

Supplementary figure 1: Phylogenetic tree of the *Arabidopsis thaliana* **GT31 family clade I-VI based on Qu et al., 2008.** The amino acid sequences of *Arabidopsis thaliana* GT31 family members with (putative) roles in AGP glycan synthesis based on Qu et al., (2008). The amino acid sequences were aligned via ClustalW. Phylogenic trees were constructed using minimum-evolution method of MEGA-X in default mode with bootstrap test of 1000 replicates. The numbers beside the branches correspond to % bootstrap values.



Supplementary figure 2: RT-PCR analysis of transcripts from the first 10 cm of 20 cm long inflorescence stems from *Col-0, cage1-1/cage2-1* and *cage1-2/cage2-2*.



Supplementary figure 3: Cross sections of 10-week-old inflorescence stem from *Col-0* and *cage1cage2* lines stained with toluidine blue. *Col-0* (A, B), *cage1-1/cage2-1* (C, D) and *cage1-2/cage2-2* (E, F). The yellow arrows indicate collapsed xylem vessels. Scale bar A,C,E = 100 μ m and scale bar B,D,F = 50 μ m.



Supplementary figure 4: Cross sections of 10-week-old inflorescence stems from *cage1* and *cage2* lines stained with toluidine blue. *cage1-1* (A, C), *cage1-2* (B, D), *cage2-1* (E,G), *cage2-2* (F,H). Scale bar A,B,E,F = 100 μ m and scale bar C,D,G,H = 50 μ m.



Supplementary figure 5: Transmission Electron Microscope (TEM) images of transverse sections of xylem fibers and vessels of Col-0, cage2-1, cage2-2, cage1-1 and cage1-2. Scale bar = $5 \mu m$



Supplementary figure 6: Interfascicular fiber thickness of Col-O and the cage lines

In the boxplot, the middle horizontal line represents the median, the two boxes denote the 25th and 75th percentiles, the whiskers denote the 1.5 interquartile range limits. The dots represent the individual measurements (n= 4 biological replicates, 6 individual cells per biological replicate). Significance is calculated with a two-sided student's T-test. *** ≤ 0.001 .



Supplementary figure 7: Length of four-day old etiolated Col-0 and CAGE1-YFP and CAGE2-YFP lines. (A) Four-day old etiolated seedlings of Col-0 and *cage1-1cage2-1* lines expressing *cage1-YFP* or *cage2-YFP*. Scale bar = 5mm

(B) Hypocotyl length of 4-day old etiolated Col-0 and *cage1-1cage2-1* lines expressing *CAGE1-YFP* or *CAGE2-YFP*. In the boxplot, the middle horizontal line represents the median, the two boxes denote the 25th and 75th percentiles, the whiskers denote the 1.5 interquartile range limits. The dots represent the individual measurements (n = 25 biological replicates). Significance is calculated with a two-sided student's T-test. * ≤ 0.05 , *** ≤ 0.001 .



cage1-1/cage2-1

Supplementary figure 8: Ten-week-old Col-0 and *cage1-1cage2-1* plants transformed with *CAGE1-YFP* or *CAGE2-YFP* under their native promoters.



Supplementary figure 9: Crystalline cellulose content (% of dry weight) in 10-week-old Col-0 and *cage1-1cage2-1* lines expressing *CAGE1-YFP* or *CAGE2-YFP*.

In the boxplot, dark horizontal lines represent the median, the two grey boxes denote the 25th and 75th percentiles, the whiskers denote the 1.5 interquartile range limits, and the dots are outliers. (n = 6 biological replicates)





Lignin content (% of dry weight) in 10-week-old inflorescence stems. In the boxplot, dark horizontal lines represent the median, the two grey boxes denote the 25th and 75th percentiles, the whiskers denote the 1.5 interquartile range limits, and the dots are outliers. *P < 0.05, ***P < 0.001 (unpaired t test, n = 6 biological replicates).



Supplementary figure 11: Western blot analysis using CESA4, 7 and 8 specific antibodies. 10ug of protein was loaded from from the bottom 10 cm of 20 cm inflorescence stems of *Col-0* and the *cage1cage2* plants. The *CESA null* controls correspond to the published lines *irx5-4* (*CESA4*), *irx3-7* (*CESA7*) and *irx 1-7* (*CESA 8*).

CAGE1 CAGE2	MSAKIKGEYSSRSFVS <mark>KKWTILLCLGSFCVGMEETNR</mark> MWNIPESKGMSHPSVTEAE MSLKHHHRGLELSASKSFVS <mark>KKWTLFLCIGFFCAGILFSDR</mark> MWPEPESNVVSRDTVASDE ** : :*:****:*************************	56 60
CAGE1 CAGE2	RLKLVSEGCNPKAKEVKRDPQALFGEVANTHIALQTLDKTISSLEMELAAARSVQESLQN RLRLESEDCDSSKKGLKRESKDILGDVYKSPDAIQTLDKTISKLETELADARAAQESIMN **:* **.*: * :**: : ::*:* :: *:*********	116 120
CAGE1 CAGE2	GAPLSDDMGKKQPQEQRRFLMVVG <mark>INTAFSSRKRRDSIRATWMPQGEKRKRLEEEKGIII</mark> GSPVSDDFKLPETVTKRKYLMVVG <mark>VNTAFSSRKRRDSVRATWMPPGEERKKLEEEKGIVM</mark> *:*:****::::::*****::****************	176 180
CAGE1 CAGE2	RFVIGHSATTGGILDRAIEAEDRKHGDFLRLDHVEGYLELSGKTKTYFSTAFSMWDADFY RFVIGHSSTPGGILDRAIQAEESKHGDFLRLDHVEGYLELSAKTKTYFTTAFAMWDADFY ******** ****************************	236 240
CAGE1 CAGE2	VKVDDDVHVNIATLGETLVRHRKKPRVYIGCMKSGPVLSQKGVRYHEPEYWKFGENGNKY VKVDDDVHVNIATLGAELARYRMKPRVYIGCMKSGPVLAQKGVRYHEPEYWKFGEEGNKY ***************	296 300
CAGE1 CAGE2	FRHATGQLYAISRDLASYISINQHVLHKYANEDVSLGAWFIGIDVKHIDDRRLCCGTPPD FRHATGQLYAISRELASYISINQNVLHKYVNEDVSLGSWFLGLDVEHVDDRRLCCGT-TD ************************************	356 359
CAGE1 CAGE2	CEWKAQAGNICVASFDWSCSGICRSADRIKEVHRRCGEGEKALWSATF* 404 CEWKAQAGNICVASFDWSCSGICRSADRMKDVHRRCGEGEKALLAASF* 407 ************************************	

Supplementary figure 12: Amino acid sequence alignment of CAGE1 and CAGE2.

The amino acid sequences of CAGE1 and CAGE2 were aligned with ClustalW using standard settings . The sequence highlighted in red annotates the predicted membrane domain. The sequence highlighted in blue marks the predicted DUF4094 domain. The sequence highlighted in green marks the predicted galactosyltransferase domain.

Supplementary table 2: Primers used in this study				
Genotyping primers				
cage2-1	LP	TTTTCGAAACACCCAGATTTG		
cage2-2	RP	GAATTCCACCGGGAGTAGAAC		
cage1-1	LP	AAGGCATGTCTCATCCATCTG		
	RP	GTTTTCGTCTTGCCTGACAAC		
cage1-2	LP	ATCTTTGAGCTGTTTTTGGGG		
	RP	AATGTCTGCAGCAACTCTTGG		
LB1.3	LB	ATTTTGCCGATTTCGGAAC		
RT-PCR				
CAGE1_FW	ATGTCTGCGAAGATCAAAGG			
CAGE1_RV	TCAAAACGTGGCACTCCA			
CAGE2 FW	ATGTCGTTGAAGC	CATCATCA		
CAGE2_RV TCAGAAAGATGC		AGCCAG		
UBQ1_FW	BQ1_FW TTCCTTGATGATGCTTGCTC			
UBQ1_RV	JBQ1_RV TTGACAGCTCTTGGGTGAAG			
YFP constructs				
CAGE2-YFP	LP	AGAAAGCTGGGTCGGCGAAAGATGCAGCCA		
		GAAGGG		
	RP	CAAAAAGCAGGCTCCATGAAGCATGAGCT		
		ATGGGGA		
CAGE1-YFP	LP	AGAAAGCTGGGTCGGCAAACGTGGCACTC		
		CAAAGA		
	RP	CAAAAAGCAGGCTCCTCAGAAACCTGGT		
		TGAGAAAA		