidopsis, available as Supplemental File S1. An Excel sheet lists the wheat gene family classification resulting from this research, the gene name, and the species for each gene in the study in the first through third columns, respectively, and is available as Supplemental File S2. Gene families not described were small or did not show drastically different gene numbers of wheat compared to maize or rice and have been previously described in many publications.

It is beyond the scope of this study to determine whether subfunctionalization or neofunctionalization has occurred in each gene family. However, other literature has demonstrated specific instances where neo/subfunctionalization has occurred in cell wall gene families. When duplicated genes diverge to give one gene a new function (neofunctionalization) or the two genes each perform part of the original gene function (subfunctionalization), it is through protein-coding differences [71]. This would likely manifest in protein sequences diverging and could cause them to align to the periphery of the gene family or form a new clade. All included wheat genes that were at the tree periphery or appeared unrelated to the family were checked using Clustal Omega [86] for at least 60% protein identity and for the correct gene family-related protein motifs using the EMBL-EBI protein motif viewer or prosite [14,15]. An outlying clade in Arabidopsis PAE genes was responsive to water and salt stress, representing novel functions for the gene family through neofunctionalization [57]. Nineteen wheat genes also form a grass-only clade in the PAE family, which may represent further changes in function [71]. Similarly, promoter sequences rather than protein sequences may diverge. As the protein sequences would remain nearly the same, genes would likely group together regardless of species in tree clades, as occurred with the laccases. Poplar laccases showed different tissue expression patterns compared to Arabidopsis, suggesting the expression subfunctionalization of the gene family as poplar and Arabidopsis genes were together in clades of a tree rather than separate clades [72,97]. The laccase gene family between wheat, maize, rice, and Arabidopsis showed consistent clades with members in each clade (Supplemental File S1), just like poplar and Arabidopsis [97]. The PP genes are known to be under finer control in multiple tissues or at different times. Dicot PAL genes use only phenylalanine as a substrate, but some grass PAL genes can use phenylalanine or tyrosine as a substrate. This may have led to multiple function and expression changes over time, explaining the large PP gene expansion, potentially expressing subfunctionalization and neofunctionalization [72,98,99]. Peroxidases were hypothesized to have increased numbers in plants to maintain a diverse set of promoters for different tissues and timing of expression and protein distribution associated with large and diverse projected activities in the plant, which could be neofunctionalization and/or expression subfunctionalization [94].

As an example of the utility of the information presented, a comparison was performed with the bioinformatics available for wheat genes in a recent RNA sequencing expression analysis paper on pre-harvest sprouting resistance. The author found that many genes differentially expressed in the seed at 35 days after anthesis in a pre-harvest sprouting

resistant variety, Scotty, between the control plants and those under conditions conducive to pre-harvest sprouting [10]. Many cell wall-related genes were reported in the Supplemental Files but could not be classified for potential function due to vague descriptions in available bioinformatics searches. By converting the version 1 protein to version 2.1 and searching Supplemental File S2, the specific gene families were revealed (Table 3). Five genes listed only as "cellulose synthase-like protein" became three members of CslC and two from CslF. All three CslCs are related to maize CslC12a and CslC12b, while the CslFs are related to rice CslF6 and maize CSlF2 and CSlF4 (Supplemental File S1). Five other genes were variously listed as "UDP-glucose:glycoprotein glucosyltransferase", "Glycosyltransferase", and "glycosyltransferase family exostosin protein" but belong to the GT families GT24, GT61, and GT47 E, respectively. The CslC family generates the xylan backbone, while some GT47 genes are needed for xylan backbone synthesis [26,47]. The GT61 genes are involved in forming GAXs. A transgenic Arabidopsis plant expressing a wheat GT61 gene was able to produce arabinoxylans from a xyloglucan backbone [30]. Genes belonging to these families are highly expressed during seed development in spring wheat [100]. This suggests the involvement of CslC, GT47E, and GT61, which combine to produce GAXs, as having a possible role in pre-harvest sprouting. Also, CslFs produce a grass-unique mixed linkage (1,3;1,4)-β-D-glucan that may be involved [25]. Finally, several members of the PP were also differentially expressed in Scotty in control versus pre-harvest sprouting conducive conditions. Most were listed with correct gene families, but "O-methyl transferase" was not family specific. The gene was revealed as belonging to the COMT family here (Table 3). Thus, the PP pathway may also have a role in pre-harvest sprouting through the production of monolignols. Similar searches could be performed quickly and easily for any group of wheat cell wall-related genes that were unclear from a basic bioinformatic search.

Table 3. Cell wall gene names improved from conventional bioinformatics searches using Supplemental File S2.

ProteinV1	Description from Penning 2023	ProteinV2_1	Gene Family
TraesCS5A01G405700.1	Cellulose synthase-like protein	TraesCS5A03G0962500.1	GT2–CSL C
TraesCS5B01G410400.1	Cellulose synthase-like protein	TraesCS5B03G1012800.1	GT2–CSL C
TraesCS5D01G415700.1	Cellulose synthase-like protein	TraesCS5D03G0916900.2	GT2–CSL C
TraesCS7A01G298600.1	Cellulose synthase-like protein	TraesCS7A03G0733400.1	GT2–CSL F
TraesCS7B01G188400.2	Cellulose synthase-like protein	TraesCS7B03G0538200.1	GT2–CSL F
TraesCS1A01G152500.1	UDP-glucose:glycoprotein glucosyltransferase	TraesCS1A03G0409400.1	GT24
TraesCS1D01G149400.1	UDP-glucose:glycoprotein glucosyltransferase	TraesCS1D03G0377700.1	GT24
TraesCS6A01G102100.1	Glycosyltransferase	TraesCS6A03G0237900.1	GT61
TraesCS3A01G440800.1	glycosyltransferase family exostosin protein	TraesCS3A03G1025100.1	GT47 E
TraesCS3D01G433400.1	glycosyltransferase family exostosin protein	TraesCS3D03G0953500.2	GT47 E
TraesCS2A01G051600.2	O-methyltransferase	TraesCS2A03G0099800.2	1.3 COMT

4. Conclusions

The purpose of this research was to produce a curated set of wheat cell wall-related genes captured in 81 gene families and compare them to homologous genes in maize, rice, and Arabidopsis previously curated. The added information can be used to gain insight into potential wheat gene function from bioinformatic searches. An example showed how information in this research better classified the cell wall-related wheat genes from an expression analysis. Some wheat cell wall-related gene families showed a large expansion beyond what would be expected due to their hexaploid compared to the diploid species maize, rice, and Arabidopsis. Some expansions appeared to be due to tandem duplication events in wheat. This work presents a more complete picture of a large and diverse set of

gene families in wheat responsible for an important plant function, the production and maintenance of the cell wall.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/d15111135/s1, Supplemental File S1. All cell wall-related gene family trees for wheat, maize, rice, and Arabidopsis. Supplemental File S2. All gene names for wheat, rice, maize, and Arabidopsis and their cell wall-related gene families.

Funding: This research was supported by the U.S. Department of Agriculture, Agricultural Research Service project # 5082-43440-001-00D.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Full gene names have been given, and the respective protein gene sequences can be downloaded from the websites for the versions of the wheat, maize, rice, and Arabidopsis as described in Section 2. The websites are long-term repositories for protein sequences.

Acknowledgments: The 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 U.S. Department of Agriculture. USDA is an equal opportunity provider and employer. The findings and conclusions in this report are those of the author and should not be construed to represent any official USDA or U.S. Government determination or policy.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

4-coumarate CoA ligase (4CL), arabinogalactan protein (AGP), basepairs (bp), cinnamate 4-hydroxylase (C4H), (hydroxy)cinnamyl alcohol dehydrogenase (CAD), caffeoyl CoA O-methyltransferase (CCoAOMT), caffeic acid (5-hydroxy-coniferaldehyde) O-methyltransferase (COMT), carbohydrate esterase (CE) cellulose synthase (CesA), (hydroxy)cinnamoyl CoA reductase (CCR), cellulose synthase-like (Csl), cysteine (CYS), ferulate (coniferaldehyde) 5-hydroxylase (F5H), glucoarabinoxylan (GAX), GDP-4-keto-6-deoxy-D-mannose 3,5-epimerase-4-reductase (GER), GDP-D-mannose-4,6-dehydratase (GMD), GDP-D-mannose 3,5-epimerase (GME), glycosyl hydrolases (GH), glycosyl transferases (GT), homogalaturonan (HG), hydroxycinnamoyl CoA:shikimate hydroxycinnamoyl transferase (HCT), kilobases (kb), low confidence wheat gene (LC), membrane-anchored UDP-D-glucuronate decarboxylase (AUD), *p*-coumaroyl shikimate 3′-hydroxylase (C3′H), pectin acetylesterase (PAE), pectin methylesterases (PME), phenylalanine ammonia lyase (PAL), Phenylpropanoid pathway (PP), polygalacturonase (PGase), polysaccharide lyases (PL), prolyl-4-hydroxylase (P4H), rhamnogalacturonan (RG), skewed growth proteins (SKU), soluble UDP-D-glucuronate decarboxylase (SUD) genes, UDP-D-glucose 4-epimerase (AXS), UDP-D-glucuronate 4-epimerase (GAE), UDP-L-rhamnose synthase (RHM), UDP-4-keto-6-deoxy-D-glucose 3,5-epimerase-4-reductase (UER), UDP-D-glucose dehydrogenase (UGD), UDP-D-glucose 4-epimerase (UGE), and UDP-D-xylose 4-epimerase (UXE), xyloglucan endotransglucosylase/hydrolase (XTH), xyloglucan (XyG).

References

1. Asseng, S.; Guarin, J.R.; Raman, M.; Monje, O.; Kiss, G.; Despommier, D.D.; Meggars, F.M.; Gauthier, P.P.G. Wheat yield potential in controlled-environment vertical farms. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 19131–19135. [CrossRef]
2. Wheat Explorer. Available online: https://ipad.fas.usda.gov/cropexplorer/cropview/commodityView.aspx?cropid=0410000 (accessed on 21 August 2023).
3. Tripathi, M.K.; Karim, S.A.; Chaturvedi, O.H.; Verma, D.L. Nutritional value of animal feed grade wheat as a replacement for maize in lamb feeding for mutton production. *J. Sci. Food Agric.* **2007**, *87*, 2447–2455. [CrossRef]
4. Patwa, N.; Penning, B.W. Environmental impact on cereal crop grain damage from pre-harvest sprouting and late maturity alpha-amylase. In *Sustainable Agriculture in the Era of Climate Change*; Roychowdhury, R., Choudhury, S., Hasanuzzaman, M., Srivastava, S., Eds.; Springer Nature: Cham, Switzerland, 2020; pp. 23–41. [CrossRef]
5. Farrokhi, N.; Burton, R.A.; Brownfield, L.; Hrmova, M.; Wilson, S.M.; Bacic, A.; Fincher, G.B. Plant Cell wall biosynthesis: Genetic, biochemical and functional genomics approaches to the identification of key genes. *Plant Biotechnol. J.* **2006**, *4*, 145–167. [CrossRef]
6. Shrivastava, B.; Jain, K.K.; Kalra, A.; Kuhad, R.C. Bioprocessing of wheat straw into nutritionally rich and digested cattle feed. *Sci. Rep.* **2014**, *4*, 6360. [CrossRef] [PubMed]

7. Tsegaye, B.; Balomajumder, C.; Roy, P. Optimization of microwave and NaOH pretreatments of wheat straw for enhancing biofuel yield. *Energy Convers. Manag.* **2019**, *186*, 82–92. [CrossRef]
8. Yong, W.; Link, B.; O'Malley, R.; Tewari, J.; Hunter, C.T.; Lu, C.A.; Li, X.; Bleecker, A.B.; Koch, K.E.; McCann, M.C.; et al. Genomics of plant cell wall biogenesis. *Planta* **2005**, *221*, 747–751. [CrossRef]
9. Nirmal, R.C.; Furtado, A.; Rangan, P.; Henry, R.J. Fasciclin-like arabinogalactan protein gene expression is associated with yield of flour in the milling of wheat. *Sci. Rep.* **2017**, *7*, 12539. [CrossRef]
10. Penning, B.W. Gene expression differences related to pre-harvest sprouting uncovered in related wheat varieties by RNAseq analysis. *Plant Gene* **2023**, *33*, 100404. [CrossRef]
11. Drula, E.; Garron, M.L.; Dogan, S.; Lombard, V.; Henrissat, B.; Terrapon, N. The carbohydrate-active enzyme database: Functions and literature. *Nucl. Acids Res.* **2022**, *50*, D571–D577. [CrossRef] [PubMed]
12. Penning, B.W.; McCann, M.C.; Carpita, N.C. Evolution of the cell wall gene families of grasses. *Front. Plant Sci.* **2019**, *10*, 1205. [CrossRef]
13. Penning, B.W.; Hunter, C.T.; Tayengwa, R.; Eveland, A.L.; Dugard, C.K.; Olek, A.T.; Vermerris, W.; Koch, K.E.; McCarty, D.R.; Davis, M.F.; et al. Genetic Resources for maize cell wall biology. *Plant Physiol.* **2009**, *151*, 1703–1728. [CrossRef]
14. Potter, S.C.; Luciani, A.; Eddy, S.R.; Park, Y.; Lopez, R.; Finn, R.D. HMMER webserver: 2018 update. *Nucl. Acids Res.* **2018**, *46*, W200–W204. [CrossRef] [PubMed]
15. Sigrist, C.J.A.; de Castro, E.; Cerutti, L.; Cuche, B.A.; Hulo, N.; Bridge, A.; Bougueleret, L.; Xenarios, I. New and continuing developments at PROSITE. *Nucl. Acids Res.* **2013**, *41*, D344–D347. [CrossRef]
16. Carpita, N.C. Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1996**, *47*, 445–476. [CrossRef]
17. Penning, B.W.; Shiga, T.M.; Klimek, J.F.; SanMiguel, P.J.; Shreve, J.; Thimmapuram, J.; Sykes, R.W.; Davis, M.F.; McCann, M.C.; Carpita, N.C. Expression profiles of cell-wall related genes vary broadly between two common maize inbreds during stem development. *BMC Genom.* **2019**, *20*, 785. [CrossRef] [PubMed]
18. CAZY. Available online: http://www.cazy.org/ (accessed on 10 August 2023).
19. Orellana, A.; Moraga, C.; Araya, M.; Moreno, A. Overview of nucleotide sugar transporter gene family functions across multiple species. *J. Mol. Biol.* **2016**, *428*, 3150–3165. [CrossRef] [PubMed]
20. Bonawitz, N.D.; Chapple, C. The genetics of lignin biosynthesis: Connecting genotype to phenotype. *Annu. Rev. Genet.* **2010**, *44*, 337–363. [CrossRef]
21. Reiter, W.D.; Vanzin, G.F. Molecular genetics of nucleotide sugar interconversion pathways in plants. *Plant Mol. Biol.* **2001**, *47*, 95–113. [CrossRef]
22. Yin, Y.; Huang, J.; Gu, X.; Bar-Peled, M.; Xu, Y. Evolution of plant nucleotide-sugar interconversion enzymes. *PLoS ONE* **2011**, *6*, e27995. [CrossRef]
23. Favery, B.; Ryan, E.; Foreman, J.; Linstead, P.; Boudonck, K.; Steer, M.; Shaw, P.; Dolan, L. KOJAK encodes a cellulose synthase-like protein required for root hair cell morphogenesis in Arabidopsis. *Genes Dev.* **2001**, *15*, 79–89. [CrossRef]
24. Holland, N.; Holland, D.; Helentjaris, T.; Dhugga, K.; Xoconostle-Cazares, B.; Delmer, D.P. A comparative analysis of the plant cellulose synthase (CesA) gene family. *Plant Physiol.* **2000**, *123*, 1313–1323. [CrossRef]
25. Burton, R.A.; Wilson, S.M.; Hrmova, M.; Harvey, A.J.; Shirley, N.J.; Medhurst, A.; Stone, B.A.; Newbigin, E.J.; Bacic, A.; Fincher, G.B. Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)-b-D-glucans. *Science* **2006**, *311*, 1940–1942. [CrossRef]
26. Cocuron, J.C.; Lerouxel, O.; Drakakaki, G.; Alonso, A.P.; Liepman, A.H.; Keegstra, K.; Raikhel, N.; Wilkerson, C.G. A gene from the cellulose synthase-like C family encodes a β-1,4 glucan synthase. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8550–8555. [CrossRef]
27. Dhugga, K.S.; Barreiro, R.; Whitten, B.; Stecca, K.; Hazebroek, J.; Randhawa, G.S.; Dolan, M.; Kinney, A.J.; Tomes, D.; Nichols, S.; et al. Guar seed β-mannan synthase is a member of the cellulose synthase super gene family. *Science* **2004**, *303*, 363–366. [CrossRef]
28. Liu, X.; Zhang, H.; Zhang, W.; Xu, W.; Li, S.; Chen, X.; Chen, H. Genome-wide bioinformatics analysis of cellulose synthase gene family in common bean (*Phaseolus vulgaris* L.) and the expression in the pod development. *BMC Genom. Data* **2022**, *23*, 9. [CrossRef] [PubMed]
29. Wang, B.; Andargie, M.; Fang, R. The function and biosynthesis of callose in high plants. *Heliyon* **2022**, *8*, e09248. [CrossRef]
30. Anders, N.; Wilkinson, M.D.; Lovegrove, A.; Freeman, J.; Tryfona, T.; Pellny, T.K.; Weimar, T.; Mortimer, J.C.; Stott, K.; Baker, J.M.; et al. Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 989–993. [CrossRef] [PubMed]
31. Chiniquy, D.; Sharma, V.; Schultink, A.; Baidoo, E.E.; Rautengarten, C.; Cheng, K.; Carroll, A.; Ulvskov, P.; Harholt, J.; Keasling, J.D.; et al. XAX1 from glycosyltransferase family 61 mediates xylosyl transfer to rice xylan. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17117–17122. [CrossRef] [PubMed]
32. Feijao, C.; Morreel, K.; Anders, A.; Tryfona, T.; Busse-Wicher, M.; Kotake, T.; Boerjan, W.; Dupree, P. Hydroxycinnamic acid-modified xylan side chains and their cross-linking products in rice cell walls are reduced in the *Xylosyl arabinosyl substitution of xylan 1* mutant. *Plant J.* **2022**, *109*, 1152–1167. [CrossRef]

33. Egelund, J.; Obel, N.; Ulvskov, P.; Geshi, N.; Pauly, M.; Bacic, A.; Petersen, B.L. 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.* **2007**, *64*, 439–451. [CrossRef] [PubMed]
34. Scheller, H.V.; Jensen, J.K.; Sørensen, S.O.; Harholt, J.; Geshi, N. Biosynthesis of pectin. *Physiol. Plant.* **2007**, *129*, 283–295. [CrossRef]
35. Sterling, J.D.; Atmodjo, M.A.; Inwood, S.E.; Kolli, V.S.K.; Quigley, H.F.; Hahn, M.G.; Mohnen, D. Functional identification of an Arabidopsis pectin biosynthesic homogalacturonan galacturonosyltransferase. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5236–5241. [CrossRef]
36. Lee, C.; Zhong, R.; Richardson, E.A.; Himmelsbach, D.S.; McPhail, B.T.; Ye, Z.H. The PARVUS gene is expressed in cells undergoing secondary wall thickening and is essential for glucuronoxylan biosynthesis. *Plant Cell Physiol.* **2007**, *48*, 1659–1672. [CrossRef] [PubMed]
37. Qu, Y.; Egelund, J.; Gilson, P.R.; Houghton, F.; Gleeson, P.A.; Schultz, C.J.; Bacic, A. Identification of a novel group of putative *Arabidopsis thaliana* β-(1,3)-galactosyltransferases. *Plant Mol. Biol.* **2008**, *68*, 43–59. [CrossRef]
38. Strasser, R.; Bondili, J.S.; Vavra, U.; Schoberer, J.; Svoboda, B.; Glössl, J.; Léonard, R.; Stadlmann, J.; Altmann, F.; Steinkellner, H.; et al. A unique β1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*. *Plant Cell* **2007**, *19*, 2278–2292. [CrossRef]
39. Cavalier, D.M.; Lerouxel, O.; Neumetzler, L.; Yamauchi, K.; Reinecke, A.; Freshour, G.; Zabotina, O.A.; Hahn, M.G.; Burgert, I.; Pauly, M.; et al. Disrupting two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell* **2008**, *20*, 1519–1537. [CrossRef] [PubMed]
40. Sarria, R.; Wagner, T.A.; O'Neill, M.A.; Faik, A.; Wilkerson, C.G.; Keegstra, K.; Raikhel, N.V. Characterization of a family of Arabidopsis genes related to xyloglucan fucosyltransferase1. *Plant Physiol.* **2001**, *127*, 1595–1606. [CrossRef] [PubMed]
41. Vanzin, G.F.; Madson, M.; Carpita, N.C.; Raikhel, N.V.; Keegstra, K.; Reiter, W.D. The *mur2* mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3340–3345. [CrossRef] [PubMed]
42. Harholt, J.; Jensen, J.K.; Sørensen, S.O.; Orfila, C.; Pauly, M.; Scheller, H.V. *ARABINAN DEFICIENT1* is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in Arabidopsis. *Plant Physiol.* **2006**, *140*, 49–58. [CrossRef]
43. Jensen, J.K.; Sørensen, S.O.; Harholt, J.; Geshi, N.; Sakuragi, Y.; Møller, I.; Zandleven, J.; Bernal, A.J.; Jensen, N.B.; Sørensen, C.; et al. Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in Arabidopsis. *Plant Cell* **2008**, *20*, 1289–1302. [CrossRef]
44. Madson, M.; Dunand, C.; Li, X.; Verma, R.; Vanzin, G.F.; Caplan, J.; Shoue, D.A.; Carpita, N.C.; Reiter, W.-D. The *MUR3* gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. *Plant Cell* **2003**, *15*, 1662–1670. [CrossRef]
45. Peña, M.J.; Zhong, R.; Zhou, G.K.; Richardson, E.A.; O'Neill, M.A.; Darvill, A.G.; York, W.S.; Ye, Z.H. Arabidopsis *irregular xylem8* and *irregular xylem9*: Implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* **2007**, *19*, 549–563. [CrossRef] [PubMed]
46. Zhong, R.; Peña, M.J.; Zhou, G.K.; Nairn, C.J.; Wood-Jones, A.; Richardson, E.A.; Morrison, W.H.; Darvill, A.G.; York, W.S.; Ye, Z.-H. Arabidopsis *fragile fiber8*, which encodes a putative glucuronosyltransferase, is essential for normal secondary wall synthesis. *Plant Cell* **2005**, *17*, 3390–3408. [CrossRef] [PubMed]
47. Brown, D.M.; Zhang, Z.; Stephens, E.; Dupree, P.; Turner, S.R. Characterization of *IRX10* and *IRX10*-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *Plant J.* **2009**, *57*, 732–746. [CrossRef]
48. Cosgrove, D.J. Loosening of plant cell walls by expansins. *Nature* **2000**, *407*, 321–326. [CrossRef] [PubMed]
49. Rose, J.K.C.; Braam, J.; Fry, S.C.; Nishitani, K. The *XTH* family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: Current perspectives and a new unifying nomenclature. *Plant Cell Physiol.* **2002**, *43*, 1421–1435. [CrossRef]
50. Akiyama, T.; Pillai, M.A.; Sentoku, N. Cloning, characterization and expression of *OsGLN2*, a rice endo-1,3-beta-glucanase gene regulated developmentally in flowers and hormonally in germinating seeds. *Planta* **2004**, *220*, 129–139. [CrossRef]
51. Liu, B.; Lu, Y.; Xin, Z.; Zhang, Z. Identification and antifungal assay of a wheat B-1,3-glucanase. *Biotechnol. Lett.* **2009**, *31*, 1005–1010. [CrossRef]
52. Minic, Z. Physiological roles of plant glycoside hydrolases. *Planta* **2008**, *227*, 723–740. [CrossRef]
53. Nawaz, M.A.; Rehman, H.M.; Imtiaz, M.; Baloch, F.S.; Lee, J.D.; Yang, S.H.; Lee, S.I.; Chung, G. Systems identification and characterization of cell wall reassembly and degradation related genes in *Glycine max* (L.) Merill, a bioenergy legume. *Sci. Rep.* **2017**, *7*, 10862. [CrossRef]
54. Schlumbaum, A.; Mauch, F.; Vögeli, U.; Boller, T. Plant chitinases are potent inhibitors of fungal growth. *Nature* **1986**, *324*, 365–367. [CrossRef]
55. Jindou, S.; Xu, Q.; Kenig, R.; Shulman, M.; Shoham, Y.; Bayer, E.A.; Lamed, R. Novel architecture of family-9 glycoside hydrolases identified in cellulosal enzymes of Acetivibrio cellulyticus and Clostridium thermocellum. *FEMS Microbiol. Lett.* **2006**, *254*, 308–316. [CrossRef] [PubMed]
56. Jenkins, E.S.; Paul, W.; Craze, M.; Whitelaw, A.; Weigand, A.; Roberts, J.A. Dehiscence-related expression of an *Arabidopsis thaliana* gene encoding a polygalacturonase in transgenic plants of *Brassica napus*. *Plant Cell Environ.* **1999**, *22*, 159–167. [CrossRef]

57. Sampedro, J.; Gianzo, C.; Iglesias, N.; Guitián, E.; Revilla, G.; Zarra, I. AtBGAL10 is the main xyloglucan β-galactosidase in Arabidopsis, and its absence results in unusual xyloglucan subunits and growth defects. *Plant Physiol.* **2012**, *158*, 1146–1157. [CrossRef] [PubMed]
58. Gou, Y.Y.; Miller, L.M.; Hou, G.; Yu, X.H.; Chen, X.Y.; Liu, C.J. Acetylesterase-mediated deacetylation of pectin impairs cell elongation, pollen germination, and plant reproduction. *Plant Cell* **2012**, *24*, 50–65. [CrossRef]
59. Philippe, F.; Pelloux, J.; Rayon, C. Plant pectin acetylesterase structure and function: New insights from bioinformatic analysis. *BMG Genom.* **2017**, *18*, 456. [CrossRef]
60. Louvet, R.; Cavel, E.; Gutierrez, L.; Guénin, S.; Roger, D.; Gillet, F.; Guerineau, F.; Pelloux, J. Comprehensive expression profiling of the pectin methylesterase gene family during silique development in *Arabidopsis thaliana*. *Planta* **2006**, *224*, 782–791. [CrossRef]
61. Domingo, C.; Roberts, K.; Stacey, N.J.; Connerton, I.; Ruíz-Teran, F.; McCann, M.C. A pectate lyase from *Zinnia elegans* is auxin inducible. *Plant J.* **1998**, *13*, 17–28. [CrossRef]
62. Ochoa-Jiménez, V.A.; Berumen-Varela, G.; Burgara-Estrella, A.; Orozco-Avitia, J.A.; Ojeda-Contreras, A.J.; Trillo-Hernández, E.A.; Rivera-Domínguez, M.; Troncoso-Rojas, R.; Báez-Sañudo, R.; Datsenka, T.; et al. Functional analysis of tomato rhamnogalacturonan lyase gene Solyc11g011300 during fruit development and ripening. *J. Plant Physiol.* **2018**, *231*, 31–40. [CrossRef]
63. Canut, H.; Albenne, C.; Jamet, E. Post-translational modifications of plant cell wall proteins and peptides: A survey from a proteomics point of view. *Biochim. Biophys. Acta–Proteins Proteom.* **2016**, *1864*, 983–990. [CrossRef]
64. Sterjiades, R.; Dean, J.F.D.; Gamble, G.; Himmelsbach, D.S.; Eriksson, K.L. Extracellular laccases and peroxidases from sycamore maple (*Acer pseudoplatanus*) cell-suspension cultures. *Planta* **1993**, *190*, 75–83. [CrossRef]
65. Brown, D.M.; Zeef, L.A.H.; Ellis, J.; Goodacre, R.; Turner, S.R. Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* **2005**, *17*, 2281–2295. [CrossRef] [PubMed]
66. Li, Y.; Qian, Q.; Zhou, Y.; Yan, M.; Sun, L.; Zhang, M.; Fu, Z.; Wang, Y.; Han, B.; Pang, X.; et al. *BRITTLE CULM1*, which encodes a COBRA-like protein, affects the mechanical properties of rice plants. *Plant Cell* **2003**, *15*, 2020–2031. [CrossRef] [PubMed]
67. Velasquez, S.M.; Ricardi, M.M.; Poulsen, C.P.; Oikawa, A.; Dilokpimol, A.; Halim, A.; Mangano, S.; Juarez, S.P.D.; Marzol, E.; Salter, J.D.S.; et al. Complex regulation of prolyl-5-hydroxylases impacts root hair expansion. *Mol. Plant* **2015**, *8*, 734–746. [CrossRef]
68. Califar, B.; Sng, N.J.; Zupanska, A.; Paul, A.L.; Ferl, R.J. Root skewing-associated genes impact the spaceflight response of *Arabidopsis thaliana*. *Front. Plant Sci.* **2020**, *11*, 239. [CrossRef]
69. Sedbrook, J.C.; Carroll, K.L.; Hung, K.F.; Masson, P.H.; Somerville, C.R. The Arabidopsis *SKU5* gene encodes an extracellular glycosyl phosphatidylinositol-anchored glycoprotein involved in directional root growth. *Plant Cell* **2002**, *14*, 1635–1648. [CrossRef]
70. Zhou, C.; Dong, Z.; Zhang, T.; Wu, J.; Yu, S.; Zeng, Q.; Han, D.; Tong, W. Genome-scale analysis of homologous genes among subgenomes of bread wheat (*Triticum aestivum* L.). *Int. J. Mol. Sci.* **2020**, *21*, 3015. [CrossRef]
71. Johnson, D.A.; Thomas, M.A. The monosaccharide transporter gene family in Arabidopsis and rice: A history of duplications, adaptive evolution, and functional divergence. *Mol. Biol. Evol.* **2007**, *24*, 2412–2423. [CrossRef]
72. Xu, A.F.; Molinuevo, R.; Fazzari, E.; Tom, H.; Zhang, Z.; Menendez, J.; Casey, K.M. Subfunctionalized expression drives evolutionary retention of ribosomal protein paralogs *Rps27* and *Rps27l* in vertebrates. *eLife* **2023**, *12*, e78695. [CrossRef]
73. Zhu, T.; Wang, L.; Rimbert, H.; Rodriguez, J.C.; Deal, K.R.; De Oliveira, R.; Choulet, F.; Keeble-Gagnère, G.; Tibbits, J.; Rogers, J.; et al. Optical maps refine the bread wheat *Triticum aestivum* cv Chinese Spring genome assembly. *Plant J.* **2021**, *107*, 303–314. [CrossRef]
74. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef] [PubMed]
75. International Wheat Genome Sequencing Consortium. Available online: https://www.wheatgenome.org/ (accessed on 21 August 2023).
76. MaizeGDB. Available online: https://www.maizegdb.org/ (accessed on 21 August 2023).
77. Rice Genome Annotation Project. Available online: http://rice.uga.edu/ (accessed on 21 August 2023).
78. The Arabidopsis Information Reseource. Available online: https://www.arabidopsis.org/ (accessed on 21 August 2023).
79. Chenna, R.; Sugawara, H.; Koike, T.; Lopez, R.; Gibson, T.J.; Higgins, D.G.; Thompson, J.D. Multiple sequence alignment with the Clustal series of programs. *Nucl. Acids Res.* **2003**, *31*, 3497–3500. [CrossRef]
80. Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [CrossRef]
81. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **1994**, *22*, 4673–4680. [CrossRef] [PubMed]
82. Chevenet, F.; Brun, C.; Banuls, A.L.; Jacq, B.; Christen, R. TreeDyn: Towards dynamic graphics and annotations for analyses of trees. *BMC Bioinform.* **2006**, *7*, 439. [CrossRef]
83. HMMER. Available online: https://www.ebi.ac.uk/Tools/hmmer/ (accessed on 11 August 2023).
84. Prosite. Available online: https://prosite.expasy.org (accessed on 11 August 2023).
85. Clustal Omega. Available online: https://www.ebi.ac.uk/Tools/msa/clustalo/ (accessed on 11 August 2023).
86. Madeira, F.; Pearce, M.; Tivey, A.R.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucl. Acids Res.* **2022**, *50*, W276–W279. [CrossRef]

87. Pellny, T.K.; Patil, A.; Wood, A.J.; Freeman, J.; Halsey, K.; Plummer, A.; Kosik, O.; Temple, H.; Collins, J.D.; Dupree, P.; et al. Loss of TaIRX9b gene function in wheat decreases chain length and amount of arabinoxylan in grain but increases cross-linking. *Plant Biotech. J.* **2020**, *18*, 2316–2327. [CrossRef]
88. Gille, S.; Hänsel, U.; Ziemann, M.; Pauly, M. Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. *Proc. Natl. Acad. Sci. USA* **2009**, *25*, 14699–14704. [CrossRef]
89. Knoch, E.; Dilokpimol, A.; Tryfona, T.; Poulsen, C.P.; Xiong, G.; Harholt, J.; Petersen, B.L.; Ulvskov, P.; Hadi, M.Z.; Kotake, T.; et al. A β-glucuronosyltransferase from *Arabidopsis thaliana* involved in biosynthesis of type II arabinogalactan has a role in cell elongation during seedling growth. *Plant J.* **2013**, *76*, 1016–1029. [CrossRef]
90. Sampedro, J.; Cosgrove, D.J. The expansin superfamily. *Genome Biol.* **2005**, *6*, 242. [CrossRef]
91. Tenhaken, R. Cell wall remodeling under abiotic stress. *Front. Plant Sci.* **2015**, *5*, 771. [CrossRef]
92. Chono, M.; Honda, I.; Shinoda, S.; Kushiro, T.; Kamiya, Y.; Nambara, E.; Kawakami, N.; Kaneko, S.; Watanabe, Y. Field studies on the regulation of abscisic acid content and germinability during grain development of barley: Molecular and chemical analysis of pre-harvest sprouting. *J. Exp. Bot.* **2006**, *57*, 2421–2434. [CrossRef] [PubMed]
93. Wang, J.; Wang, C.; Zhu, M.; Yu, Y.; Zhang, Y.; Wei, Z. Generation and characterization of transgenic poplar plants overexpressing a cotton laccase gene. *Plant Cell Tissue Organ Cult.* **2008**, *93*, 303–310. [CrossRef]
94. Shigeto, J.; Tsutsumi, Y. Diverse functions and reactions of class III peroxidases. *New Phytol.* **2016**, *209*, 1395–1402. [CrossRef]
95. Zipor, G.; Duarte, P.; Carqueijeiro, I.; Shahar, L.; Ovadia, R.; Teper-Bamnolker, P.; Eshel, D.; Levin, Y.; Doron-Faigenboim, A.; Sottomayor, M.; et al. *In planta* anthocyanin degradation by a vacuolar class III peroxidase in *Brunfelsia calycina* flowers. *New Phytol.* **2015**, *205*, 653–665. [CrossRef]
96. Martinez, M.; Gómez-Cabellos, S.; Giménez, M.J.; Barro, F.; Diaz, I.; Diaz-Mendoza, M. Plant Proteases: From key enzymes in germination to allies for fighting human gluten-related disorders. *Front. Plant Sci.* **2019**, *10*, 721. [CrossRef]
97. Berthet, S.; Thevin, J.; Baratiny, D.; Demont-Caulet, N.; Debeaujon, I.; Bidzinski, P.; Leple, J.; Huis, R.; Hawkins, S.; Gomez, L.D.; et al. Chapter 5—Role of plant laccases in lignin polymerization. In *Advances in Botanical Research*; Jouanin, J., Lapierre, C., Eds.; Academic Press: London, UK, 2012; Volume 61, pp. 145–172. [CrossRef]
98. Pant, S.; Huang, Y. Genome-wide studies of *PAL* genes in sorghum and their responses to aphid infestation. *Sci. Rep.* **2022**, *12*, 22537. [CrossRef]
99. Riaz, M.W.; Yousaf, M.I.; Hussain, Q.; Yasir, M.; Sajjad, M.; Shah, L. Role of lignin in wheat plant for the enhancement of resistance against lodging and biotic and abiotic stresses. *Stresses* **2023**, *3*, 434–453. [CrossRef]
100. Pellny, T.K.; Lovegrove, A.; Freeman, J.; Tosi, P.; Love, C.G.; Knox, J.P.; Shewry, P.R.; Mitchell, R.A.C. Cell walls of developing wheat starchy endosperm: Comparison of composition and RNA-seq transcriptome. *Plant Phyisol.* **2012**, *158*, 612–627. [CrossRef]