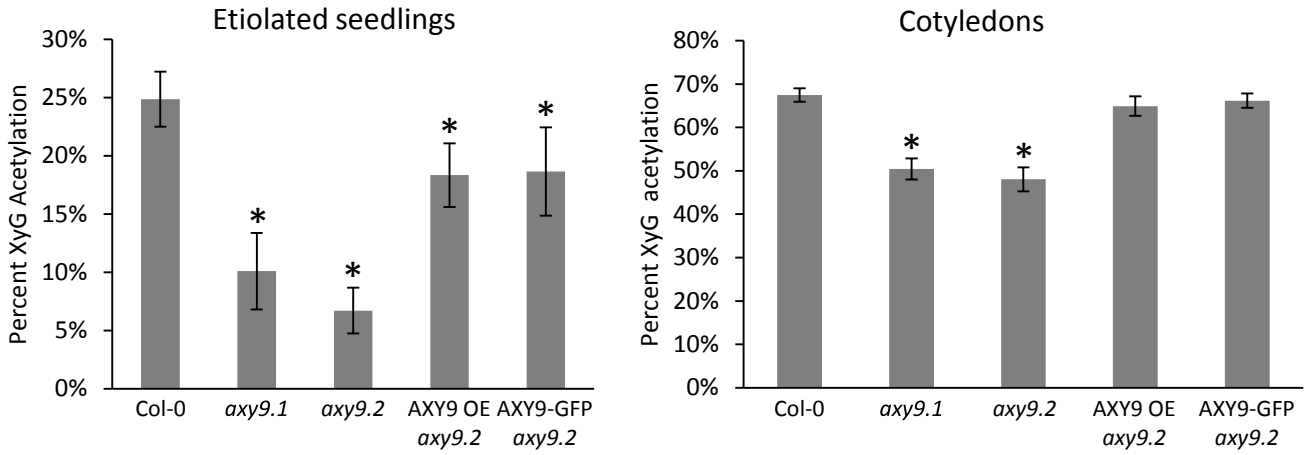
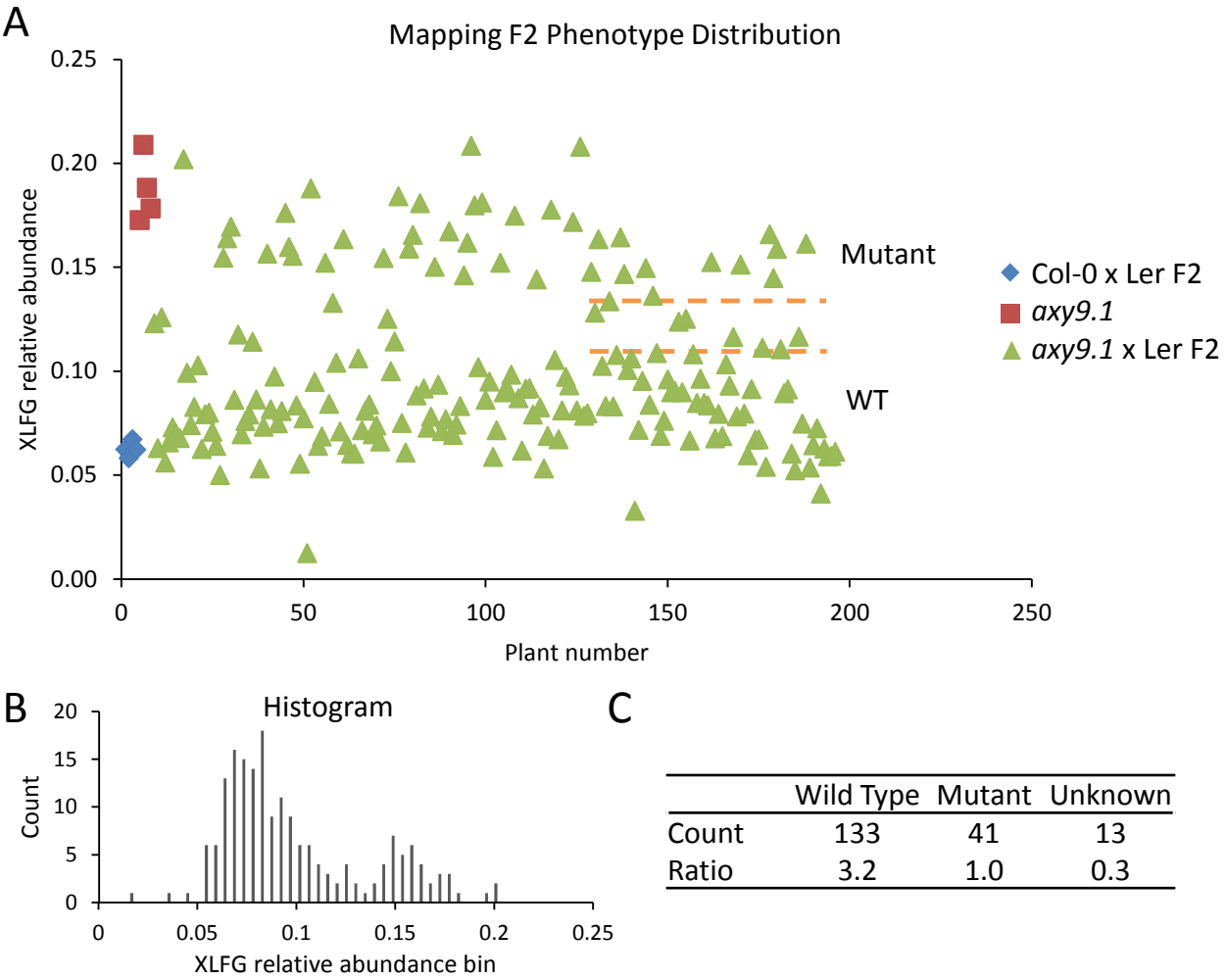


Supplemental Figure 1. Relative amount of XyG *O*-acetylation. Quantification is based on the relative abundance of XyG oligosaccharides detected by MALDI-TOF MS after digestion of wall material with a XyG-specific endoglucanase. The percent acetylation is the total relative quantity of the acetylated oligosaccharides (XXLG-Ac, XXFG-Ac, and XLFG-Ac) divided by the sum of oligosaccharides that can be *O*-acetylated (XXLG, XXLG-Ac, XXFG, XXFG-Ac, XLFG, XLFG-Ac). Error bars indicate standard deviation ($n \geq 6$); * indicates statistically significant difference from Col-0 (Student's T-Test, $p<0.01$).

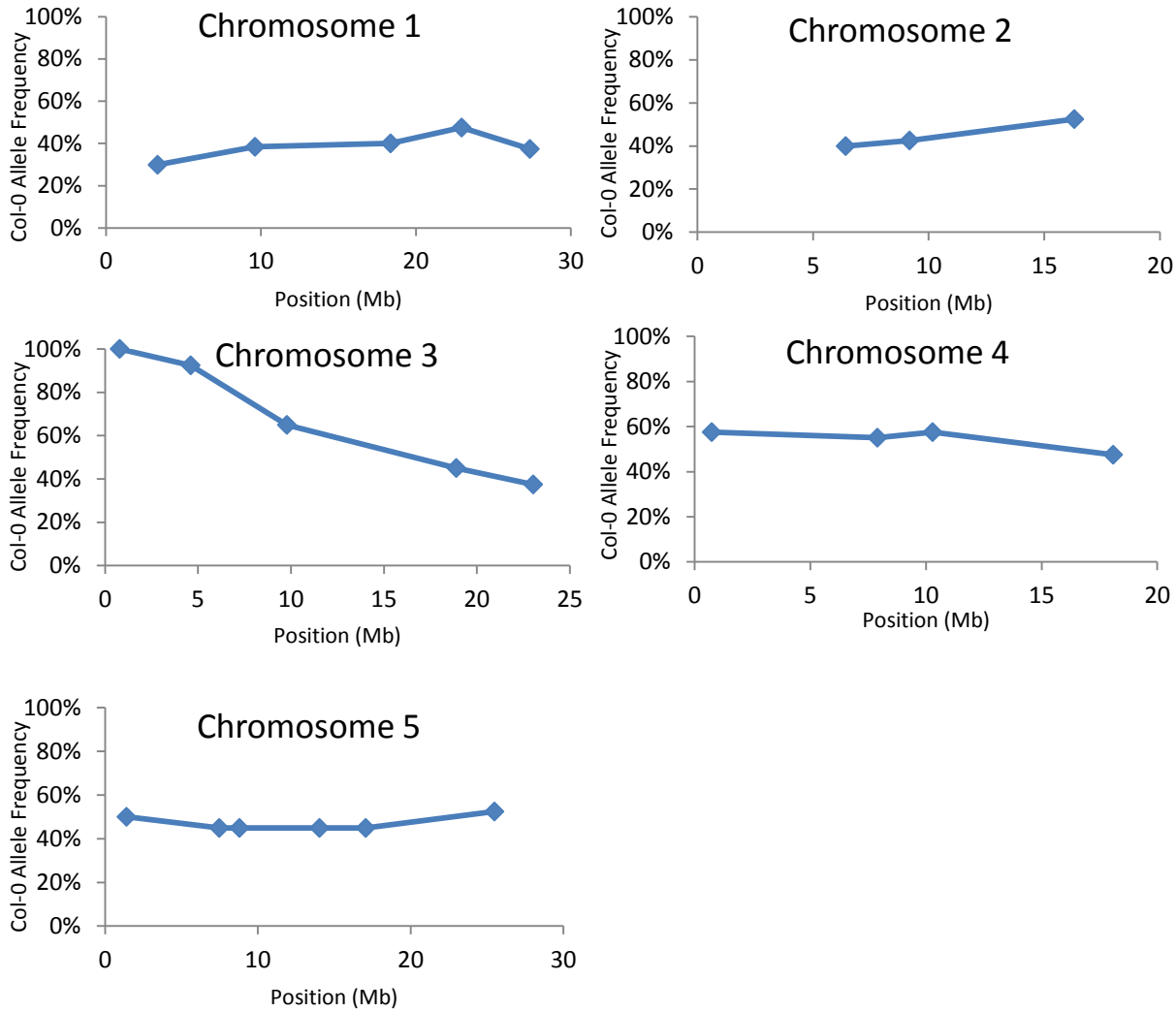


C

	Wild Type	Mutant	Unknown
Count	133	41	13
Ratio	3.2	1.0	0.3

Supplemental Figure 2. XyG oligosaccharide profiling of F2 plants from a cross between *axy9.1* and *Arabidopsis* ecotype Landsberg erecta (Ler). (A) Relative abundance of non-acetylated XLFG based on MALDI-TOF MS. Dashed lines indicate cutoff values used for calling the phenotype as mutant (*axy9.1*-like) or wild type (Col-0-like). (B) Histogram of the relative abundance of non-acetylated XLFG in the F2 population. (C) Segregation analysis of the F2 population.

A



B

Marker	9fm1	9fm2	NGA172	9fm4	9fm5	9fm8	NGA162
Position (Mb)	0.1	0.7	0.8	1.6	1.8	4.4	4.6
Plant 21	C	C	C	C	C	H	H
38	C	C	C	C	C	H	H
78	C	C	C	C	H	H	H
130	C	C	C	C	C	H	H
162	C	C	C	H	H	H	H
171	C	C	C	C	C	H	H

Mapped Region

Col-0 C Heterozygous H Landsberg erecta L

Supplemental Figure 3. Mapping of *axy9.1*. (A) Col-0 allele frequency based on SSLP marker analysis (Supplemental Table 1). (B) Fine mapping of the region of Col-0-enriched region on chromosome 3 using six mutant plants with recombination events within the region of interest.

Name	Position (Mb)	Primer sequence Forward	Reverse	Product length Col-0	Ler
Chromosome 1					
T27I1	3.3	GGATAAGTTGTCTCTCTAAGTGC	CTTTACTCTGATGGTGGTCC	199	164
CIW12	9.6	AGGTTTTATTGCTTTTCACA	CTTTCAAAAGCACATCACA	128	115
CIW1	18.4	ACATTTTCTCAATCCTTACTC	GAGAGCTTCTTTATTTGTGAT	159	135
F19K23-438	22.9	GGTCTAATTGCCGTTGTTGC	GAATTCTGTAACATCCCATTTCC	200	?
NGA111	27.4	TGTTTTTTAGGACAAATGGCG	CTCCAGTTGGAAGCTAAAGGG	128	162
Chromosome 2					
CIW3	6.4	GAAACTCAATGAAATCCACTT	TGAACTTGTTGTGAGCTTTGA	230	200
PLS5	9.2	GATGCCTTTCTCCTGGTTG	AATATAGCCGTCGTCTTCATCA	189	?
NGA168	16.3	GAGGACATGTATAGGAGCCTCG	TCGTCTACTGCACTGCCG	151	135
Chromosome 3					
9fm1	**0.1**	**GACAAATCACAATAAGCACA**	**TTCAAACCCTAATTATCGAC**	**482, 135**	**617**
9fm2	**0.7**	**CAAAGATCGGTAGTCAAACT**	**TTTGTATAGGCCCTAAGGTA**	**457, 389**	**846**
NGA172	0.8	CATCCGAATGCCATTGTTC	AGCTGCTTCCTTATAGCGTCC	162	136
9fm4	**1.6**	**CCTTCCTTTATTAGCTGTTG**	**CCAATATTACGAGTAAAGGC**	**322, 404**	**726**
9fm5	**1.8**	**GAGCTAATTCCCTCCTAAAT**	**GAAAACTTGGTTGGTAAGAA**	**345, 317**	**662**
9fm8	**4.4**	**ATCAGTAATGCTTCTATGGC**	**AAATGCTAGTTATGAGTGGG**	**304, 489**	**793**
NGA162	4.6	CTCTGTCACTCTTTTCCTCTGG	CATGCAATTTGCATCTGAGG	107	89
CIW11	9.8	CCCCGAGTTGAGGTATT	GAAGAAATTCCTAAAGCATTC	179	230
CIW4	18.9	GTTCATTAAACTTGCGTGTGT	TACGGTCAGATTGAGTGATTC	190	215
NGA6	23.0	ATGGAGAAGCTTACACTGATC	TGGATTTCTTCCTCTCTTCAC	143	123
Chromosome 4					
CIW5	0.7	GGTTAAAAATTAGGGTTACGA	AGATTTACGTGGAAGCAAT	164	144
CIW6	7.9	CTCGTAGTGCACTTTCATCA	CACATGGTTAGGGAAACAATA	162	148
F28A21-23	10.3	CTTCGGACGCATAAAAGAGTC	CAAGACGCCAATATCCGTAC	131	108
NGA1107	18.1	CGACGAATCGACAGAATTAGG	GCGAAAAAACAAAAAAATCCA	150	140
Chromosome 5					
MUK11D	1.4	GGAGAAGGCTTTGTGTCTGTATC	CTTTCTCTTCCACTGAATCTCCTC	402	390
CIW8	7.5	TAGTGAAACCTTTCTCAGAT	TTATGTTTTCTTCAATCAGTT	100	135
NGA139	8.8	GGTTTCGTTTCACTATCCAGG	AGAGCTACCAGATCCGATGG	174	132
PHYC.3	14.0	AAACTCGAGAGTTTTGTCTAGATC	CTCAGAGAATTCCCAGAAAAATCT	207	222
CIW9	17.1	CAGACGTATCAAATGACAAATG	GACTACTGCTCAAACTATTCGG	165	145
MBK-5	25.5	GAGCATTTCACAGAGACG	ATCACTGTTGTTTACCATTA	207	180

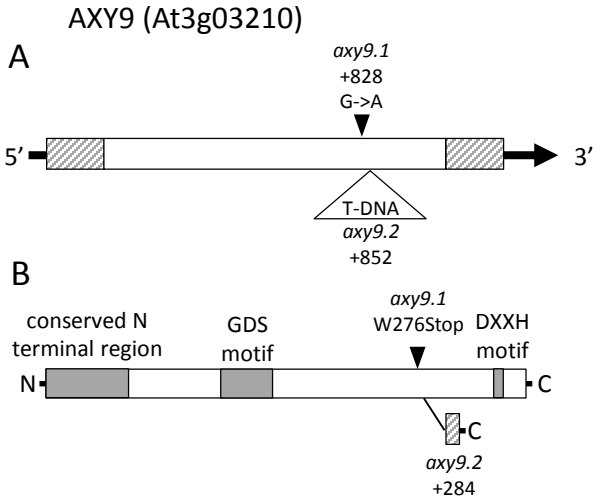
Supplemental Table 1. Markers used for genetic linkage analysis of *axy9.1*. Simple sequence length polymorphism (SSLP) markers were were obtained from www.arabidopsis.org. Cleaved amplified polymorphic sequence (CAPS) markers (bold) were based on the Col-0 reference genome and a list of SNPs present in the Landsberg erecta ecotype (www.arabidopsis.org). CAPS markers require digestion with EcoRI following the PCR reaction.

Criteria (additive)	# Candidate SNPs
Total SNPs	8827
Unique to *axy9.1*	1987
Homozygous	652
Causing AA change	148
Mapped region (Chr3, 0-1.6 Mb)	2

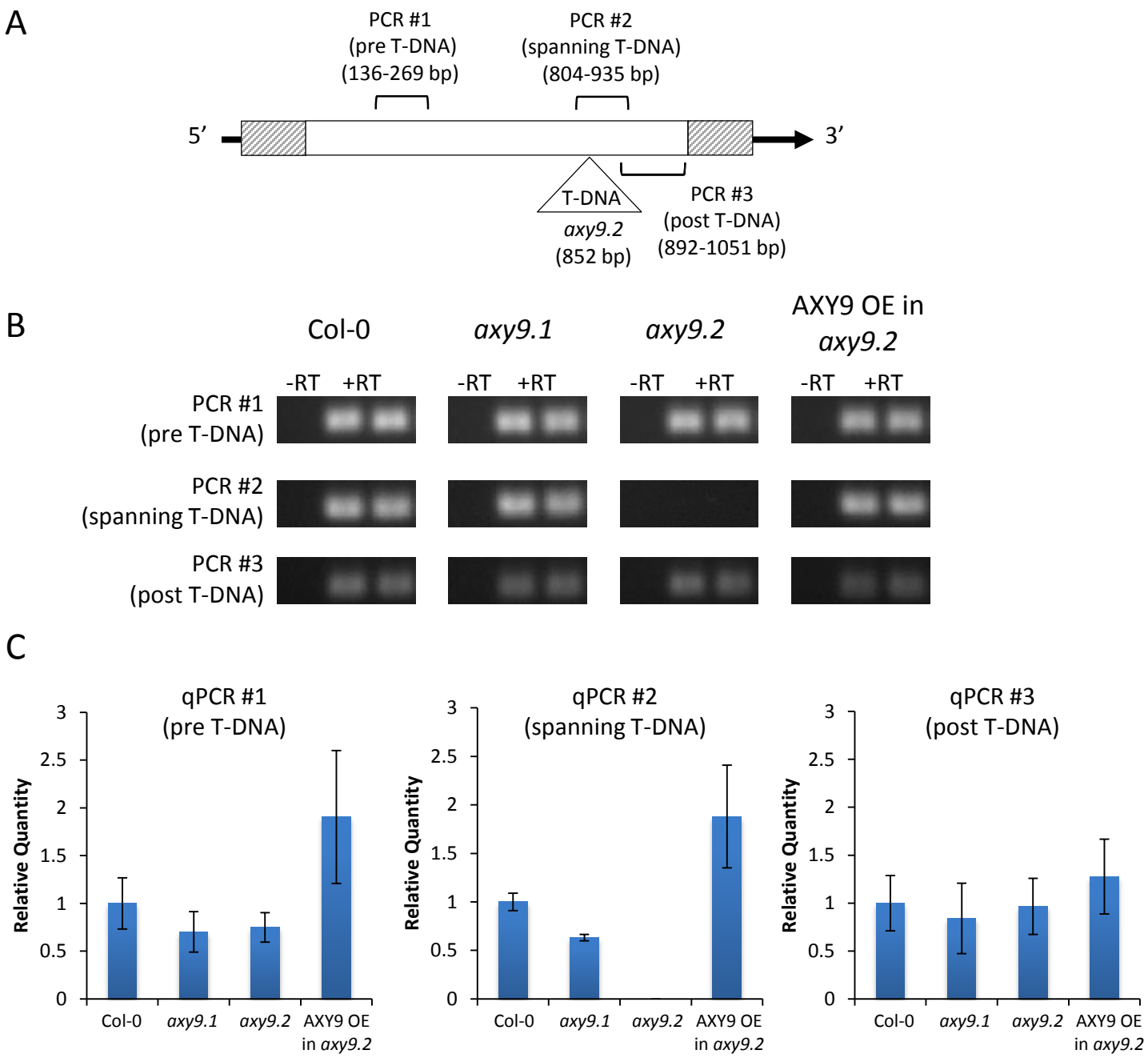
Supplemental Table 2. Genomic sequencing of *axy9.1.* Detected SNPs, relative to the Col-0 reference genome, were filtered against sequencing data from other *axy* mutants (Günl et al., 2011, Gille et al., 2011) to obtain a specific *axy9.1* SNP set. 652 of these SNPs were found to be homozygous based on detection frequency and 148 of the homozygous SNPs were predicted to result in amino acids changes. Of these 148 SNPs, two were located within the mapped region on chromosome 3.

Position (bp)	Gene	AA change	Gene Annotation	XyG Acetylation[1] (relative to wild type)
121,177	At3g01323		ECA1-like gametogenesis related family protein	
210,881	At3g01530	Thr62Ile	Member of the R2R3 factor gene family	98 ± 8
364,726	At3g02080		Ribosomal protein S19e family protein	
494,882	At3g02410		Isoprenylcysteine methylesterase-like 2 (ICME-LIKE2)	104 ± 7
603,885	At3g02780		Isopentenyl diphosphate:dimethylallyl diphosphate isomerase	
741,814	At3g03210	Trp276Stop	unknown protein	69 ± 2
1,401,626	At3g05040		Member of importin/exportin family. Involved in shoot maturation.	
1,506,767	At3g05290		Peroxisomal adenine nucleotide transporter, involved in fatty acid beta-oxidation	

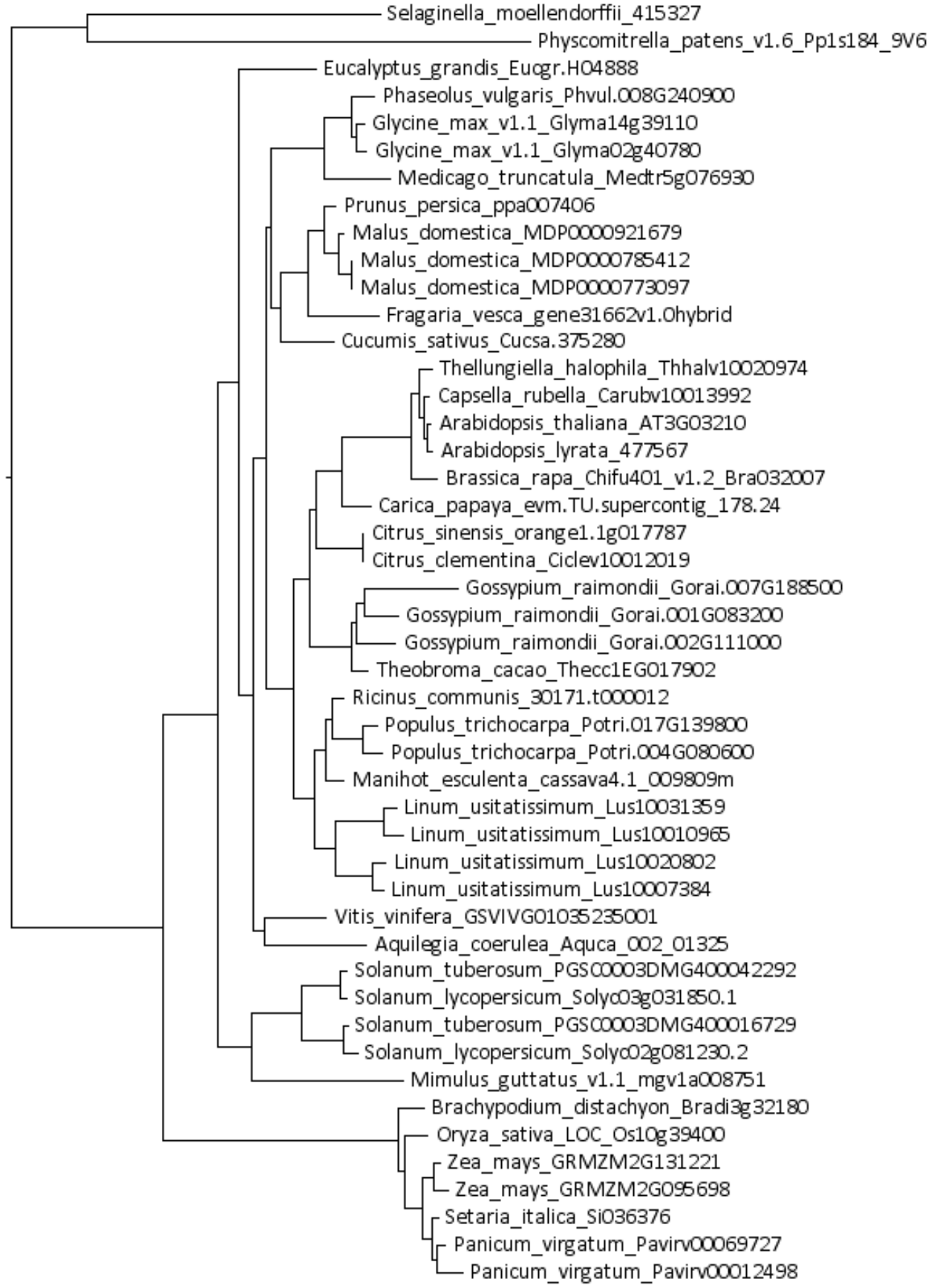
Supplemental Table 3. Candidate SNPs for the basis of the *axy9.1* phenotype in the mapped region. Two out of the eight SNPs detected in genes within the mapped region result in predicted amino acid changes. T-DNA lines with insertions in these candidate genes were obtained and XyG acetylation was measured of leaf tissue relative to wild type. [1]Percent XyG acetylation of possible as a percentage relative to wild type, ± standard deviation, n = 4.



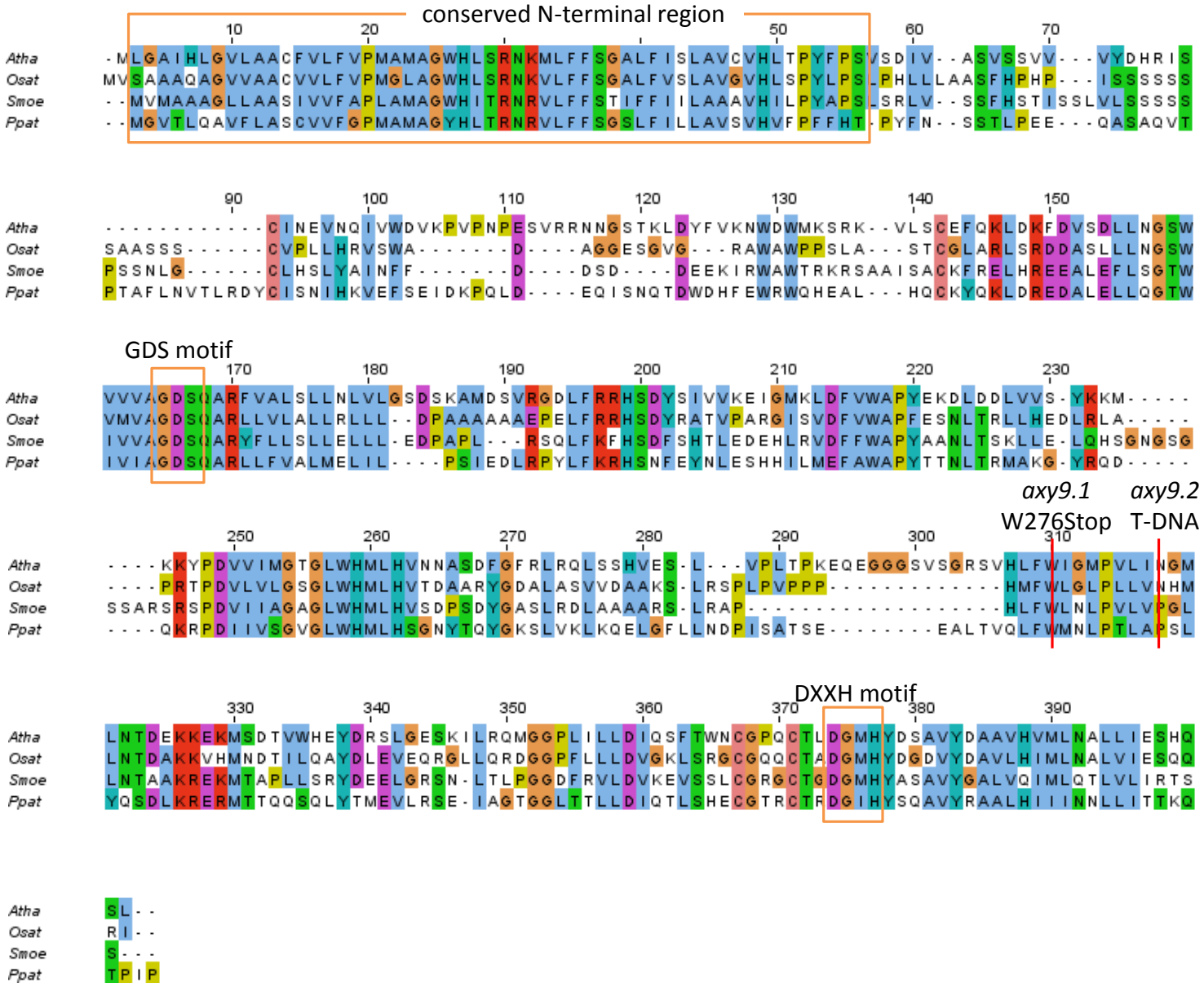
Supplemental Figure 4. Gene and protein model of AXY9. (A) Gene model of *AXY9* (*At3g03210*) including the introduced point mutation in *axy9.1* and T-DNA insertion location in *axy9.2*. Striped boxes – untranslated regions; white box – coding sequence. (B) Protein model of AXY9 including the stop codon in *axy9.1* and the insertion in *axy9.2*. Shaded grey box – protein regions of interest. Striped box – aberrant amino acids for the predicted protein in *axy9.2.*



Supplemental Figure 5. Expression profiling of *AXY9* in the *AXY9* mutants in leaves. (A) Gene model of *AXY9* showing the location of the T-DNA insertion in the *axy9.2* allele and the positions of three PCR primer pairs before, spanning, and after the T-DNA insertion (PCR #1, #2, and #3 respectively) relative to the direction of transcription. (B) Qualitative RT-PCR using the three primer pairs for the indicated genotypes. Reverse transcriptase enzyme was include (+RT) or excluded (-RT) from cDNA preparations to check for the presence of genomic DNA contamination. Two independent replicates are shown for the cDNA samples prepared with reverse transcriptase. (C) Quantitative RT-PCR using the three primer pairs with SYBR green for quantification during amplification. The results were normalized to an internal control (*ACTIN 2*, *At3g18780*) and to Col-0. Error bars indicate standard error (n = 3).

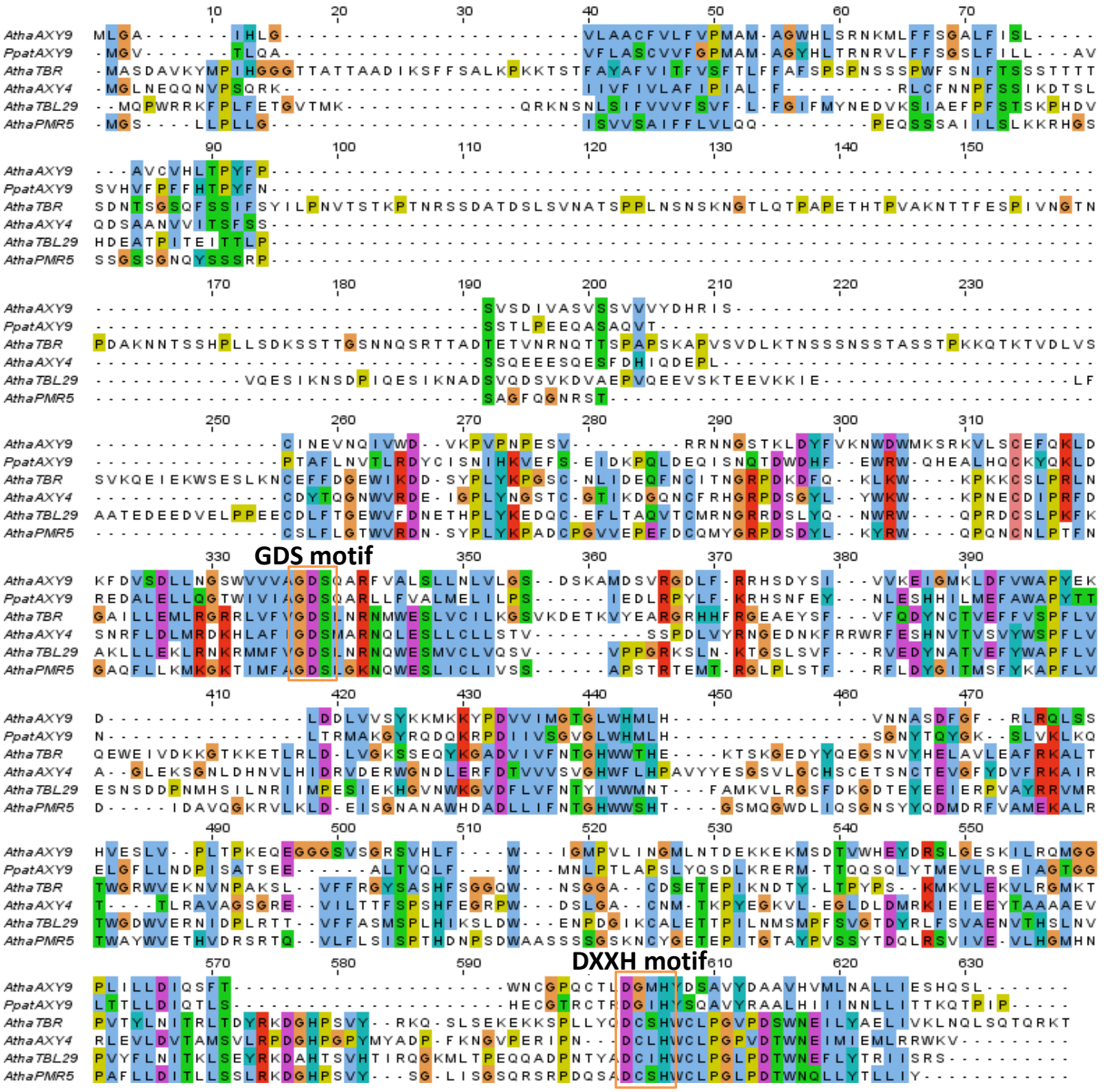


Supplemental Figure 6. AXY9 maximum-likelihood phylogenetic tree. Between one and four putative AXY9 orthologs were identified in each land plant species from the Phytozome database.



Supplemental Figure 7. Alignment of AXY9 protein sequences. Protein sequences were obtained from Phytozome (www.phytozome.net) and aligned using MUSCLE (Edgar, 2004). Regions of note include the highly conserved N-terminus, the putative catalytic GDS motif (position 157), and the conserved DXXH motif at the C-terminus. The highly conserved N-terminus contains two hydrophobic stretches separated by four hydrophilic residues (SRNK in Arabidopsis). The mutation present in *axy9.1*, W276Stop, is marked with a red line. *Atha – Arabidopsis thaliana, Osat – Oryza sativa*, *Ppat – Physcomitrella patens*, *Smoe – Selaginella moellendorffii*.

Edgar, Robert C. "MUSCLE: multiple sequence alignment with high accuracy and high throughput." *Nucleic acids research* 32.5 (2004): 1792-1797.

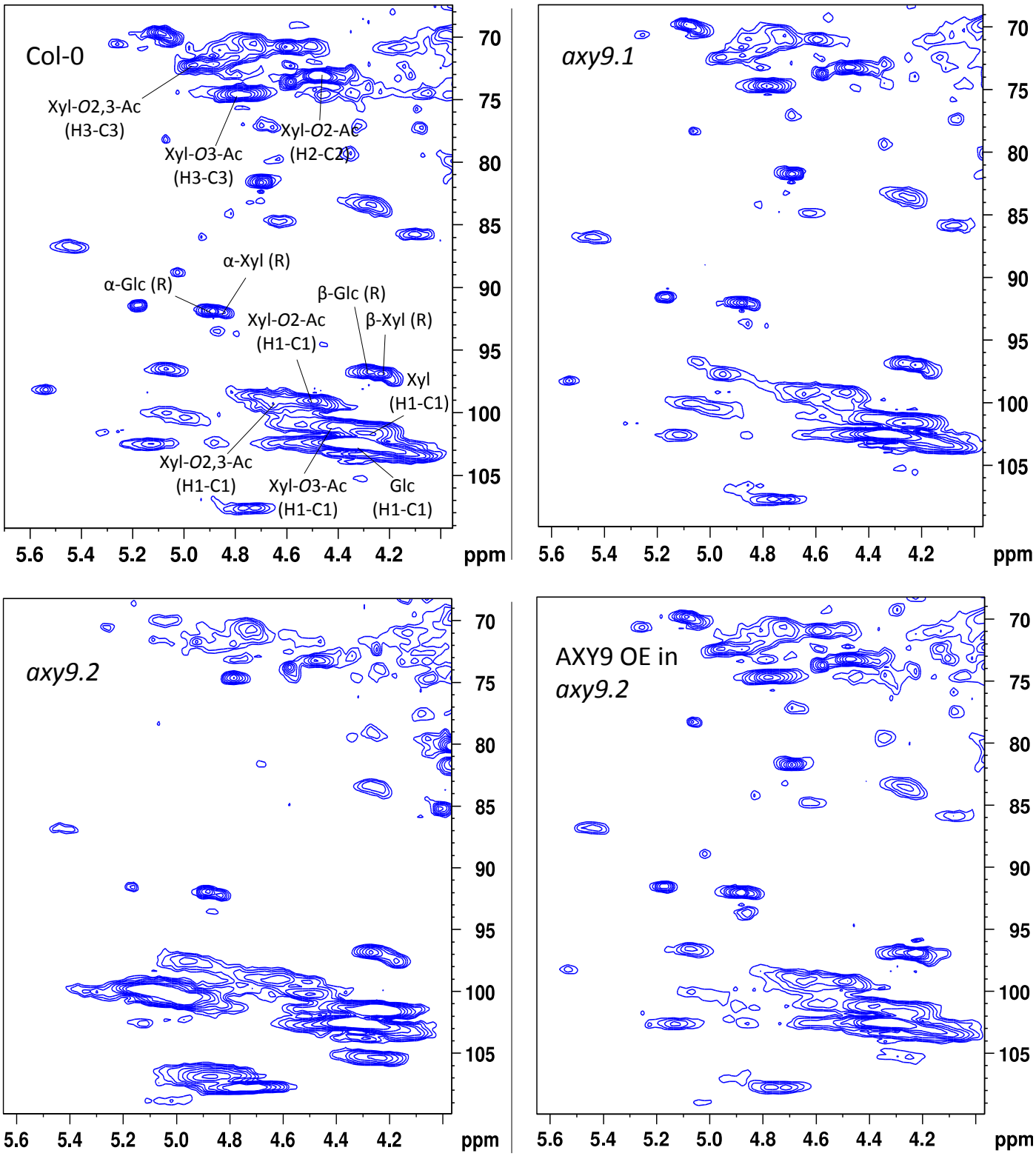


Pairwise Sequence Identity					
	PpatAXY9	AthaTBR	AthaAXY4	AthaTBL29	AthaPMR5
AthaAXY9	35	14	17	14	16
PpatAXY9		14	17	15	18
AthaTBR			21	26	26
AthaAXY4				24	27
AthaTBL29					31

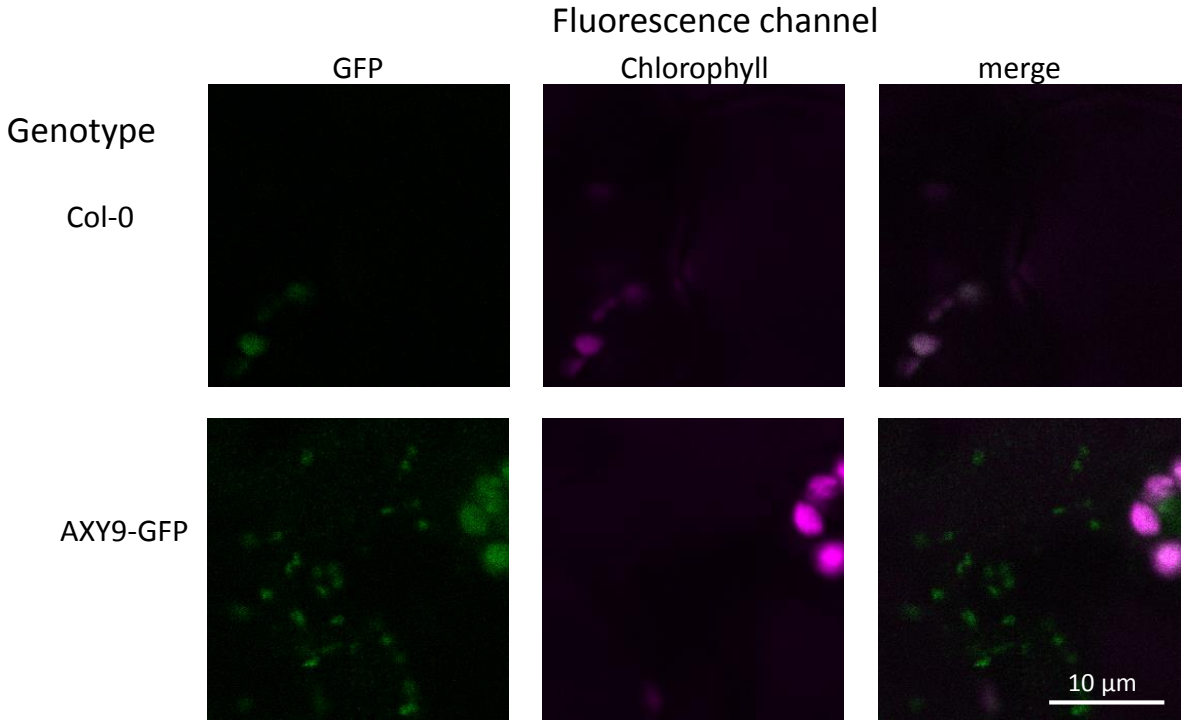
Pairwise Sequence Similarity					
	PpatAXY9	AthaTBR	AthaAXY4	AthaTBL29	AthaPMR5
AthaAXY9	57	26	32	28	30
PpatAXY9		26	31	29	32
AthaTBR			34	40	37
AthaAXY4				39	41
AthaTBL29					43

Supplemental Figure 8. Alignment and sequence similarity of AXY9 and TBL protein sequences. Multiple sequence alignment (top) and pairwise percent identity and similarity values (bottom) of protein sequences obtained from Phytozome (www.phytozome.net). Multiple and pairwise alignments were done using MUSCLE (Edgar, 2004). Conservation is observed in the GDS and to a lesser extent the DXXH motifs. *Atha – Arabidopsis thaliana, Ppat – Physcomitrella patens*.

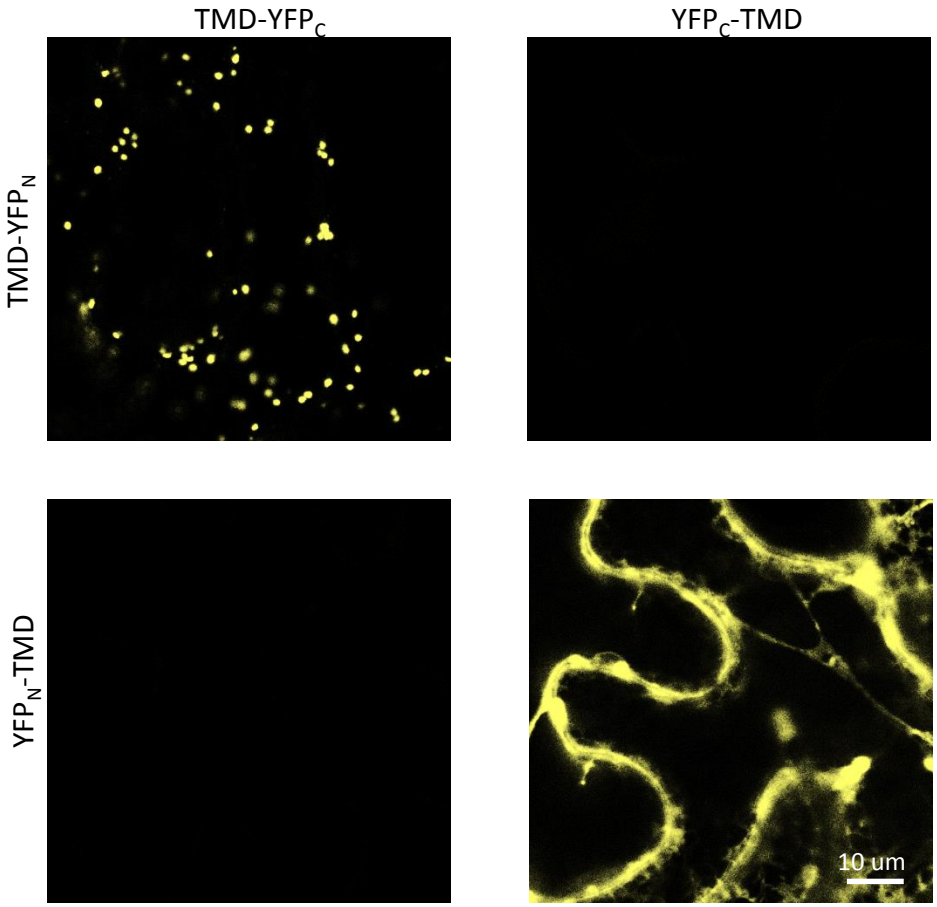
Edgar, Robert C. "MUSCLE: multiple sequence alignment with high accuracy and high throughput." *Nucleic acids research* 32.5 (2004): 1792-1797.



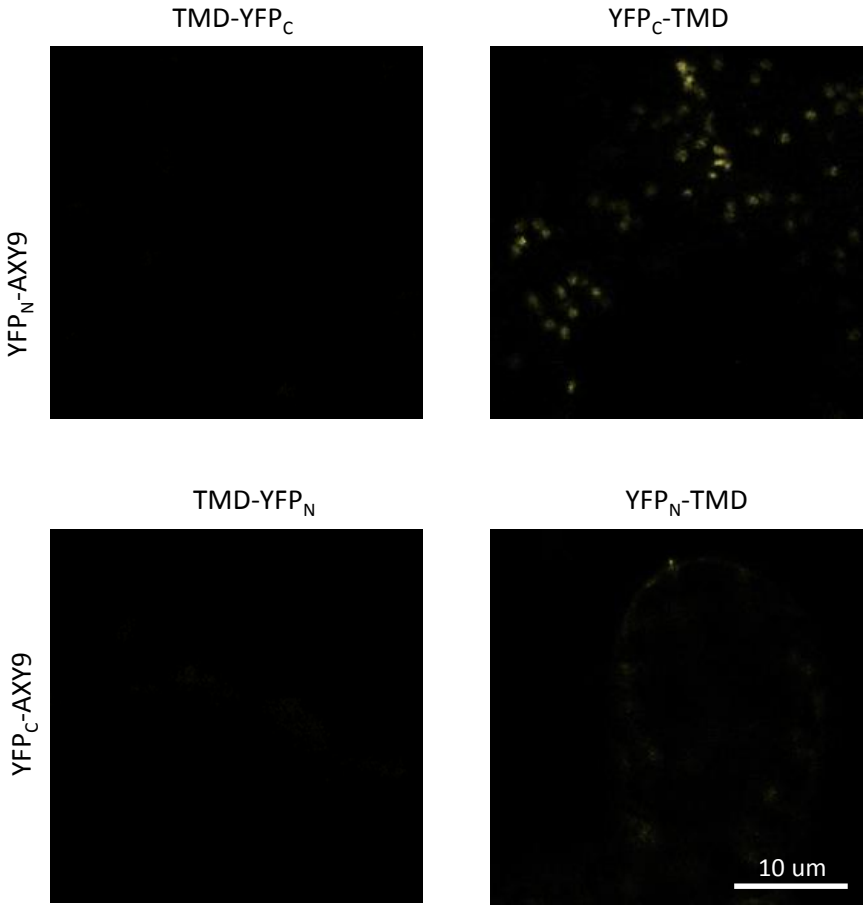
Supplemental Figure 9. 2D HSQC NMR spectrum of dissolved stem tissue. The ^{1}H (x-axis) / ^{13}C (y-axis) NMR spectra for wild type (WT), Per-*O*-acetylated arabinoxylan from wheat, *axy9.1*, *axy9.2* and AXY9 OE in *axy9.2* mutants. Prominent peaks corresponding to known polysaccharide linkages are labeled (Cheng et al., 2013 & Chong et al., 2014). A quantification of the peaks related to acetylated polysaccharides is shown in Table 1.



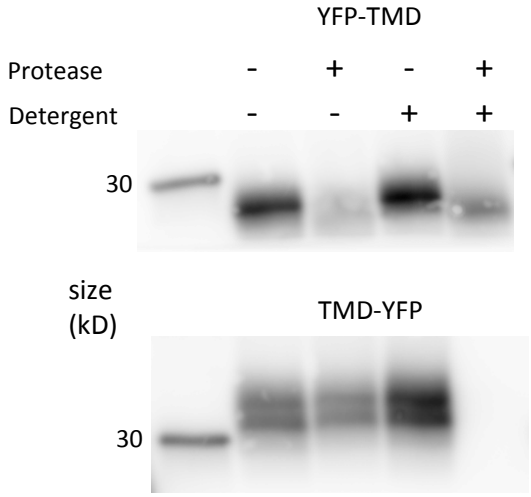
Supplemental Figure 10. Subcellular localization of AXY9-GFP in *Arabidopsis.* Confocal microscopy was used to image leaf tissue of stably transformed *Arabidopsis* plants expressing AXY9-GFP.



Supplemental Figure 11. Split-YFP controls. Confocal microscopy was used to image *N. benthamiana* leaf tissue infiltrated with split-YFP fusion constructs.



Supplemental Figure 12. Split-YFP AXY9 N-terminal fusions. Split-YFP AXY9 protein fusions (YFP$_N$-AXY9 and YFP$_C$-AXY9) were coexpressed with TMD-YFP$_{N/C}$ and YFP$_{N/C}$-TMD in *N. benthamiana*. Leaf tissue was imaged for YFP fluorescence using a confocal microscope.



Supplemental Figure 13. Protease protection assay of internal and external Golgi controls. The YFP-TMD and TMD-YFP peptide fusions have been reported to result in cytosolic and lumenal orientations respectively for the YFP domain (Søgaard et al., 2012).

Purpose	Sequence
Genotyping *axy9.2* (mutant forward)	TGGTTCACGTAGTGGGCCATCG
Genotyping *axy9.2* (WT forward)	TGGGTCAGACTTTTGCATTTC
Genotyping *axy9.2* (reverse)	TTGAATCTTGTTTTGGGTTCG
Genotyping AXY9 OE (forward)	CAAACTCTCTCGAACCTTCC
Genotyping AXY9-GFP (forward)	TGTAAAACGACGGCCAGT
Genotyping AXY9 OE, AXY9-GFP (reverse)	TTGATACAAGAGATACGATG
forward for AXY9 into pDONR221 P1-P2	GGGGACAAGTTTGTACAAAAAAGCAGGCTatgctaggagcgattcatttg
reverse for AXY9 into pDONR221 P1-P2	GGGGACCACTTTGTACAAGAAAGCTGGGTtaaagattgatgagattcaataag
forward to clone entcup2 into P1-P3r	GGGGACAAGTTTGTACAAAAAAGCAGGCTgggatcttctgcaagca
reverse to clone entcup2 into P1-P3r	GGGGACAACTTTATTATACAAAGTTGtccggtgggttttgagg
forward for axy9 in p3-p2	GGGGACAACTTTGTATAATAAAGTTGatgctaggagcgattcatttg
reverse for axy9 in P3-P2	GGGGACCACTTTGTACAAGAAAGCTGGGTAtcataaagattgatgagattcaataag
forward for AXY9-GFP into pICH31070	actttttgttcttattgttgcaggtATGCTAGGAGCGATTCATTTG
forward for AXY9-GFP into pICH31070	tgcgcaccacgcctagtcttcaagcTTAGTGGTGGTGGTGGTG
forward for AXY9-YFP$_N$, forward for AXY9-YFP$_C$	aaccgggctcaggcctggcgcgccaATGCTAGGAGCGATTCATTTG
reverse for AXY9-YFP$_N$	ttgctccatcccgggagcggtacccgtTAAAGATTGATGAGATTCAATAAGC
reverse for AXY9-YFP$_C$	tgggtacatcccgggagcggtacccgtTAAAGATTGATGAGATTCAATAAGC
forward for YFP$_N$-AXY9	tctgaggaggatcttgggcccaggcctagcATGCTAGGAGCGATTCATTTG
reverse for YFP$_N$-AXY9, YFPC-AXY9 with stop	gagctcctacccgggagcggtacccTCATAAAGATTGATGAGATTCAATAAG
forward for YFP$_C$-AXY9	ccagattacgctgggcccaggcctagcATGCTAGGAGCGATTCATTTG
forward for tmd-YFP into pICH31070	actttttgttcttattgttgcaggtATGATTCATACCAACTTGAAGAAAAAG
reverse for YFP-tmd into pICH31070	tgcgcaccacgcctagtcttcaagcTCAGGCCACTTTCTCCTGGC
forward for TMD-YFP$_N$, TMD-YFP$_C$	aaccgggctcaggcctggcgcgccaATGATTCATACCAACTTGAAGAAAAAG
reverse for TMD-YFP$_N$	ttgctccatcccgggagcggtacccgtGGCCACTTTCTCCTGGC
reverse for TMD-YFP$_C$	tgggtacatcccgggagcggtacccgtGGCCACTTTCTCCTGGC
forward for YFP$_N$-TMD	tctgaggaggatcttgggcccaggcctagcATGATTCATACCAACTTGAAGAAAAAG
forward for YFP$_C$-TMD	ccagattacgctgggcccaggcctagcATGATTCATACCAACTTGAAGAAAAAG
reverse for YFP$_N$-TMD, YFP$_C$-TMD with stop	gagctcctacccgggagcggtacccTCAGGCCACTTTCTCCTGGC
reverse for tmd-YFP into pICH31070	tgcgcaccacgcctagtcttcaagcTTACTTGTACAGCTCGTCCA
forward for YFP-tmd into pICH31070	actttttgttcttattgttgcaggtATGGTGAGCAAGGGCGAG
RT-PCR and qRT-PCR primer pair 1 forward	TGTGTTCATCTCACGCCTTA
RT-PCR and qRT-PCR primer pair 1 reverse	GGTACCGGTTTAACATCCCA
RT-PCR and qRT-PCR primer pair 2 forward	CGGTAGATCTGTGCATCTGT
RT-PCR and qRT-PCR primer pair 2 reverse	TTACTCTCCCCGAGTGATCT
RT-PCR and qRT-PCR primer pair 3 forward	GATACGGTGTGGCATGAGTA
RT-PCR and qRT-PCR primer pair 3 reverse	CATCATAGACCGCAGAGTCA
qRT-PCR internal control (*At3g18780*) forward	TCTTCCGCTCTTTCTTTCCAAGC
qRT-PCR internal control (*At3g18780*) reverse	ACCATTGTCACACACGATTGGTTG

Supplemental Table 4. Primer sequences used for genotyping, expression profiling and plasmid construction.

Description	Template	Backbone	Construction Method	Template Reference	Backbone reference
Split-YFP					
TMD-YFP_N	TMD-YFP_N	GW-VYNE	Gibson (SpeI, XhoI)	Søgaard et al., 2012	Gehl et al., 2009
TMD-YFP_C	TMD-YFP_C	GW-VYCE	Gibson (SpeI, XhoI)	Søgaard et al., 2012	Gehl et al., 2009
YFP_N-TMD	YFP_N-TMD	VYNE-GW	Gibson (SpeI, XhoI)	Søgaard et al., 2012	Gehl et al., 2009
YFP_C-TMD	YFP_C-TMD	VYCE-GW	Gibson (SpeI, XhoI)	Søgaard et al., 2012	Gehl et al., 2009
AXY9-YFP_N	AXY9 cds	GW-VYNE	Gibson (SpeI, XhoI)		Gehl et al., 2009
YFP_N-AXY9	AXY9 cds	VYNE-GW	Gibson (SpeI, XhoI)		Gehl et al., 2009
AXY9-YFP_C	AXY9 cds	GW-VYCE	Gibson (SpeI, XhoI)		Gehl et al., 2009
YFP_C-AXY9	AXY9 cds	VYCE-GW	Gibson (SpeI, XhoI)		Gehl et al., 2009
Over-expression / complementation					
entcup2 entry	pORE E4	pDONR221 P1-P3r	Gateway BP	Coutu et al., 2007	Gateway Manual (Life Technologies)
AXY9 entry	AXY9 cds	pDONR221 P3-P2	Gateway BP		Gateway Manual (Life Technologies)
entcup2:AXY9		pGWB1	Gateway LR (2 piece)		Nakagawa et al., 2007
Localization					
AXY9 entry	AXY9 cds	pDONR 221 P1-P2	Gateway BP		Gateway Manual (Life Technologies)
AXY9-GFP		pMDC84	Gateway LR		Curtis et al., 2003
Protease protection					
AXY9-GFP	AXY9-GFP	pICH31070	Gibson (BsaI)		Icon Genetics, Germany
YFP-TMD	YFP-TMD	pICH31070	Gibson (BsaI)	Søgaard et al., 2012	Icon Genetics, Germany
TMD-YFP	TMD-YFP	pICH31070	Gibson (BsaI)	Søgaard et al., 2012	Icon Genetics, Germany

Supplemental Table 5. Plasmid construction. The cloning method and destination vector used for construction of each plasmid are listed along with the template DNA and relevant references.

Coutu, Catherine, et al. "pORE: a modular binary vector series suited for both monocot and dicot plant transformation." *Transgenic research* 16.6 (2007): 771-781.

Curtis, Mark D., and Grossniklaus, U. "A gateway cloning vector set for high-throughput functional analysis of genes in planta." *Plant physiology* 133.2 (2003): 462-469.

Gehl, Christian, et al. "New GATEWAY vectors for high throughput analyses of protein–protein interactions by bimolecular fluorescence complementation."*Molecular plant* 2.5 (2009): 1051-1058.

Nakagawa, Tsuyoshi, et al. "Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation." *Journal of bioscience and bioengineering* 104.1 (2007): 34-41.

Søgaard, Casper, et al. "GO-PROMTO illuminates protein membrane topologies of glycan biosynthetic enzymes in the Golgi apparatus of living tissues." *PloS one* 7.2 (2012): e31324.