UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Identification and characterization of genes involved in the biosynthesis of the plant cell wall polysaccharide xyloglucan

Permalink
https://escholarship.org/uc/item/2m42k754

Author
Schultink, Alexander

Publication Date
2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Identification and characterization of genes involved in the biosynthesis of the plant cell wall polysaccharide xyloglucan

By

Alexander Christiaan Schultink

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Plant Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Markus Pauly, Chair
Professor Chris Somerville
Professor Michelle Chang

Fall 2013

Abstract

Identification and characterization of genes involved in the biosynthesis of the plant cell wall polysaccharide xyloglucan

by

Alexander Christiaan Schultink

Doctor of Philosophy in Plant Biology

University of California, Berkeley

Professor Markus Pauly, Chair

Xyloglucan (XyG) represents the most abundant hemicellulose in the primary cell wall of many dicot and non-graminaceous monocot plant species. Found throughout the land plant lineage, the structure of this polysaccharide varies by species, tissue, and developmental stage. Some proteins responsible for the biosynthesis and metabolism of XyG have been previously identified but additional genes remain to be uncovered. In the work presented here, several genetic and genomic approaches were used to identify previously unknown genes involved in XyG biosynthesis and metabolism. A forward genetic screen of mutagenized *Arabidopsis thaliana* resulted in the identification of the *Altered Xyloglucan 9* (*AXY9*) gene, coding for a protein of unknown function, as part of an apparently plant-specific pathway for the acetylation of polysaccharides. The AXY9 protein may be part of a system for shuttling activated acetyl groups across the Golgi membrane for use by polysaccharide-specific acetyltransferases. A transcriptional profiling approach used on developing XyG-rich *Tropaeolum majus* seeds revealed a galactosyltransferase that acts on XyG, as demonstrated by mutant analysis in *A. thaliana*. By utilizing comparative genomics, two putative XyG arabinofuranosyltransferase genes were discovered from tomato, which are able to act *in vivo* to arabinosylate XyG. Expression of these genes rescued mechanical and morphological phenotypes of an *A. thaliana* mutant deficiency for XyG galactosylation, providing insight into the function of XyG substitution. To provide a testing platform for candidate genes involved in XyG biosynthesis and to investigate the activities of these genes independently from endogenous plant proteins, an attempt was made to reconstruct the XyG biosynthetic pathway in a heterologous host. This effort resulted in the production of a glucan backbone and the UDP-xylose donor substrate required for the xylosyltransferases, but a XyG polysaccharide was not detected.

Table of Contents

List of figures ix

List of tables xi

1 Introduction 1

1.1 Plant cell wall polysaccharides 1

1.1.1 Function 1

1.1.2 Structures 2

1.1.3 Biosynthesis 5

1.2 XyG 7

1.2.1 Function 7

1.2.2 Structures and diversity 8

1.2.3 XyG Biosynthesis 11

1.2.4 XyG Metabolism 13

1.2.5 Applications 13

1.3 Thesis objective 14

2 Materials and methods 15

2.1 Media and growth conditions 15

2.1.1 Growth media 15

2.1.2 Induction of transgene expression in *Pichia pastoris* 16

2.1.3 Sterilization of *A. thaliana* seed surface 16

2.1.4 Growth of *A. thaliana* on plates and soil 16

2.1.5 Growth of *N. benthamiana* on soil 16

2.2 Molecular biology 16

2.2.1 Extraction of genomic DNA from plant tissue 16

2.2.2 Genotyping *A. thaliana* and *Pichia pastoris* using PCR 16

2.2.3 Genetic mapping with SSLP and CAPS markers 17

2.2.4 Amplification of genes for cloning using PCR 17

2.2.5 Cloning methods 17

2.2.6 Restriction digest analysis of plasmids 18

2.2.7 DNA sequencing of plasmids and gene fragments 18
2.2.8 Transformation of *P. pastoris* 18
2.2.9 Excision of selectable marker from genome of *P. pastoris* 19
2.2.10 Transformation of *A. tumefaciens* 19
2.2.11 Transformation of *A. thaliana* 19
2.2.12 Infiltration of *N. benthamiana* 19
2.2.13 Transcriptional analysis using qRT-PCR 20
2.2.14 Transcriptional analysis using RT-PCR 20
2.2.15 High-throughput sequencing of *A. thaliana* 20
2.3 Microscopy 20
2.3.1 Subcellular localization 20
2.3.2 Light microscopy (for stem sections) and staining 21
2.4 Bioinformatic methods 21
2.4.1 Analysis of high-throughput sequencing data 21
2.4.2 Phylogenetic tree construction 21
2.5 Analytical methods 21
2.5.1 Preparation of alcohol insoluble residue (AIR) from plant tissue 21
2.5.2 Preparation of alcohol insoluble residue (AIR) from *P. pastoris* 22
2.5.3 Analysis of XEG-released oligosaccharides by MALDI-TOF MS (XyG OLIMP) 22
2.5.4 Analysis of XEG-released oligosaccharides using HPAEC-PAD 22
2.5.5 Desalting / purification of XyG oligosaccharides 23
2.5.6 Purification of novel oligosaccharide (XXSG) 23
2.5.7 Glycosidic linkage analysis 23
2.5.8 Extraction of XyG from olive fruit 24
2.5.9 Nucleotide sugar extraction and analysis 24
2.5.10 Measurement of acetic acid content of AIR material 25
2.5.11 NMR spectroscopy of the oligosaccharide 25
2.5.12 OREX protein purification and digestion 25
2.5.13 Leaf extensibility 25

2.5.14 Immunodetection of XyG by dot blot 26
3 A forward genetic screen to identify genes involved in XyG biosynthesis and metabolism 27
3.1 Background 27
3.1.1 Efficacy of forward genetics in identifying cell wall related genes 27
3.1.2 The AXY (Altered XYloglucan) Screen 29
3.2 Results 30
3.2.1 Identification of the genetic basis of four mutants with an *axy3*-like phenotype . 32
3.2.2 Characterization of three mutants with an *axy8*-like phenotype 33
3.2.3 Identification of the genetic basis of *mur1*-like and *mur3*-like *axy* mutants 34
3.2.4 The *axy6* mutant has reduced XyG fucosylation and is distinct from *axy8* 38
3.2.5 Identification of the genetic basis of the *axy9.1* mutant 44
3.2.6 Characterization of the *axy9* mutants 51
3.3 Discussion 57
3.3.1 Identification of additional *axy3*, *axy8*, *mur1* and *mur3* alleles 57
3.3.2 Genetic basis of the *axy6* mutant 58
3.3.3 Effect of AXY9 deficiency on plant morphology and growth 58
3.3.4 Model of AXY9 involvement in plant cell wall polysaccharide acetylation 61
4 A transcriptional profiling approach to identify XyG biosynthetic genes 64
4.1 Background 64
4.2 Results 65
4.2.1 Quantification and timing of XyG deposition in developing nasturtium seeds 65
4.2.2 Transcriptional profiling of developing seeds 66
4.2.3 Selection of a candidate XyG galactosyltransferase 67
4.2.4 At5g62220 affects XyG galactosylation 73
4.2.5 Growth phenotypes 74
4.2.6 XLT2 subcellular localization 75
4.3 Discussion 76
4.3.1 Efficacy of the transcriptional profiling approach 76
4.3.2 XLT2 (At5g62220) represents a putative XyG galactosyltranferase 76

4.3.3 XyG structure and growth phenotype of the *mur3.1 xlt2* double mutant 77
5 Identification of a XyG arabinofuranosyltransferase from tomato 78
5.1 Background 78
5.2 Results 81
5.2.1 Identification of candidate XyG arabinofuranosyltransferases from tomato 81
5.2.2 Effect of expression of tomato genes on XyG structure 82
5.2.3 Structural characterization of the novel XyG oligosaccharide 85
5.2.4 Expression of XST rescues the growth phenotype of *mur3.1 xlt2* 90
5.2.5 XyG substitution pattern affects cell wall biomechanics 91
5.2.6 Susceptibility to powdery mildew 92
5.3 Discussion 93
5.3.1 SlMUR3 – a XyG galactosyltransferase from tomato 93
5.3.2 XST1 and XST2 – putative XyG arabinofuranosyltransferases from tomato 93
5.3.3 Origin of XST1 and XST2 94
5.3.4 Implications for the function of XyG substitution 96
5.3.5 Role of XyG diversity 96
6 Heterologous production of XyG 98
6.1 Introduction 98
6.1.1 Limitations of direct protein purification, forward genetic, and reverse genetic approaches 98
6.1.2 Heterologous pathway reconstruction and *in vivo* activity assays 99
6.1.3 Considerations for the reconstruction of the XyG biosynthetic pathway in *P. pastoris* 100
6.2 Results 103
6.2.1 Sequential gene insertion in *P. pastoris* 103
6.2.2 Glycosidic linkage analysis of *P. pastoris* expressing CSLC4 and XXT2 105
6.2.3 Nucleotide sugar analysis of *P. pastoris* expressing UGD and UXS 105
6.2.4 Expression of CSLC4, XXT2, UGD and UXS in *P. pastoris* 106
6.2.5 Addition of nucleotide sugar transporters 109
6.2.6 Addition of TmXXT5 109

6.2.7 Subcellular localization of XXT2 and CSLC4 in *P. pastoris* 109
6.3 Discussion 111
6.3.1 Efficacy of the mazF markerless cassette for sequential gene insertion 111
6.3.2 Production of a 4-linked glucan in *P. pastoris* 112
6.3.3 Production of UDP-xylose in *P. pastoris* 113
6.3.4 Lack of detectable XyG after coexpression of CSLC4, XXT2, UGD and UXS 113
6.3.5 Conclusion 114
7 Concluding remarks 115
8 References 117
9 Appendix 1: Gene sequences 137
10 Appendix 2: Vector maps 146
11 Appendix 3: Genotyping primer combinations 153
12 Appendix 4: Primer sequences. 157
13 Appendix 5: Miscellaneous data. 161

List of abbreviations

A	deoxyadenosine monophosphate
A. thaliana	*Arabidopsis thaliana*
A. tumefaciens	*Agrobacterium tumefaciens*
Ac	acetyl
acetyl-CoA	acetyl-coenzyme A
AIR	alcohol insolbule residue
Api	apiose
Ara*f*	arabinofuranose
Ara*p*	arabinopyranose
AXS	UDP-apiose synthase
AXY	altered xyloglucan
BLAST	basic local alignment search tool
BMGY	buffered complex media containing glycerol
BMMY	buffered complex media containing methanol
bp	base pair
C	deoxycytosine monophosphate
CAPS	cleaved ampified polymorphic sequence
cDNA	complementary deoxyribonucleic acid
CFP	cyan fluorescent protein
Col0	*Arabidopsis thaliana* cv. Columbia 0
CSLC	cellulose synthase like class C
CTAB	cetyltrimethylammonium bromide
D2O	deuterium oxide
dCAPS	derived cleaved amplified polymorphic sequence
DHB	2,5-dihydroxybenzoic acid
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleic triphsophates
DSS	3-(Trimethylsilyl)-1-propanesulfonic acid
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EI	electron impact ionization
EMS	ethyl methanesulfonate
Frc	fructose
Fuc	fucose
g	gram
g	gravity
G	deoxyguanosine monophosphate
GAE	UDP-glucuronic acid epimerase
Gal	galactose
GalA	galacturonic acid
GC-MS	gas chromatography mass spectometry
GDP	guanosine diphosphate.
GER	GDP-4-keto-6-deoxy-mannose epimerase reductase
GFP	green fluorescent protein
GH	glycosyl hydrolase
Glc	glucose
GlcA	glucuronic acid
GMD	GDP-mannose dehydratase
GME	GDP-mannose epimerase

GMP	GDP-mannose pyrophosphorylase
GT	glycosyltransferase
H_2O	water
HCl	hydrochloric acid
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
HPAEC-PAD	high-performane anion exchange chromatograph with pulsed amperometric detection
HSQC	heteronuclear single quantum coherence
IRX	irregular xylem
kb	kilobase pairs
kDa	kilodalton
KOH	potassium hydroxide
LB	lysogeny broth
LC-MS	liquid chromatography mass spectometry
Ler	*Arabidopsis thaliana* cv. Landsberg erecta
m / z	mass to charge ratio
MALDI-TOF	matrix assisted laser desorption ionization time of flight
Man	mannose
Mb	megabase pairs
MES	4-Morpholineethanesulfonic acid
min	minute
mL	milliliter
MM	minimal media containing methanol
MS	Murashige and Skoog
n	number of replicates / sample size
N. benthamiana	*Nicotiana benthamiana*
n.d.	not determined
NaOH	sodium hydroxide
nasturtium	*Tropaeolum majus*
NCBI	national center for biotechnology information
nd	not detected
NMR	nuclear magnetic resonance
NST	nucleotide sugar transporter
OD	optical density
OLIMP	oligosaccharide mass profiling
OREX	oligoxyloglucan reducing end-specific xyloglucanase
P	phosphate
P. pastoris	*Pichia pastoris*
PCR	polymerase chain reaction
PGI	phosphoglucose isomerase
PGM	phosphoglucomutase
PMI	phosphomannose isomerase
PMM	phosphomannomutase
qRT-PCR	quantitative reverse transcription polymerase chain reaction
RGP	reversibly glycosylated protein / UDP-arabinopyranomutase
Rhm	rhamnose
RHM	rhamnose synthase
RNA	ribonucleic acid
Rnase	ribonuclease
S. lycopersicum	*Solanum lycopersicum*
SNP	single nucleotide polymorphism
SSLP	simple sequence length polymorphism
SUS	sucrose synthase
T	deoxythymidine monophosphate

TAE	tris, acetic acid and EDTA
TAIR	the arabidopsis information resource
T-DNA	transfer DNA
TE	tris EDTA
TOCSY	total correlation spectroscopy
tomato	*Solanum lycopersicum*
tris	2-Amino-2-(hydroxymethyl)-1,3-propanediol
U	unit
UDP	uridine diphosphate
UGD	UDP-glucose dehydrogenase
UGE	UDP-glucose epimerase
UGP	UDP-glucose pyrophosphorylase
UXE	UDP-xylose epimerase
UXS	UDP-xylose synthase
WT	wild type
XEG	xyloglucan endoglucanase
XLT	xyloglucan L-side-chain glycosyltransferase
XST	xyloglucan S-side-chain glycosyl transferase
XTH	xyloglucan transglycosylase / hydrolase
XXT	xyloglucan xylosyl transferase
XyG	xyloglucan
Xyl	xylose
YNB	yeast nitrogen base
YPD	yeast extract, peptone and dextrose

List of figures

Figure 1-1. Model of the polymers present in the primary cell wall. 2
Figure 1-2. Cell wall polysaccharide structures. 4
Figure 1-3. Nucleotide sugar biosynthesis. 6
Figure 1-4. XyG side-chain diversity. 9
Figure 1-5. Methods for the analysis of XyG structure. 10
Figure 1-6. Genes involved in the biosynthesis of XyG. 11
Figure 3-1. XyG OLIMP spectra *axy* mutant classes. 31
Figure 3-2. *AXY3* gene model. 32
Figure 3-3. Genetic complementation of *axy1.3*. 33
Figure 3-4. *AXY8* gene model. 34
Figure 3-5. MUR3 gene model. 35
Figure 3-6. Mutant selection for mapping of *axy2.4*. 35
Figure 3-7. Genetic linkage of the *axy2.4* locus. 36
Figure 3-8. Fine mapping of *axy2.4*. 37
Figure 3-9. *MUR1* gene model. 38
Figure 3-10. XyG OLIMP of *axy6* root tissue. 39
Figure 3-11. Mutant selection for genetic mapping of *axy6*. 40
Figure 3-12. Genetic linkage analysis of the *axy6* locus. 41
Figure 3-13. Fine mapping of *axy6*. 42
Figure 3-14. Allelism test between *axy9.1* and *axy4*. 45
Figure 3-15. Selection of mutants for mapping of *axy9.1*. 46
Figure 3-16. Genetic linkage analysis of the *axy9* locus. 47
Figure 3-17. Fine mapping of *axy9.1*. 48
Figure 3-18. XyG OLIMP of *axy9* etiolated seedlings. 49
Figure 3-19. *AXY9* gene model. 50
Figure 3-20. Morphological phenotype of the *axy9.2* mutant. 52
Figure 3-21. Vascular tissue of *axy9.2*. 53
Figure 3-22. XyG OLIMP phenotypes and complementation lines of *axy9*. 54
Figure 3-23. Total acetyl-ester content of *axy9* wall material. 54
Figure 3-24. 2D HSQC NMR spectrum of *axy9* stem tissue. 55
Figure 3-25. Subcellular localization of AXY9-GFP in *N. benthamiana*. 56
Figure 3-26. Subcellular localization of AXY9-GFP in *A. thaliana*. 57
Figure 3-27. Model of plant cell wall polysaccharide acetylation. 63
Figure 4-1. Accumulation of storage XyG in developing nasturtium seeds. 65
Figure 4-2. Immunodetection of peak XyG production in developing nasturtium seeds. 66
Figure 4-3. Transcript abundance of *TmCSLC4*, *TmXXT2* and *TmMUR3*. 68
Figure 4-4. Glycosyltransferase transcipt induction. 70

Figure 4-5. Glycosyl hydrolase transcipt induction. 72
Figure 4-6. RT-PCR of an *A. thaliana* insertional mutant for At5g62220. 73
Figure 4-7. XyG structural analysis of *A. thaliana* mutants. 74
Figure 4-8. Growth phenotype of *A. thaliana* mutants. 75
Figure 4-9. Subcellular localization of XLT2. 76
Figure 5-1. XyG structures of *A. thaliana* and *S. lycopersicum*. 79
Figure 5-2. Phylogenetic tree of GT47 sequences from *A. thaliana*. 80
Figure 5-3. Identification of candidate XyG arabinofuranosyltransferases from tomato. 81
Figure 5-4. Relative quantification of tomato GT expression in transgenic *A. thaliana*. 82
Figure 5-5. XyG OLIMP of *A. thaliana* expressing tomato GTs. 83
Figure 5-6. HPAEC-PAD of XyG oligosaccharides from *A. thaliana* expressing tomato GTs. 84
Figure 5-7. Glycosidic linkage analysis of the novel XyG oligosaccharide. 86
Figure 5-8. OREX digestion of the novel oligosaccharide. 87
Figure 5-9. NMR of the oligosaccharide. 88
Figure 5-10. Comparison of the novel oligosaccharide to XXSG isolated from olive fruit. 89
Figure 5-11. Growth phenotypes of *A. thaliana* plants expressing tomato GTs. 90
Figure 5-12. Acid-induced creep of leaf petioles. 91
Figure 5-13. Wall compliance of leaf petioles of *A. thaliana* plants with altered XyG. 92
Figure 5-14. Images of leaves showing signs of powdery mildew infection. 93
Figure 5-15. Phylogenetic analysis of the XLT2 and XLT2-related clade of GT family 47. 95
Figure 6-1. Method for analysis of nucleotide sugars. 101
Figure 6-2. Schematic of XyG production in transgenic *P. pastoris*. 102
Figure 6-3. Markerless modification of *P. pastoris*. 103
Figure 6-4. Sequential gene insertion results. 104
Figure 6-5. Linkage analysis of transgenic *P. pastoris*. 105
Figure 6-6. Nucleotide sugar analysis of transgenic *P. pastoris*. 106
Figure 6-7. Glycosidic linkage analysis of transgenic *P. past*oris. 107
Figure 6-8. Nucleotide sugar analysis of transgenic *P. pastoris*. 108
Figure 6-9. Localization of CSLC4 and XXT2 within *P. pastoris*. 110
Figure 6-10. Functionality of GFP tagged CSLC4 and XXT2. 110
Figure 10-1. Vector map for *P. pastoris* constructs with the excisable markerless cassette. 146
Figure 10-2. Vector map for *P. past*oris expression plasmids with the ZeoR cassette. 147
Figure 10-3. Vector map for *P. past*oris constructs with GFP fusions and the ZeoR cassette. 148
Figure 10-4. Vector map for GFP localization in *N. benthamiana* and *A. thaliana*. 149
Figure 10-5. Vector map for plasmid used for complementation of the *xlt2* mutant. 150
Figure 10-6. Vector map for a plasmid used for overexpression of genes in *A. thaliana*. 151
Figure 10-7. Vector map for a plasmid used for plant overexpression. 152
Figure 13-1. XyG OLIMP of *axy3*-like mutants. 161

Figure 13-2. MALDI-TOF MS analysis of XEG-released material from *axy8*-like mutants. 162
Figure 13-3. XyG OLIMP of candidate mutants *196-70* and *8-64*. 163
Figure 13-4. XyG OLIMP of candidate *axy9* alleles. 164
Figure 13-5. XyG OLIMP of *axy6* leaves. 165
Figure 13-6. AXY9 protein sequence alignment. 167
Figure 13-7. XyG OLIMP of *mur3.1* and a *mur3 T-DNA* allele. 167

List of tables

Table 2-1. Growth media composition. 15
Table 2-2. Antibiotics used for selection. 15
Table 3-1. Summary of the *axy* mutants. 30
Table 3-2. Allelism test for *axy2.4* and *mur1.1*. 38
Table 3-3. Genomic sequencing of *axy6*. 43
Table 3-4. *axy6* candidate genes. 44
Table 3-5. Genomic sequencing of *axy9.1*. 48
Table 3-6. Candidate SNPs for the basis of the *axy9.1* phenotype. 49
Table 3-7. Summary of *axy* mutants characterized in this work. 50
Table 4-1. Sequencing of cDNA from developing nasturtium seeds. 67
Table 4-2. Contig assembly and read mapping of nasturtium transcriptome. 67
Table 4-3. Top ten most-induced contigs annotated with the closest *A. thaliana* homolog. 69
Table 6-1. Frequency of gene loss following transformation. 104
Table 11-1. Simple Sequence Length Polymorphism (SSLP) Markers. 153
Table 11-2. Cleaved Amplified Polymorphic Sequence (CAPS) Markers. 154
Table 11-3. Genotyping primer combinations for *A. thaliana* lines. 155
Table 11-4. Genotyping primer combinations for *P. pastoris*. 156

1 Introduction

The plant cell wall is a complex polymer composite that must fulfill a diverse set of important functions for the cell. The wall must be strong enough to provide mechanical support for the cell and the entire plant, resist osmotic pressure, and act as a barrier to pathogens and environmental stresses, while still being capable of being remodeled to allow for cell elongation (Cosgrove 2005). The diverse nature of these functions, which vary by developmental stage, tissue type, and species, is reflected in the variability of wall composition and structures observed (Burton et al. 2010; Knox 2008). Despite this variability in wall type, most plant cell walls are composed of three major types of components - glycoproteins, polyphenols, and polysaccharides. Of these, the polysaccharides are the most abundant, structurally complex, and dominant structural component, particularly for the primary cell wall (the often extensible wall produced by a plant cell during growth) (McNeil et al. 1984).

1.1 Plant cell wall polysaccharides

1.1.1 Function

Plant cell wall polysaccharides can have important structural roles in the wall, perhaps most prominently exemplified by cellulose (Somerville 2006). Polysaccharides can also have roles in energy storage (Buckeridge et al. 2000), developmental signaling (Dumville and Fry 2000; Fry et al. 1993a; Hall and Cannon 2002; Motose et al. 2004), and/or pathogen defense (Aist 1976; Hematy et al. 2009). In the primary cell wall, cellulose microfibrils are believed to be crosslinked by xyloglucan (XyG) and/or other hemicelluloses, forming a hemicellulose-cellulose network that can act as the major load bearing structure of the wall (Figure 1-1) (Hayashi 1989; Pauly et al. 1999a; Somerville et al. 2004). This network is embedded in more amorphous matrix of pectic polysaccharides. The relative proportion of these components varies, but in dicots may be approximately equal whereas as some families of monocots have relatively little pectin (Jarvis 1987). Secondary cell walls, produced in many cells after the cessation of cell elongation, also contain cellulose and hemicellulose but have less pectin, instead having the polyphenol lignin as a significant component (Vanholme et al. 2010). This provides the secondary wall with additional strength and water impermeability. In addition to these structural roles, cell wall polysaccharides can act as energy storage polysaccharides. XyG, mixed-linkage glucan, and mannan have all been demonstrated or are believed to act as storage polysaccharides in seed tissue (Buckeridge 2010; Guillon et al. 2011). Wall components, or rather their breakdown products, also serve as indicators of pathogen attack as fungal pathogens often secrete hydrolytic enzymes to degrade the wall during attempted hyphal penetration (Annis and Goodwin 1997; Huckelhoven 2007). Oligogalacturonides, derived from the breakdown of pectic homogalacturonan, have been shown to be potent signaling molecules that lead to the induction of defense responses (Bishop et al. 1981; Hahn et al. 1981; Ridley et al. 2001). XyG and mannan oligosaccharides have also been shown to have signaling properties and may have roles in plant development (Fry et al. 1993a; Mcdougall and Fry 1989; York et al. 1984; Zhao et al. 2013).

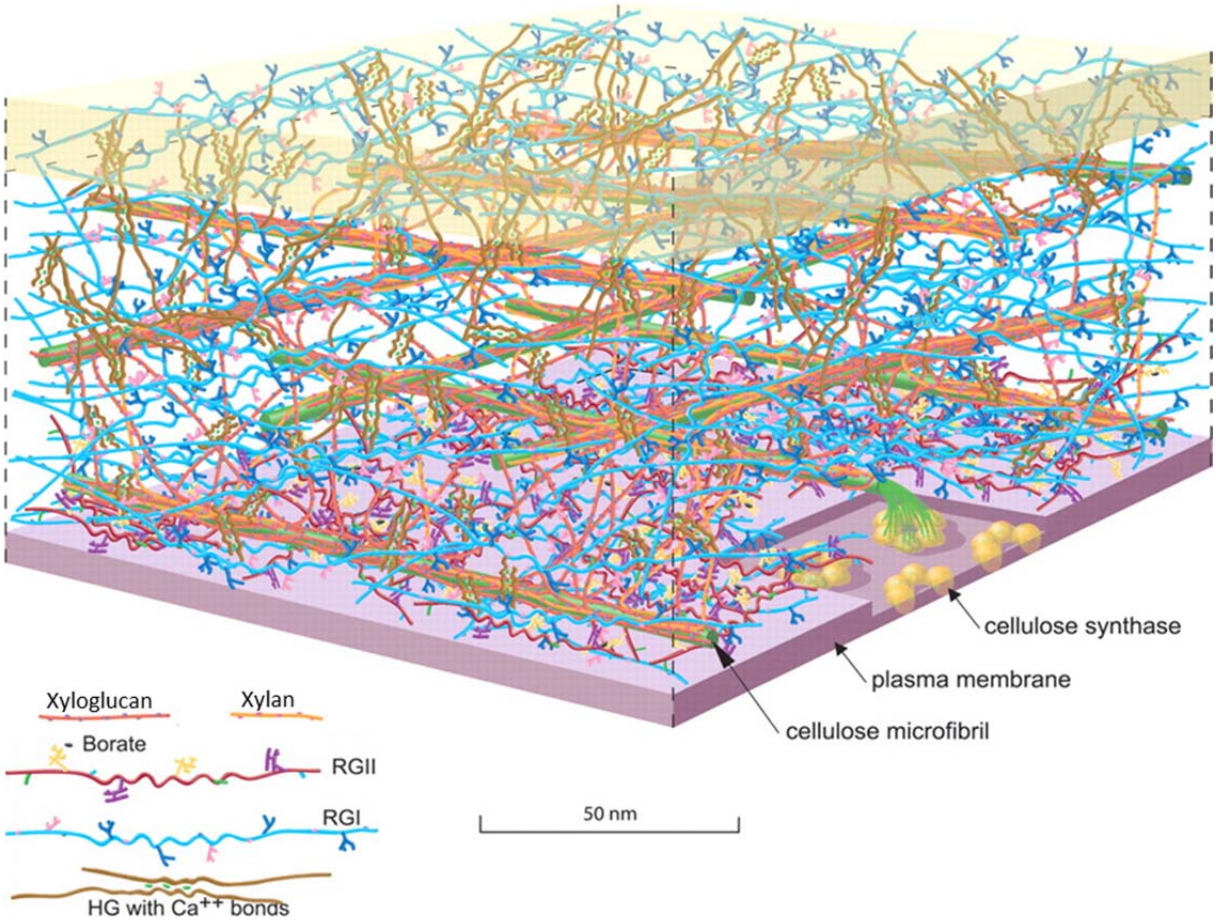


Figure 1-1. Model of the polymers present in the primary cell wall. Cellulose microfibrils are the primary load-bearing elements. The microfibrils are crosslinked by hemicelluloses such as xylan and XyG. Pectins form an amorphous matrix in which the hemicellulose and cellulose network is embedded. Rhamnogalacuronan I (RGI), rhamnogalacturonan II (RGII), and homgalacturonan (HG) are pectic polysaccharides. RGII and HG can form borate and calcium-mediated cross links respectively which contribute to the strength of the wall. Figure reprinted from Somerville et al. 2004 (Somerville et al. 2004).

1.1.2 Structures

Plant cell wall polysaccharides are divided into three main categories based on structure and chemical properties. These classes are cellulose, hemicelluloses, and pectic polysaccharides. Cellulose consists of β-1,4 glucan chains bound together by hydrogen-bonds to form semi-crystalline microfibrils. These microfibrils are a major structural component of the cell wall (Somerville 2006). The space between cellulose microfibrils, on the order of 10 – 50 nm, is filled with the other wall components to prevent the cellulose microfibrils from aggregating and thereby allowing for an extensible wall network (Figure 1-1).

Hemicelluloses are typically defined as polymers that have a β-1,4 linked glycan backbone with additional substitutions or linkages (Pauly et al. 2013; Scheller and Ulvskov 2010). The β-1,4 linked glycan backbone of hemicelluloses allows molecular interaction with cellulose microfibrils through hydrogen bonds, while the additional substitutions allow them to be partially soluble and not aggregate. Examples of hemicelluloses include heteroxylan, heteromannans, mixed-linkage glucan, and XyG (Figure 1-2). Xylan consists of a β-1,4 linked backbone of xylosyl residues that can be *O*-acetylated or substituted with glucuronosyl-, arabinosyl-, or additional xylosyl-residues depending on the plant species (Faik 2010). Xylan is the dominant hemicellulose in both the primary and secondary cell walls of grasses and in the secondary cell wall in most angiosperms. The backbone of mannan and glucomannan consists entirely of mannosyl-residues or mannosyl and glucosyl residues, respectively, and can be further modified by *O*-acetylation or galactosylation (Moreira and Filho 2008). Mannan is found throughout the plant kingdom including in some algal species, but is particularly abundant in the seeds of some species where it functions as a seed storage polysaccharide, such as guar (Kooiman 1971), or in the secondary walls of gymnosperms (Samuels et al. 2002). Mixed-linkage glucan has a backbone consisting of both β-1,3 and β-1,4 linked glucosyl residues. Unlike the other hemicelluloses, side branches or other substituents have not been observed on mixed-linkage glucan (Sorensen et al. 2008). XyG, whose structure is discussed in more detail below, consists of a β-1,4 linked glucan backbone with a regular pattern of xylosyl substituents, which can be further substituted with various moieties depending on the position, tissue, and species (Pauly et al. 2013).

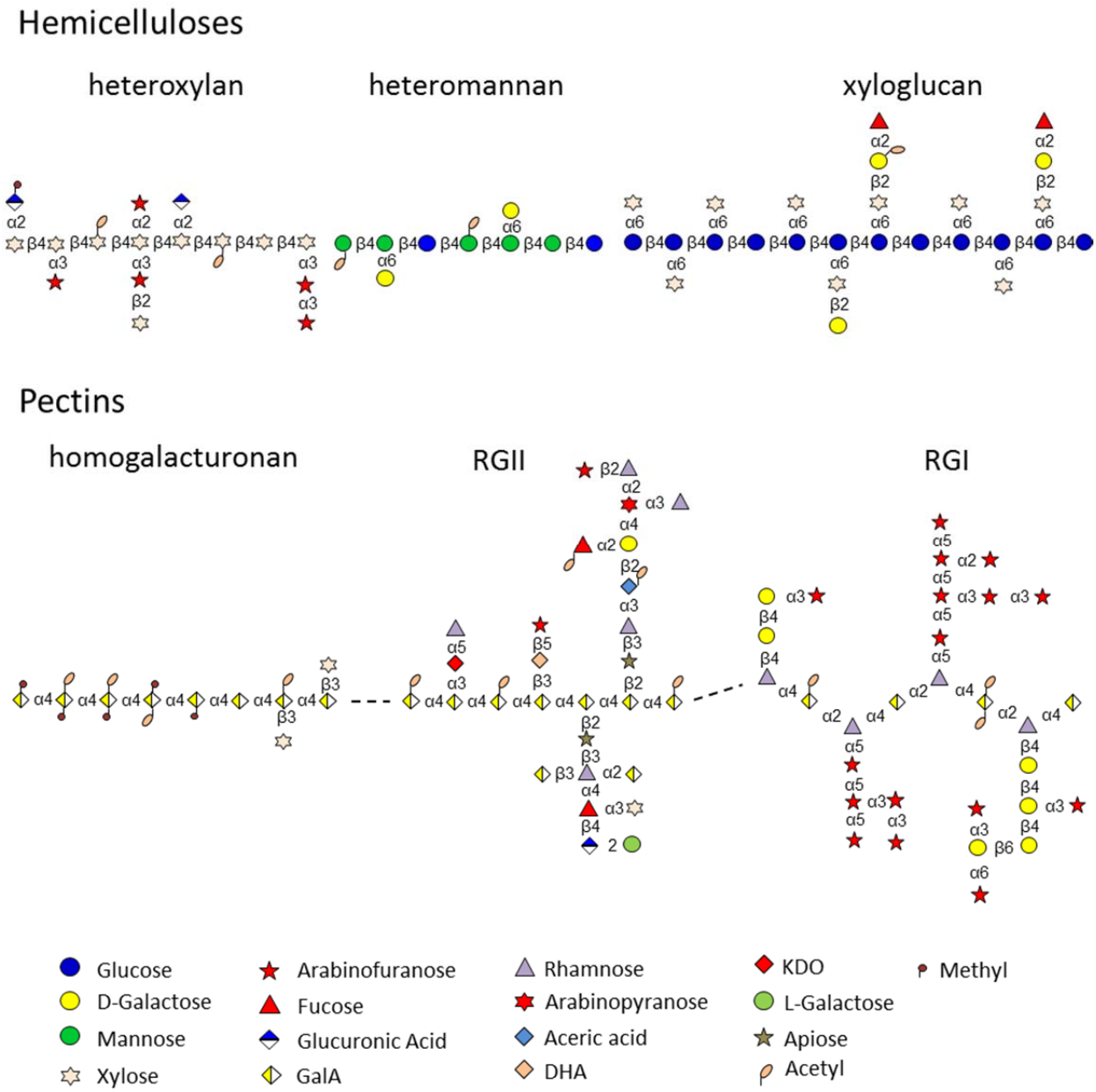


Figure 1-2. Cell wall polysaccharide structures. Representative structures of hemicellulosse (heteroxylan, heteromannan, and XyG) and pectin [homogalacturonan, rhamnogalacturonan I (RGI), rhamnogalacturonan II (RGII)] are shown. The text between the glycosyl moieties indicates the orientation of the glycosidic linkage (α or β) and the carbon number associated with the acceptor oxygen atom. The actual side-chain composition and pattern within a polymer can vary between polymer, plant tissues and plant species. The dotted lines between homogalacturonan, RGII and RGI indicate that the nature of the linkages between these various pectic fractions are not precisely known. Hemicellulose structures reprinted with modification from Pauly et al. 2013 (Pauly et al. 2013). Pectin structures adapted from Harholt et al. 2010 and Mohnen 2008 (Harholt et al. 2010; Mohnen 2008).

The pectic polysaccharides are the most structurally diverse group of polysaccharides in the plant cell wall. A total of 12 different monosaccharides have been identified from various pectins and it has been estimated that 67 distinct enzymatic activities are required for pectin biosynthesis (Mohnen 2008). The diversity of pectin structure makes it difficult to precisely define this group based on linkages or backbone composition, although pectins usually contain galacturonic acid as a major backbone component. The polysaccharides homogalacturonan (HG), rhamnogalacturonan I (RGI) and rhamnogalacturonan II (RGII) are the three major types of pectic polysaccharides, all of which are likely at least partially covalently linked to each other in the wall as enzymatic treatment with an α-1,4-endo-polygalacturonase can release them (Figure 1-2) (Harholt et al. 2010). HG consists of a backbone of α-1,4 galacturonosyl residues that can be modified with methylesters neutralizing the charged galacturonosyl-moiety. Calcium bridges can form between free galacturonosyl residues leading to the formation of gels. The cell can utilize apoplastic pectin methylesterases and pectin methylesterase inhibitors to control the degree of methylesterification, thereby affecting the ability of the polymer to form calcium bridges and altering strength of the pectin matrix (Anthon and Barrett 2010; Jarvis 1984; Muller et al. 2013; Saez-Aguayo et al. 2013). HG can be modified with xylosyl- or apiosyl-substiutents, in which case it is named xylogalacturonan or apiogalacturonan, respectively (Harholt et al. 2010). RGI has a backbone which consists of alternating rhamnose and galacturonic acid residues. Side-chains rich in galactose or arabinose can be present on the rhamnosyl residues, whereas the galacturonosyl residues can be *O*-acetylated (Mohnen 2008). Like HG, RGII has a backbone of α-1,4 linked galacturonic acid (Perez et al. 2003). The galacturonic acid residues can harbor complex side-chains that contain glycosyl units of low abundance in the wall including apiose, aceric acid, 3-deoxy-D-manno-oct-2-ulosonic acid, and 3-deoxy-D-lyxo-heptulosaric acid in addition to more common ones such as galacturonic acid, rhamnose, galactose and fucose (Figure 1-2). The numerous monosaccharide constituents and the diverse linkages in RGII make it the most structurally complex polysaccharide yet identified (Perez et al. 2003).

1.1.3 Biosynthesis

The biosynthesis of plant cell wall polysaccharides involves the activity of glycosyltransferases (GTs), which utilize activated sugars – usually nucleotide sugars – to form glycosidic bonds. UDP-glucose is an important nucleotide sugar for cellulose and XyG biosynthesis. It can be formed by the enzyme sucrose synthase, which converts sucrose to UDP-glucose and fructose (Amor et al. 1995). UDP-glucose can be acted upon by a series of enzymes, both cytosolic and Golgi-localized, to form a variety of the nucleotide sugars required for wall polysaccharide synthesis (Figure 1-3) (Bar-Peled and O'Neill 2011; Seifert 2004). Cellulose biosynthesis occurs at the plasma membrane, where cellulose synthases reside in large complexes, the rosette structures (Mueller and Brown 1980). In contrast, the matrix polysaccharides are produced in the Golgi apparatus (Lerouxel et al. 2006). Some of the enzymes, such as the cellulose synthases and the cellulose synthase-like enzymes, have multiple transmembrane domains and in most cases appear to utilize a cytosolic pool of nucleotide sugars, with the assembled polymer being extruded through the membrane to the other side. Other GTs present in the Golgi-apparatus are type II transmembrane proteins with a single N-terminal transmembrane helix (Lerouxel et al. 2006). These proteins are likely to utilize luminal pools of nucleotide sugars, which are

transported into the lumen of the Golgi by nucleotide sugar transporters (Seifert 2004). Genes required for the formation of most of the nucleotide sugars are known, but the specificities of most nucleotide sugar transporters have not been elucidated (Reyes and Orellana 2008). Wall polysaccharides produced in the Golgi are delivered to the wall by exocytosis where they are incorporated into the extracellular matrix (Sandhu et al. 2009). Once in the apoplast wall polysaccharides can be further modified by apoplastic esterases and/or glycosyl hydrolases / transglycosylases (Giovane et al. 2004; Gunl et al. 2011; Kerr and Fry 2003).

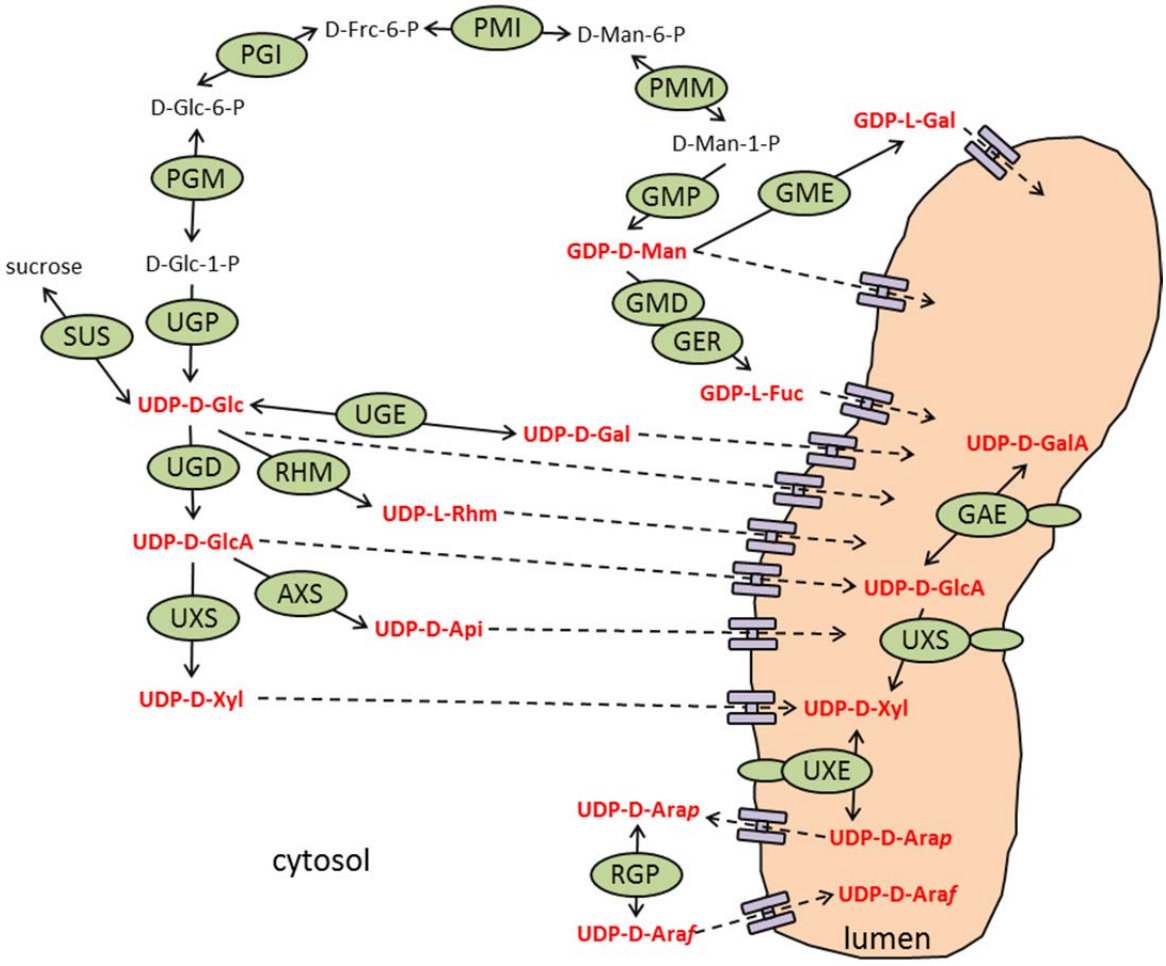


Figure 1-3. Nucleotide sugar biosynthesis. Most of the enzymes required for nucleotide sugar interconversion pathways required for plant cell wall polysaccharide biosynthesis are known. Many of the nucleotide sugars are synthesized in the cytosol and imported into the endomembrane system by nucleotide sugar transporters. Adapted from Seifert 2004 (Seifert

2004). Abbreviations: Api – apiose, Araf – arabinofuranose, Arap – arabinopyranose, AXS – UDP-apiose synthase, Frc – fructose, Fuc – fucose, GAE – UDP-glucuronic acid epimerase, Gal – galactose, GalA – galacturonic acid, GDP – guanosine diphosphate, GER – GDP-4-keto-6-deoxy-mannose epimerase reductase, Glc – glucose, GlcA – glucuronic acid, GMD – GDP-mannose dehydratase, GME – GDP-mannose epimerase, GMP – GDP-mannose pyrophosphorylase, Man – mannose, P – phosphate, PGI – phosphoglucose isomerase, PGM – phosphoglucomutase, PMI – phosphomannose isomerase, PMM – phosphomannomutase, RGP – reversibly glycosylated protein / UDP-arabinopyranomutase, Rhm – rhamnose, RHM – rhamnose synthase, SUS – sucrose synthase, UDP – uradine diphosphate, UGD – UDP-glucose dehydrogenase, UGE – UDP-glucose epimerase, UGP – UDP-glucose pyrophosphorylase, UXE – UDP-xylose epimerase, UXS – UDP-xylose synthase, Xyl – xylose.

1.2 XyG

1.2.1 Function

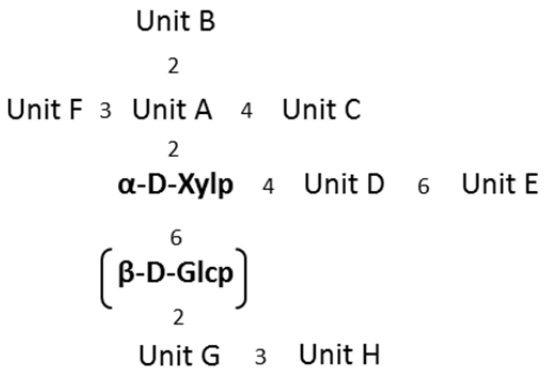
XyG has been shown to crosslink cellulose microfibrils and in some models of the primary wall the cellulose/XyG network is proposed to be the major load-bearing structure (Carpita and Gibeaut 1993; Hayashi 1989; Pauly et al. 1999a). In these models, the XyG crosslinks can be remodeled by XTH's to incorporate new XyG polysaccharides or weaken the network, allowing for control cell elongation (Takeda et al. 2002). The XyG polysaccharides that span the distance between cellulose microfibrils are accessible to enzymatic remodeling, whereas additional XyG can be closely associated with or entrapped within the microfibrils (Pauly et al. 1999a). Expansins are believed to disrupt interactions between XyG and cellulose, also allowing for alterations to this cellulose/XyG network (Cosgrove 1998). These models would predict that a plant lacking XyG would have drastically altered cell wall architecture and likely a severe growth phenotype. However, most *A. thaliana* mutants with altered XyG structure have relatively minor growth phenotypes. This includes an *xxt1 xxt2* double mutant which lacks detectable XyG and has approximately normal growth, with the exception of having aberrant root hairs and being slightly smaller than wild type (Cavalier et al. 2008; Park and Cosgrove 2012a). This suggests that XyG may not be as critical for cell wall integrity and elongation in *A. thaliana* as previously thought and a revised model has been proposed in which XyG instead functions at the interface of adjacent cellulose microfibrils (Park and Cosgrove 2012b).

In addition to having a structural role, XyG is known to act as a seed storage polymer in several species including nasturtium (*Tropaealum majus*) (Edwards 1985), tamarind (White and Rao, Constitution of the polysaccharide from tamarind seed, 1953), and *Hymenaea courbaril* (Buckeridge 1990). Seed storage XyG typically lacks fucosylation and acetylation common in structural XyG found in the wall (Gidley et al. 1991). The reason why storage XyG lacks these substituents is not known, but it has been proposed that lacking fucosylation can prevent a potential anti-auxin activity of released oligosaccharides (Buckeridge et al. 2000; Mcdougall and Fry 1989) or reduce XyG-cellulose interactions (Levy et al. 1997). Alternatively, the lack of these modifications on seed storage XyG may allow for the plant to selectively degrade storage XyG without affecting the structural XyG also presence in the wall (Marcus et al. 2008) or may simplify the suite of enzymes required for the hydrolysis and metabolism the polysaccharide.

XyG oligosaccharides have been implicated as signaling molecules as fucosylated oligosaccharides have been shown to suppress the elongation effect of auxin in pea stem (Mcdougall and Fry 1989; York et al. 1984). These XyG oligosaccharins, a term applied to oligosaccharides with signaling activity (Fry et al. 1993a), have effects at nM concentrations suggesting a receptor-mediated mechanism of action though such a receptor has not been reported. Large concentrations (1 mM) of XyG oligosaccharides actually have a growth-promoting effect, possibly due to a weakening of the XyG matrix due to the incorporation of these oligosaccharides by the action of XTHs and subsequent decrease in the XyG degree of polymerization or by activation of endogenous hydrolytic enzymes (Fry et al. 1992; Mcdougall and Fry 1990). As XyG oligosaccharides can be produced *in vivo* by the action of XTHs promoting cell elongation, it has been proposed that a subset of these can feedback to act as a form of negative regulation to keep cell growth in balance (Fry et al. 1993a). *A. thaliana* mutants having decreased (*mur2*) or increased (*axy8*) XyG fucosylation do not show strong growth phenotypes (Gunl et al. 2011; Vanzin et al. 2002), questioning the physiological relevance of fucosylated XyG oligosaccharins as signaling molecules or suggesting that redundant regulatory mechanisms exist.

1.2.2 Structures and diversity

XyG is a β-1,4 glucan with a regular pattern of xylosyl-substituents. The xylosyl units are often further substituted with various glycosyl groups. A nomenclature has been developed to describe the substitution pattern of each backbone glucosyl residue (Figure 1-4) (Fry et al. 1993b). An unsubstituted glucosyl residue is represented by a G, whereas a glucosyl substituted at *O*-2 with a xylosyl residue is known as an X. The xylosyl residue can be further substituted with a galactosyl-moiety (L) or an arabinofuranosyl-residue (S). The galactosyl residue of an L side-chain can be additionally fucosylated (F) or *O*-acetylated (underlined <u>L</u>). In total, more than 18 different side-chains have been discovered on XyG in various plant species and tissues (Figure 1-4).



Abbrev.	Substituent Position								Occurrence	Reference
	A	B	C	D	E	F	G	H		
X										
L	β-D-Gal*p*									
F	β-D-Gal*p*	α-L-Fucp								
S	α-L-Ara*f*								Solanales, Lamiales, Gentianales	Vierhuis 2001, Hoffman 2005
D	α-L-Ara*p*								*S. kraussiana*	Pena 2008
P	β-D-GalpA			β-D-Gal*p*					*P. patens*	Pena 2008
Q	β-D-GalpA		β-D-Galp	β-D-Gal*p*					*P. patens*	Pena 2008
E	α-L-Ara*p*	α-L-Fuc*p*							*E. hyemale*	Pena 2008
M	α-L-Ara*p*			β-D-Gal*p*					*P. patens*	Pena 2008
N	α-L-Ara*p*			β-D-Gal*p*	β-D-Gal*p*				*P. patens*	Pena 2008
T	α-L-Ara*f*					β-L-Araf			Solanales	Jia 2003
Y	β-D-GalpA								*A. thaliana* root hair	Pena 2012
J	β-D-Gal*p*	α-L-Gal*p*							Jojoba	Hantus 1997
U	β-D-Xyl*p*								*A. spinosa*	Ray 2004
Z	β-D-GalpA	α-L-Fuc*p*							*A. thaliana* root hair	Pena 2012
A							α-L-Ara*f*			Kiefer 1990, Hisamatsu 1992
B							β-D-Xyl*p*			Hisamatsu 1992
C							β-D-Xyl*p*	α-L-Ara*f*		Hisamatsu 1992

Figure 1-4. XyG side-chain diversity. These structures and the corresponding nomenclature are listed including occurrence and reference. An unsubstituted glucosyl residue part of the glucan backbone is depicted as a G (omitted from this table). The core motif of substituted glucosyl-residues consists of a backbone glucosyl-residue (bold in bracket) with a xylosyl-residue attached to it (bold). The numbers between the glycosyl residues indicate the carbon number of the acceptor oxygen atom.

The study of XyG structure is aided by the availability of enzymes that hydrolyze XyG to release XyG oligosaccharides from cell wall material. Non-specific endoglucanases (Bauer et al. 1973) or a XyG-specific endoglucanase (XEG) that specifically hydrolyze the XyG glucan backbone at unsubstituted glucosyl residues where the aglycone contains substituted glucosyl-residues (Pauly et al. 1999b) can be utilized for this purpose. The solubilized XyG oligosaccharides can be

analyzed by mass spectrometry or HPLC to determine their identity and abundance to better understand the structure and substitution pattern of the polysaccharide (Figure 1-5) (Lerouxel et al. 2002). Digestion of wall material from *A. thaliana* results predominately in the oligosaccharides XXXG, XLFG, XLFG, XXFG, XXFG, XXLG, XXLG and XLXG. These structures contain a core XXXG-type motif that is typical for the XyG in many dicot species (Hoffman et al. 2005). In contrast, some species have XyG that is less xylosated, such as plant species from the Solanales (XXGG-type) and the grasses (York et al. 1996), as well as the moss *Physcomitrella patens* ($XXGG_n$-type) (Kato et al. 2004; Pena et al. 2008), or more xylosated such as *Hymenaea courbaril* (XXXXG-type) and *Helipterum eximium* (fully xylosated) (Mabusela et al. 1990).

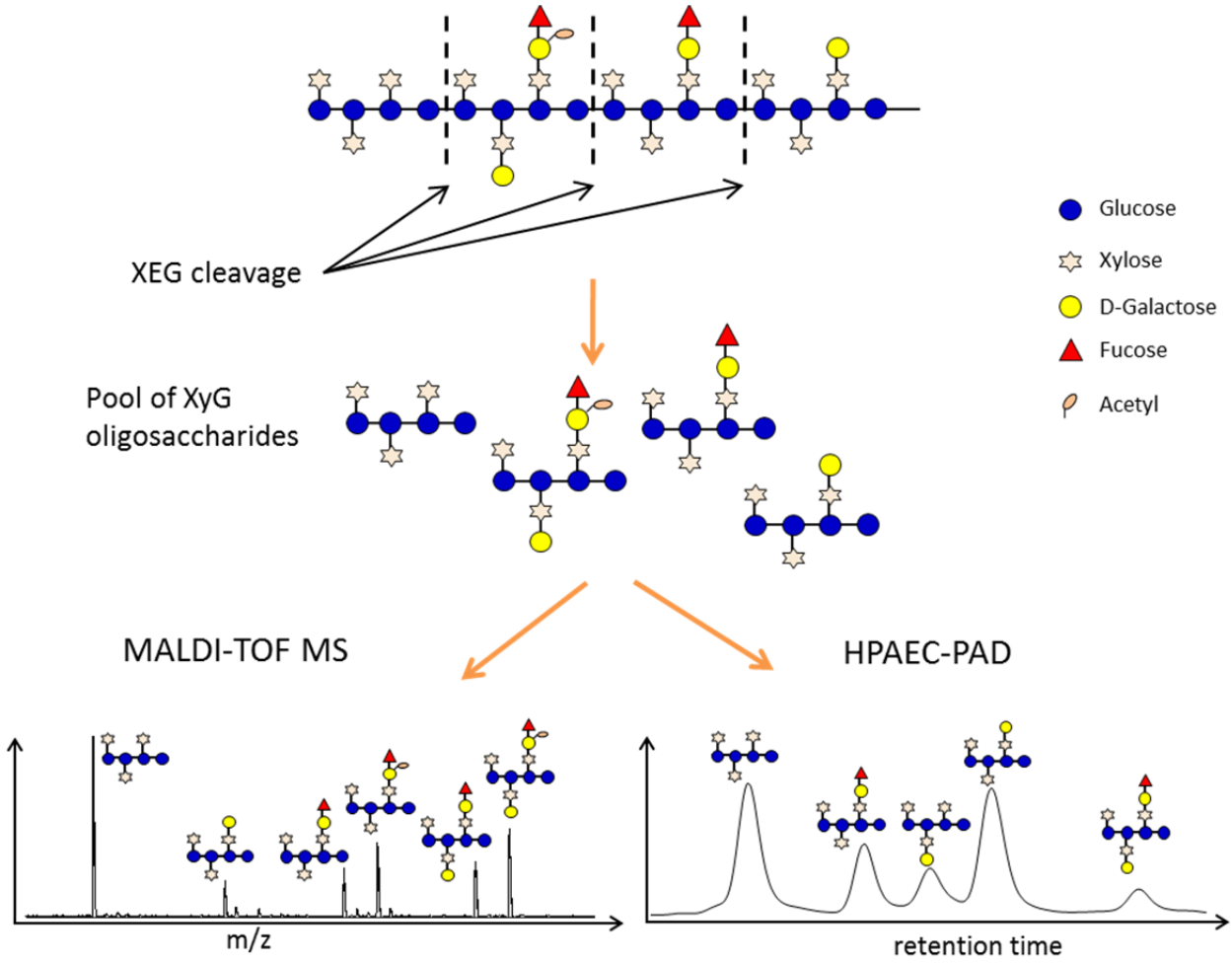


Figure 1-5. Methods for the analysis of XyG structure. To analyze XyG structure wall material can digested with a XyG specific endoglucanase (XEG). The resulting pool of solubilized oligosaccharides can then be detected by matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS) in a method known as oligosaccharide mass profiling (OLIMP) or separated by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) (Lerouxel et al. 2002).

1.2.3 XyG Biosynthesis

Numerous genes are known to be involved in the biosynthesis of XyG (Gunl et al. 2011; Gunl and Pauly 2011; Sampedro et al. 2012; Zabotina 2012). These genes include GTs required for the addition of the various glycosyl-moieties (Figure 1-6) as well as glycosyl hydrolases (GHs), which are required for maturation of the polymer and incorporation into the extracellular matrix.

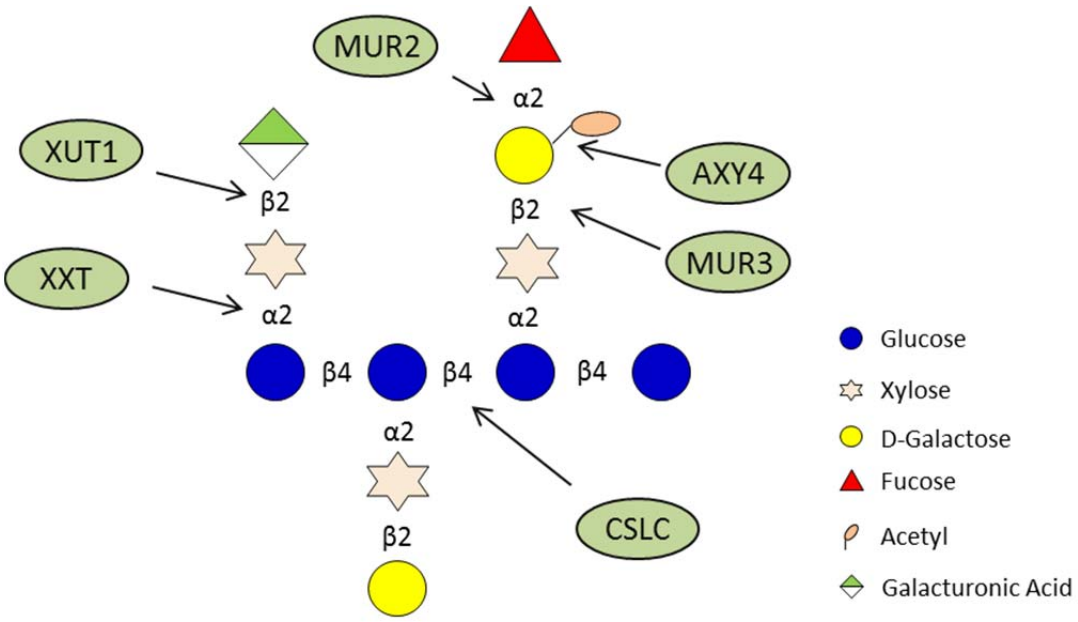


Figure 1-6. Genes involved in the biosynthesis of XyG. The fully substituted XyG motif XLFG is shown with the genes known to be required for its biosynthesis. CSLC and XXT represented classes of genes involved in glucan backbone synthesis and xylosyl transfer, respectively. The MUR3 enzyme is responsible for galactose transfer specifically to the third xylosyl unit and XUT1 can transfer a galacturonosyl to the first xylosyl unit. The galactosyltransferase required for galactosylation at the second position had not been identified. AXY4 is believed to be an acetyltransferase. This figure represents linkage and the corresponding known genes from A. thaliana, but many other linkages have been reported from other species (Figure 1-4).

A XyG glucan synthase termed CSLC4 was identified using a transcriptomics approach on developing seeds of *Tropaeolum majus* (nasturtium), which produce XyG as a seed storage polymer (Cocuron et al. 2007). *A. thaliana* plants defective for CSLC4 still produce XyG, indicating that one or several of the other CSLC genes may also encode redundant XyG glucan synthases. The XyG xylosyltransferases were initially discovered by homology to a fenugreek (*Trigonella foenum-graecum*) α-1,6 galactosyltransferase from GT family 34 that acts on mannan (Faik et al. 2002). Five out of seven *A. thaliana* proteins in GT family 34 have been demonstrated by *in* vitro activity or genetic complementation of mutants to function as XyG xylosyltransferases (Faik et al. 2002; Vuttipongchaikij et al. 2012; Zabotina et al. 2008), even

though a knockout of XXT1 and XXT2 is sufficient to result in an *A. thaliana* plant that lacks detectable XyG (Cavalier et al. 2008).

Two genes from GT family 47, MUR3 and XUT1, have been shown to be responsible for the addition of glycosyl units to the *O*-2 of the XyG xylosyl residues (Madson et al. 2003; Pena et al. 2012). MUR3 was first identified in a forward genetic screen for *A. thaliana* mutants with altered cell wall monosaccharide composition. Despite being a XyG galactosyltransferase, the *mur3* mutant was initially identified by having a lack of fucose due to the absence of the galactose required for XyG fucosylation (Figure 1-6). MUR3 was heterologously expressed and purified for *in vitro* activity assays. These assays indicated that MUR3 is a galactosyltransferase that utilizes UDP-galactose and acts at the third xylose of the XXXG motif of the XyG polymer. XUT1 was identified from *A. thaliana* as a XyG galacturonsyltransferase (Pena et al. 2012). *A. thaliana* was not previously known to contain galacturonic acid as a substituent on XyG, but the authors were able to identify it by specifically isolating negatively charged XyG oligosaccharides released from root walls by an XEG digest.

The XyG fucosyltransferase (Fut1) gene was first identified by purification of the corresponding protein from pea (Perrin et al. 1999). Partial peptide sequences from the purified protein led to the identification of the Fut1 gene. In parallel, the gene was also identified in a forward genetic screen for mutants with altered cell wall monosaccharide content and is therefore also known as MUR2 (Vanzin et al. 2002). In addition to fucosylating the galactosyl residue at the third position of the XXXG motif, MUR2 can also fucosylate galacturonic acid present at the first position of the motif (Pena et al. 2012).

While some plants contain acetyl substituents on the XyG backbone, in *A. thaliana* acetyl groups have only been identified on the galactosyl residue at the third position of the XXXG motif (Kiefer et al. 1989; Maruyama et al. 1996; Sims et al. 1996; York et al. 1995; York et al. 1988). A screen for *A. thaliana* plants with altered XyG resulted in the identification of a putative XyG acetyltransferase gene, named *AXY4*, which is required for acetylation of XyG in the hypocotyls, leaves, stems and roots of *A. thaliana* (Gille et al. 2011b). Interestingly, a separate gene, AXY4L, is responsible for acetylation of XyG in seeds. AXY4 and AXY4L are both members of the Trichome Birefringence Like (TBL) protein family, which has 46 members in *A. thaliana* (Bischoff et al. 2010). While *in vitro* activity of AXY4 and AXY4L has not been demonstrated, their discovery suggested that they are XyG acetyltransferases and that other polysaccharide acetyltransferase may be present within this TBL-family. Indeed, a TBL believed to be a xylan acetyltransferase has been discovered among them (Xiong et al. 2013; Yuan et al. 2013).

In the work implicating CSLC4 as the XyG glucan synthase, expression of CSLC4 only resulted in the production of a glucan when coexpressed with the XyG xylosyltransferases XXT2 (Cocuron et al. 2007). This suggests the formation of a complex between CSLC4 and XXT2 required for XyG biosynthesis. Bimolecular fluorescence complementation and coimmunoprecipitation experiments found evidence for the existence of CSLC4-CSCL4 and XXT2-XXT2 homodimers in addition to XXT5-CSLC4 and various XXT heterodimers (Chou et al. 2012). The existence of protein complexes in XyG biosynthesis parallels complexes involved in xylan (Zeng et al. 2010),

pectin (Atmodjo et al. 2011; Harholt et al. 2012), and cellulose biosynthesis (Desprez et al. 2007; Gu et al. 2010; Guerriero et al. 2010). Complexes may serve as a way to regulate or increase GT activity/specificity, sequester substrates to promote more regular product formation, or control protein localization (Atmodjo et al. 2011; El-Battari et al. 2003; Harholt et al. 2012; Nguema-Ona et al. 2006).

1.2.4 XyG Metabolism

XyG Transglycosylase / Hydrolase (XTH) enzymes can either hydrolyze or hydrolyze and reform the backbone of XyG (Rose et al. 2002). These enzymes are important for the biosynthesis of XyG, incorporation of the polysaccharide into the wall matrix, remodeling of the XyG network, and mobilization of seed storage XyG in select species (Buckeridge et al. 2000; Edwards et al. 1986; Fanutti et al. 1993; Farkas et al. 1992; Herbers et al. 2001; Liu et al. 2007; Nishikubo et al. 2011; Nishitani and Tominaga 1992; Smith and Fry 1991). XyG can form crosslinks between cellulose microfibrils forming a network believed to be important for the strength of the cell wall (1.2.1) (Hayashi 1989). By breaking and reforming these XyG crosslinks, XTHs are thought to be able to remodel this network and allow for the wall to extend in a controlled fashion during cell elongation (Fry 1989; Hayashi and Kaida 2011). Plants that use XyG as a seed storage polysaccharide must be able to rapidly degrade and utilize the extracellular seed storage XyG. Some XTHs, specifically those with high hydrolytic activity, have been shown to be highly expressed during seed germination mobilizing the polymer into energy production (Desilva et al. 1993; Edwards et al. 1986; Fanutti et al. 1993).

Glycosidases play an important role in XyG formation as well as in the degradation of seed storage XyG during seed germination. The xylosidase XYL1/AXY3 (from GH family 31) was first discovered by purification from cabbage is active on XyG oligosaccharides (Sampedro et al. 2001). The *XYL1* gene was also identified in the AXY screen for *A. thaliana* mutants with altered XyG (Gunl and Pauly 2011). The glycosyl hydrolases AXY8 and BGal10 are also active on XyG oligosaccharides, resulting in the removal of the fucosyl (Gunl et al. 2011) and galactosyl residues (Sampedro et al. 2012), respectively. While only active on XyG oligosaccharides, the oligosaccharides can be reincorporated into the XyG polymer suggesting that their activity can have significant effects on overall substitution pattern and therefore could modulate polymer solubility and interactions with other wall components (Gunl et al. 2011; Smith and Fry 1991). That both the axy3 and axy8 mutants have major changes in XyG subunit composition suggests that there is a significant amount of XyG turnover in the wall (Gunl et al. 2011; Gunl and Pauly 2011). By degrading XyG oligosaccharides, which can be released by the action of XTHs (Farkas et al. 1992), the glycosyl hydrolases allow the average degree of polymerization of the polymer to increase in the wall (Liu et al. 2007). Together, the XTHs and glycosidases involved in XyG metabolism provide a way for the cell to significantly remodel the XyG polysaccharides once deposited in the wall.

1.2.5 Applications

XyG is used in a variety of applications for food, textiles, industrial and pharmaceutical purposes (Mishra and Malhotra 2009). In the food area, XyG can act as a thickening agent and falls into the category of a soluble dietary fiber, which can have beneficial effects on health such as

lowering blood cholesterol, promoting gastrointestinal health, and increasing satiety to reduce excess food consumption (Marlett et al. 2002). XyG can also be used to functionalize cellulose surfaces, as the soluble nature of XyG, or XyG oligosaccharides that can be reincorporated into polymeric XyG by the action of XTHs, allows for easy chemical modification while retaining the ability of XyG to bind tightly to cellulose (Brumer et al. 2004; Fink et al. 2011). This approach has been used to coat bacterial cellulose, a promising material for artificial blood vessels, with a cell adhesion peptide to allow epithelial cells to better adhere (Fink et al. 2011). XyG and related polymers can be used to form membranes for food packaging, drug delivery, cosmetics, and a variety of applications (Simi and Abraham 2010; Simkovic 2013). As the properties of biopolymer composites vary based on the structure of the components used (Roubroeks and Kondo 2012; Simkovic 2013), tailoring the structure of XyG for particular applications may be important to obtain the desired properties.

1.3 Thesis objective

While significant strides have been made in the last fifteen years to identify the genes involved in biosynthesis of cell wall polysaccharides, the majority of cell wall polysaccharide linkages have not been attributed to specific genes (Keegstra 2010). XyG, as the dominant hemicellulose in the primary walls of dicots and containing a variety of linkages in various plant species (1.2.2), makes for a relevant and interesting polysaccharide to study. Several of the genes involved in the biosynthesis of this polysaccharide have been identified in the model species *A. thaliana* (Figure 1-6, 1.2.3), but as the known genes do not comprise a full set of enzymes for all the linkages it is likely that additional ones are required. Moreover, a diverse array of XyG substitution patterns have been found in other species that are not present in *A. thaliana* (Figure 1-4), but the genetic basis for and functional significance of this variation has not been explored.

The objective of this work is to identify and characterize novel genes involved in XyG biosynthesis as well as to gain insight into the function of specific XyG substitution patterns. The following specific questions will be addressed:

- Which additional genes are involved in XyG biosynthesis?
- Which are some of the genes responsible for XyG structural variation found between species?
- What is the function of XyG substitution? What effect does side-chain diversity have on polysaccharide function?
- Are the known XyG biosynthetic genes sufficient for XyG biosynthesis in a heterologous host?
- Can such a host be used as a platform for the identification of XyG biosynthetic genes and to gain insight into how the biosynthetic proteins are regulated?

2 Materials and methods

2.1 Media and growth conditions

2.1.1 Growth media

Media for growth of *P. pastoris* was prepared according to the directions in the EasySelect *Pichia* Expression Kit manual (Invitrogen). Media was prepared by autoclaving or filter sterilizing the various components and cooling to 55 °C prior to mixing and addition of appropriate antibiotics from stock solutions (Table 2-2). The contents of the four media types are listed in Table 2-1. Solid media was prepared with 1.5% (w/v) agar for bacteria and yeast or 0.8% (w/v) agar for *A. thaliana*.

Media	Contents
LB	1% sodium chloride, 1% tryptone, 0.5% yeast extract, pH 7.0
YPD	1% yeast extract, 2% peptone, 2% glucose[1]
BMMY	1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0[1], 1.34% YNB[2], 4 x 10^{-5}% biotin[2], 0.5% methanol[2]
BMGY	1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6.0[1], 1.34% YNB[2], 4 x 10^{-5}% biotin[2], 1% glycerol[1]
MM	1.34% YNB[2], 4 x 10^{-5}% biotin[2], 0.5% methanol[2]
½ MS	0.215% MS salts (Sigma), 0.06% MES, 1% sucrose pH 5.6 with KOH

Table 2-1. Growth media composition. LB was used for growth of *E. coli* and *A. tumefaciens*. *P. pastoris* was grown on YPD, BMMY, BMGY and MM media. ½ MS was used for growth of *A. thaliana*. YNB – yeast nitrogen base with ammonium sulfate without amino acids. MES – morpholine-4-ethanesulfonic acid. [1] – stock solution autoclaved separately from other components. [2] – stock solution prepared by filter sterilization rather than autoclaving.

Antibiotic	Concentrations
Kanamycin	60 µg / mL (*E. coli*, *A. tumefaciens*, *A. thaliana*)
Hygromycin	25 µg / mL (*E. coli*, *A. thaliana*)
Rifampicin	10 µg / mL (*A. tumefaciens*)
Chloramphenicol	25 µg / mL (*E. coli*)
Carbenicillin	100 µg / mL (*E. coli*, *A. tumefaciens*)
Gentamycin	25 µg / mL (*A. tumefaciens*)
Spectinomycin	100 µg / mL (*E. coli*, *A. tumefaciens*)
Zeocin	100 µg / mL (*P. pastoris*)

Table 2-2. Antibiotics used for selection. The antibiotics used for selection of transformants in the various systems are listed with the concentrations used.

2.1.2 Induction of transgene expression in *Pichia pastoris*

P. pastoris cultures were streaked from glycerol stocks onto YPD plates and placed at 30 °C for three days. BMGY cultures (1 mL) were inoculated from single colonies and placed on a shaker at 30 °C overnight. The cultures were then centrifuged at 1500 *g* for 3 min, the supernatant was discarded, and the pellets were resuspended in BMMY (5 mL). The cultures were placed on a shaker at 30 °C for three days with daily additions of 25 µL methanol to continue induction.

2.1.3 Sterilization of *A. thaliana* seed surface

For sterilization of the seed surface, *A. thaliana* seeds were incubated with 1 mL 70% ethanol with 0.05% Triton X-100 for 10 minutes. The ethanol was removed and the seeds were washed twice with 95% ethanol for five minutes each time. The seeds were then pipetted onto ethanol soaked filter paper, allowed to dry, and distributed onto MS plates for growth.

2.1.4 Growth of *A. thaliana* on plates and soil

Following seed sterilization and plating, the plates with *A. thaliana* seeds were moved to 4 °C for two days for stratification. For light-grown plants, the plates were then moved to the growth chamber with 16 hours light and 8 hours dark at 22 °C. For dark-grown plants, the plates were exposed to light in the growth chamber for six hours then wrapped in foil and left at 22 °C for one week. For growth on soil, seeds were first suspended in a sterile 0.15% agar solution for two days at 4 °C. The seeds were then put on water-saturated soil and placed in the growth chamber with 16 hours light, 8 hours dark at 22 °C. The plants were watered as needed and fertilized every two weeks with Miracle Grow All-Purpose Plant Food (Scotts) at the recommended concentration.

2.1.5 Growth of *N. benthamiana* on soil

Nicotiana benthamiana plants were grown on soil with 16 hours light at 26 °C and 8 hours dark at 22 °C with 50% humidity. One week old plants were fertilized with Miracle Grow (Scotts) following the manufacturer's instructions.

2.2 Molecular biology

2.2.1 Extraction of genomic DNA from plant tissue

Leaf material (10-30 mg fresh weight) was placed in a 1.7 mL tube (Posi-Click, Denville), frozen in liquid nitrogen and ground with two glass beads (3 mm) in a Retsch-Mill at 25 Hz for 1 min. The powdered material was then incubated with 300 µL of 2X CTAB buffer (2% w/v CTAB, 1.4 M NaCl, 100 mM Tris HCl pH 8, 20 mM EDTA) at 65 °C for 10 minutes. Chloroform (300 µL) was added and the samples were vortexed prior to centrifugation for 5 minutes at 21,000 g. The aqueous phase was transferred to a new tube, mixed with 300 µL isopropanol, and centrifuged at 21,000 *g*. The supernatant was discarded and the pellet was washed once with 70% ethanol (500 µL). The pellet was dried and dissolved in 100 µL TE buffer.

2.2.2 Genotyping *A. thaliana* and *Pichia pastoris* using PCR

The genotyping PCR reactions were performed using JumpStart Red Taq ReadyMix (Sigma) in 20 µL reactions using the manufacturer's recommendations for reaction recipe and thermo cycling conditions. Genomic DNA (5 µL of a 5X dilution) (2.2.1) or a dilute mixture of cells (5 µL) were used for genotyping of *A. thaliana* and *P. pastoris* respectively in each PCR reaction. The PCR

product was visualized by DNA gel electrophoresis on a 1% TAE agarose gel. PCR primers for genotyping T-DNA lines were designed using the T-DNA express tool (http://signal.salk.edu/cgi-bin/tdnaexpress).

2.2.3 Genetic mapping with SSLP and CAPS markers

Simple sequence length polymorphism (SSLP) marker primer sequences were obtained from The Arabidopsis Information Resource (TAIR), with selection and verification performed previously by Sascha Gille, a member of the Pauly Lab (Table 11-1). PCR for amplification of the polymorphic sequences was performed as described above for the genotyping PCR (2.2.2), except with 45 cycles of amplification rather than 35. The PCR products were visualized with DNA gel electrophoresis on 4% agarose TAE gels due to the small differences is length of the polymorphic sequences (usually 10 to 100 bp).

The identification of cleaved amplified polymorphic sequence (CAPS) markers was aided by the use of a Python script. This script used a list of SNPs between *A. thaliana* ecotypes Col0 and Ler and, by utilizing Primer3 (Rozen and Skaletsky 1999) and BLAST (Altschul et al. 1997), generated primer pairs likely to be specific and useful CAPS markers (Konieczny and Ausubel 1993). PCR with these markers was performed as described for the genotyping PCR above followed by addition of 2.1 µL water, 2.5 µL 10X EcoRI buffer, and 0.4 µL EcoRI (New England Biolabs) to each reaction for restriction digestion overnight at 37 °C. The digested products were visualized using DNA gel electrophoresis. The CAPS markers used for fine mapping are listed in Table 11-2.

2.2.4 Amplification of genes for cloning using PCR

Phusion DNA Polymerase (Finnzymes) was used for cloning DNA sequences for plasmid construction. The manufacturer's recommendations were followed initially, with optimization of the reaction recipe by addition of DMSO and alteration of annealing temperature as required.

2.2.5 Cloning methods

Several cloning methods, sometimes in combination, were used for plasmid construction include Gateway, Restriction/Ligation, and Topo/TA. Intermediate and final constructs were verified by analytical restriction digest and Sanger Sequencing.

2.2.5.1 Gateway Cloning

Gateway Cloning (Invitrogen) was performed by first amplifying the gene of interest with attB sites incorporated at the 5' end of the gene-specific primers. The PCR product was extracted from the gel after DNA gel electrophoresis using the Qiagen Gel Purification kit (Qiagen). The purified product was used in a BP reaction with 100 ng of the appropriate DONR plasmid and transformed into TopTen Oneshot chemically competent *E. coli* cells (Invitrogen). The correct entry plasmids were selected, confirmed by restriction digest analysis and Sanger sequencing, and used to move the desired gene fragment(s) into the appropriate destination vector via an LR reaction. The final constructs were again verified by restriction digest and Sanger sequencing of the insert.

2.2.5.2 Restriction/Ligation cloning

The plasmids (3 µg) containing the desired DNA fragment and the plasmid serving as the acceptor for the DNA fragment were digested with the appropriate DNA restriction enzymes in a 50 µL reaction. The digest products were purified by gel electrophoresis or PCR purification using the corresponding Qiagen purification protocols. The purified products, approximately 100 ng, were then combined with 2 µL NEB ligation buffer and water to 19 µL. The solution was incubated at 65 °C for 5 min and cooled on ice. DNA ligase (1 µL) was added to the ligation reaction and it was incubated overnight at 16 °C. The ligation reaction was then used to transform chemically competent *E. coli*.

2.2.5.3 Topo/TA cloning

The construction of several constructs was aided by the use of TOPO/TA cloning of genes or gene fragments. The desired DNA sequence was first amplified by PCR using Phusion DNA Polymerase. The product was purified using a PCR Purification (Qiagen) and A' overhangs were added by incubation of the purified product with Taq Polymerase in appropriate buffer with dNTPs at 72 °C for 10 minutes. The resulting product was then used directly for the TOPO/TA reaction with either pCR2.1 or pCR8 (Invitrogen) using the protocol recommended by the manufacturer and transformed into chemically competent *E. coli*.

2.2.6 Restriction digest analysis of plasmids

Restriction digests of plasmids purified from *E. coli* or *A. tumefaciens* utilizing the Qiagen Miniprep kit were performed using restriction enzymes from New England Biolabs following the manufacturer's recommendations. Typically 1.5 µg plasmid DNA was digested in a 25 µL reaction.

2.2.7 DNA sequencing of plasmids and gene fragments

Sanger Sequencing was used to verify plasmid constructs and to sequence individual genes and gene fragments. Plasmids were purified from *E. coli* cultures using the Qiagen Miniprep Kit. Purified plasmid (500 ng) was combined with 8 pmole primer (Appendix 4) in 15 µL total volume for sequencing by Elim Biopharmaceuticals. To sequence genes from *A. thaliana*, the gene or gene fragment was first amplified from genomic DNA (2.2.1) using PCR (using gene specific primers and either JumpStart Red Taq ReadyMix (Sigma) or Phusion Polymerase (Finnzymes) following manufacturers recommendations). The PCR product was purified using the Qiagen PCR Purification Kit and 10-20 ng of DNA was used with 8 pmole primer in a 15 µL total volume for the sequencing reaction.

2.2.8 Transformation of *P. pastoris*

The protocol for transformation of *P. pastoris* was based on the EasySelect *Pichia* Expression Kit manual (Invitrogen). Plasmid DNA (~10 µg) was linearized by digestion with XhoI and purified using the Qiagen PCR purification protocol. The DNA was lyophilized and resuspended in 10 µL water. *P. pastoris* was grown in YPD (200 mL) to an OD_{600} of 1.2 then centrifuged at 1500 *g* for 5 min at 4 °C. The pellet was washed with sterile ice cold water (200 mL and 100 mL) and 1 M sorbitol (20 mL), resuspended in 0.4 mL ice cold 1 M sorbitol, and kept on ice. *P. pastoris* cells (80 µL) were added to 10 µL of linearized DNA and electroporated in a 1 mm gap cuvette at 750

V, 175 Ω, and 50 µF. The cells were resuspended in 1 mL 1 M sorbitol, incubated at 30 °C for three hours with no shaking, and then plated on YPD with 100 µg / mL zeocin.

2.2.9 Excision of selectable marker from genome of *P. pastoris*

A cassette with selectable and counter-selectable markers was included on transgene constructs for *P. pastoris* to allow for sequential gene insertion. The selectable marker used was a zeocin resistance gene, which allows for growth on media containing the antibiotic zeocin. The counter-selectable marker was the mazF RNase gene under the control of a methanol inducible promoter, which degrades mRNAs upon induction and thereby inhibits protein translation (Yang et al. 2009; Zhang et al. 2003). Direct repeats flanking the markers allows for homologous recombination to excise the entire cassette, leaving the desired transgene incorporated in the genome while remove the zeocin resistance gene. To induce expression of the mazF gene and thereby select for cells in which the cassette had been excised, *P. pastoris* cultures were grown overnight in YPD media on a shaker at 30 °C, then plated onto minimal methanol plates. Colonies appearing after 5-7 days were restreaked for purity on YPD plates and genotyped by colony PCR.

2.2.10 Transformation of *A. tumefaciens*

Electrocompetent *A. tumefaciens* (strain GV3101) was prepared as previously described (McCormac et al. 1998). Plasmid DNA (0.5 µL, approximately 75 ng) was added to the electrocompetent *A. tumefaciens* (80 µL) on ice and incubated for 5 min at 0 °C. The cells were transferred to a 0.1 cm gap cuvette and electroporated at 1800 V, 50 µF, 150 Ω. SOC media (Invitrogen) was added to the cells and they were incubated at 30 °C for three hours with shaking before plating on LB plates containing the appropriate antibiotics for selection (gentamycin and rifampicin plus plasmid resistance).

2.2.11 Transformation of *A. thaliana*

For plant transformation, the plasmids were first transformed into the *A. tumefaciens* by electroporation (2.2.10). *A. tumefaciens* was then grown in LB media containing the appropriate antibiotics on a shaker at 30 °C to an OD_{600} of approximately 2, centrifuged at 4,000 *g* for 10 min and resuspended to an OD_{600} of 0.8 in transformation buffer (3% (w/v) sucrose, 0.02% (v/v) Silwet L-77). Bolting *A. thaliana* plants (approximately four weeks old) were dipped into the *A. tumefaciens* suspension and a vacuum was applied three times each for three minutes (Clough and Bent 1998). Transformed *A. thaliana* plants were selected by plating the T1 seeds onto ½ MS, 1% sucrose, 0.8% agar plates with the appropriate antibiotic. Resistant plants were selected after two weeks, moved to soil, and genotyped (2.2.2) to confirm the presence of the transgene.

2.2.12 Infiltration of *N. benthamiana*

A. tumefaciens cultures were grown up to an OD_{600} of 1 in LB media containing the appropriate antibiotics. The *A. tumefaciens* were collected by centrifugation at 10,000 *g* for 5 min and resuspended in infiltration buffer to a final OD_{600} of 0.05 - 0.4. Two different infiltration buffers were used with similar results, one consisting of 10 mM MES, pH 6.0, adjusted with KOH, 2 mM

Na_2PO_4, 0.5% (w/v) glucose, and 100 µM acetosyringone (solubilized in DMSO), the other buffer consisting of 10 mM MES pH 5.6, 10 mM $MgCl_2$, and 150 µM acetosyringone solubilized in DMSO (Sparkes et al. 2006). The first infiltration buffer was used for subcellular localization of XLT2 whereas the second was used for AXY9. The infiltration was performed by submerging the leaves of *N. ben*thamiana in the infiltration buffer solution (containing *A. tumefaciens*), applying vacuum for 3 minutes and then releasing the vacuum slowly over 15 seconds.

2.2.13 Transcriptional analysis using qRT-PCR

RNA was extracted from *A. thaliana* leaf material using the Plant RNeasy kit from Qiagen following the manufacturer's protocol. cDNA was prepared from this RNA by using the Superscript III First-Strand Synthesis kit from Invitrogen. The qRT-PCR reaction was performed using Maxima SYBR Green/ROX qPCR Master Mix from Thermo Scientific for reagent and a StepOnePlus Real-Time PCR System from Applied Biosciences for thermocycling and fluorescence detection. The program used was 10 minutes at 95 °C followed by 40 cycles of 95 °C for 15 s, 60 °C for 30 s, and 72 °C for 30 s. A melting curve followed the amplification program to help verify the amplification products. The C_t parameter was calculated by the Applied Biosciences software and used to calculate the relative abundance of the transcripts by comparison to an internal control (transcript of the polypyrimidine tract-binding gene) and standard curves for each transgene using the corresponding plasmids. Primers are listed in Appendix 4.

2.2.14 Transcriptional analysis using RT-PCR

RNA was extracted from *A. thaliana* leaf material using the Plant RNeasy kit (Qiagen) following the manufacturer's protocol. Genomic DNA was removed by digestion of 1.7 µg RNA with 1 µL DNAse (Roche 04716728001) in a 20 µL total volume for 15 min at 37 °C. The DNAse was inactivated by incubation at 70 °C for 15 min. cDNA synthesis was performed using the M-MLV Reverse Transcriptase kit (Invitrogen 28025-013) following the manufacturer's instructions. The cDNA (1 µL) was used in a in a PCR reaction with gene-specific primers as described (2.2.2).

2.2.15 High-throughput sequencing of *A. thaliana*

Genomic DNA of high purity was extracted from *A. thaliana* grown on plates using the DNeasy DNA Extraction kit (Qiagen). This genomic DNA (5 µg) was submitted to the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley for library preparation and one lane of sequencing on a HiSeq 2000 (Illumina) with 100 bp single end reads. The resulting data was analyzed as described (2.4.1).

2.3 Microscopy

2.3.1 Subcellular localization

Protein localization was performed by generating fluorescent protein fusion constructs for the genes of interest and examining the resulting fluorescent signal in transformed *A. thaliana* or infiltrated *N. benthamiana* plants. The genes were cloned into the vector pMDC84 using gateway cloning (2.2.5.1) to create C-terminal GFP fusions (Curtis and Grossniklaus 2003). The

vectors were transformed in to *A. tumefaciens* (2.2.10) and the transgenic *A. tumefaciens* was used to infiltrate *N. benthamiana* leaves (2.2.12). Leaves were examined two, three, and four days post-infiltration to determine optimum viewing time, which was affected by the amount of protein expression and the health of the plants following the infiltration procedure. For localization in *A. thaliana*, stably transformed plants were used (2.2.11). A Zeiss LSM 710 laser scanning confocal microscope was used to image the fluorescent signal in the leaf epidermal cells of the infiltrated or stably transformed plants. A mannosidase tagged with CFP or mCherry was used as a Golgi-marker and control (Nelson et al. 2007). Lasers (405 nm, 488 nm and 594 nm) were used to excite the fluorescent proteins and band pass filters of 454-481 nm, 504-598 nm, and 599-696 nm were used for detection of CFP, GFP and mCherry, respectively.

2.3.2 Light microscopy (for stem sections) and staining

For the examination of *A. thaliana* tissues, plants were cross sectioned by hand and stained with toluidine blue (0.02% in water) for approximately 10 seconds. The stained sections were imaged using a light microscope (DM5000 B, Leica Microsystems).

2.4 Bioinformatic methods

2.4.1 Analysis of high-throughput sequencing data

The Illumina sequencing data from the Vincent J. Coates Sequencing facility was assembled against the *A. thaliana* TAIR9 reference genome using the CLC Genomics Workbench software (CLC Bio) and single nucleotide polymorphism (SNP) detection was performed. Non-EMS SNPs were removed by comparison with sequencing data from other mutants derived from the AXY screen by the use of a Python script.

2.4.2 Phylogenetic tree construction

Gene and protein sequences were obtained from a variety of sources including Phytozome (http://www.phytozome.net), the MSU Medicinal Plants Genomics Resource, SolGenomics (Bombarely et al. 2011), NCBI, and TAIR (http://www.arabidopsis.org) using BLAST (Altschul et al. 1997) and related search tools. Analysis of protein sequences was aided by use of the SeaView software (Gouy et al. 2010). This program aligned sequences using ClustalO (Sievers et al. 2011) or MUSCLE (Edgar 2004) and phylogenetic trees were construction using PhyML (Guindon et al. 2010).

2.5 Analytical methods

2.5.1 Preparation of alcohol insoluble residue (AIR) from plant tissue

To isolate a cell wall-enriched fraction for analysis, plant tissue was snap frozen (small scale) or freeze dried (large scale) prior to grinding with metal beads in a Retsch mill. For leaf, root, and etiolated seedling tissue, 25 Hz for 2 min 30 s was used. For stem tissue, material was ground at 30 Hz for 3 min and the grinding procedure was repeated if necessary to obtain a fine powder free of visible stem fragments. The powdered material was resuspended in 70% ethanol (v/v) and centrifuged (21,000 *g*, 5 min). The supernatant was discarded and the extraction with 70% ethanol was repeated until the supernatant was colorless. The remaining pellet was then resuspended in 1:1 chloroform methanol (v/v) and centrifuged (21,000 *g*, 5 min). The

supernatant was discarded and the extraction was repeated until the supernatant was no longer colored. The remaining AIR material was dried at room temperature.

2.5.2 Preparation of alcohol insoluble residue (AIR) from *P. pastoris*

Acid-washed glass beads (approximately 100 µL total volume, 425 – 600 µm diameter, Sigma G8772) were added to the pellet from 1 mL of *P. pastoris* culture. The samples were ground in at 25 Hz for 2 min 30 s at room temperature. The material was extracted sequentially with 70% ethanol, 1:1 chloroform:methanol (v:v), and acetone with the samples vortexed, centrifuged (21,000 *g*, 10 min) and the supernatant removed each time. The remaining material AIR material was dried at room temperature.

2.5.3 Analysis of XEG-released oligosaccharides by MALDI-TOF MS (XyG OLIMP)

The XyG oligosaccharide mass profiling (OLIMP) method (Lerouxel et al. 2002) was performed on a variety of plant tissue types. In general, two metal balls (3 mm diameter) were added to the material in a 1.7 mL tube (PosiClick, Denville) and the material was snap frozen in liquid nitrogen and ground for 2 min 30 s at 25 Hz in a Retsch Mill. The material was then extracted with several solvents, with each extraction consisting of resuspension of the material in 1 mL solvent by shaking in the Retsch mill for 10 s, 3 min centrifugation at 21,000 *g*, and removal of the solvent. The solvents used were 70% ethanol for the first extraction, 1:1 (v/v) chloroform:methanol for the second, and 50 mM ammonium formate pH 4.5 for the third extraction (this buffer wash was not performed when OLIMP was performed on etiolated seedlings). Following the extractions, the pellets were dried in a vacuum centrifuge and digested with 25 µL total volume containing 0.1 U XEG in 50 mM ammonium formate (1 U XEG is defined as amount required to release 1 µmol XyG oligosaccharide per minute from a digest containing 0.1% tamarind XyG in 50 mM sodium acetate pH 5.0 digested at 21 °C for one hour (Pauly 1999)). The digest was performed on a shaker at 37 °C for at least four hours or overnight. When performing OLIMP on single hypocotyls, the digest was dried and then resuspended in 5 µL water. Washed Biorex MSZ501 cation beads were used to desalt the digest supernatant by incubation with approximately 15 beads for 45 minutes. The MALDI-TOF plate was prepared by spotting 2 µL of 10 mg / mL 2,5-dihydroxybenzoic acid (DHB) onto each well and drying under vacuum. The desalted digest supernatant (2 µL) was then spotted onto the DHB matrix, incubated for 3 minutes at room temperature, and dried under vacuum. An AXIMA Performance MALDI-TOF mass spectrometer (Shimadzu) was used in positive linear mode with an accelerating voltage of 20,000 V to obtain the mass profiles.

2.5.4 Analysis of XEG-released oligosaccharides using HPAEC-PAD

AIR (2 mg, 2.5.1) was washed with 1 mL of 20 mM ammonium formate pH 4.5 to remove buffer soluble components. The pellet was then digested with 0.4 U XEG in 100 µL of 20 mM ammonium formate, pH 4.5, overnight at 37 °C with shaking. The digest was centrifuged at 21,000 *g* and the supernatant was diluted 4X prior to injection of 25 µL on the HPAEC-PAD. The oligosaccharides were separated using a CarboPac PA200 column with an eluent of 100 mM NaOH with a gradient of 0-80 mM sodium acetate over 15 min with a flow rate of 0.4 mL / min. For analysis of XXSG (5.2.3), which elutes several minutes after XXXG, the program was extended to a gradient of 0 to 110 mM sodium acetate over 22 minutes with 100 mM NaOH.

Peaks were identified by comparison of retention time to standards and mass (determined by fraction collection and MALDI-TOF (desalted using reverse phase column, 2.5.5). Oligosaccharides were quantified by comparison to a standard curve generated from commercially available XXXG (Megazyme).

2.5.5 Desalting / purification of XyG oligosaccharides

A Supelclean ENVI-Carb reversed-phase column (57109-U, Supelco) was equilibrated with 50% acetonitrile and washed four times with water. The sample was applied to the column, the column washed with water (4 mL) and the oligosaccharides eluted with 50% acetonitrile. Light pressure from a Pasteur pipette bulb was used to push the solutions through the column. The eluent was dried under vacuum and then dissolved in water for further analysis.

2.5.6 Purification of novel oligosaccharide (XXSG)

AIR material from etiolated seedlings (150 mg) was extracted with 4 M potassium hydroxide containing 10 mM sodium borohydride for four hours with shaking. The supernatant containing extracted XyG was removed and neutralized with a mixture of 1 part acetic acid (as buffer) and 10 parts hydrochloric acid. Polysaccharides were precipitated from this solution by incubation in 70% aqueous ethanol overnight at -20 °C. The resulting pellet was washed repeatedly with 70% ethanol to remove remaining salts, dried, and digested with 16 U XEG in 4 mL of 50 mM ammonium formate, pH4.5, overnight at 37 °C. The digest supernatant was dried and the oligosaccharides were reduced by incubation with 1 M ammonium hydroxide containing 10 mg/mL sodium borohydride at room temperature for one hour. The reduction reaction was neutralized with glacial acetic acid and desalted (2.5.5). The oligosaccharides were dissolved in 400 µL water and the oligosaccharide of interest was collected after separation by HPAEC-PAD. The oligosaccharide was again desalted (2.5.5).

2.5.7 Glycosidic linkage analysis

AIR material (1 mg) or oligosaccharide (~100 µg) was dissolved in a suspension of dry DMSO and NaOH (12.5 mg / mL). The methylation procedure, based on Ciucanu 1984, was performed by two additions of 100 µL methyliodide with 1 hour stirring following each addition(Ciucanu and Kerek 1984). The reaction was quenched with water (2 mL). Dichloromethane (2mL) was used to extract the partially methylated carbohydrates. The dichloromethane was evaporated under a stream of dry nitrogen and the residue was hydrolyzed by incubation with 2M trifluoroacetic acid for 90 min at 121 °C. The trifluoroacetic acid was removed by evaporation under a stream of nitrogen at 40 °C followed by two additions and evaporations of 300 µL isopropanol. The samples were reduced by addition of 200 µL 1 M ammonium hydroxide containing 10 mg/mL sodium borodeuteride and incubation at room temperature for one hour. The reduction was halted by addition of 150 µL acetic acid. Three additions and evaporations of 250 µL of 9:1 (v:v) methanol:acetic acid, followed by four additions and evaporations of methanol to remove resulting borate side-products. The evaporations were performed under a stream of nitrogen at 40 °C. The samples were per-*O*-acetylated by addition of 50 µL acetic anhydride and 50 µL pyridine, incubated at 121 °C for 20 min. The solvents were removed by evaporation at room temperature, followed by two additions and evaporations of toluene (200 µL). The partially methylated alditol acetates were extracted using 1.2 mL ethyl acetate and 5 mL water. The organic phase was transferred to a new tube and dried under nitrogen. The sample was

subsequently dissolved in acetone and injected subjected to GC-MS analysis using a 7890A GC system with an SP-2380 column (Supelco) and a 5975C MS detector (Agilent) in EI mode. Following injection the column was held at 180 °C for 2 min, heated (20 °C per min) to 200 °C and held for 6 min, then heated (20 °C per minute) to 245 °C and held for 12 min. Peaks were assigned by comparison of retention time and fragmentation pattern to standards and utilizing the CCRC database of partially methylated alditol acetates (http://www.ccrc.uga.edu/databases/).

2.5.8 Extraction of XyG from olive fruit

Olive fruit (45 g fresh weight) was ground with mortar and pestle then divided between two 50 mL Falcon™ tubes. The material was washed 4X with 15 mL hexanes (3220 *g* centrifugation). The material was dried and then ground in the Retsch mill for 1 min at 25 Hz. The material was divided between 16 2 mL tubes and extracted 4X with 1 mL hexanes, 4X with 1 mL 70% ethanol, 3X with 1 mL 1:1 (v:v) chloroform:methanol, and once with 1 mL acetone. A centrifugation (20,000 *g*, 3 min) was performed to pellet the wall material and allow for the solvent to be removed after each extraction. The resulting AIR material (200 mg out of 2 g total yield)) was extracted with 25 mL 4 M KOH with 10 mM $NaBH_4$ with vigorous shaking for four hours at room temperature. The supernatant after centrifugation (10 min, 3220 *g*) was transferred to a new tube and neutralized with glacial acetic acid (1 part) and HCl (10 parts). Ethanol was added to a final concentration of 70% and the samples were placed at -20 °C overnight. The supernatant after centrifugation (3220 *g*, 5 min) was discarded and the pellet was washed with 70% until no visible salt remained. The resulting pellet was (approximately 36 mg total yield) was digested with 20 U XEG in 4 mL of 50 mM ammonium formate pH 4.5 at 37 °C with shaking overnight. The digest mixture was dried and 70% ethanol (1 mL) was added. The sample was vortexed thoroughly, centrifuged (5 min, 21000 *g*), and the supernatant was transferred to a new tube and dried under vacuum. The resulting material was resuspended in 1 mL H_2O and the XyG oligosaccharides were purified using a reversed-phase column (2.5.5). The resulting material was dissolved in 1 mL H_2O and diluted 20 to 100X for MALDI-TOF and injection onto the HPAEC-PAD.

2.5.9 Nucleotide sugar extraction and analysis

P. pastoris cells were collected by centrifugation of 1 mL culture at 2000 *g* for 1 min. The supernatant was removed and the cells were washed with 1 mL H_2O. The supernatant was removed and methanol (500 µL) was added with approximately 100 µL of acid-washed glass beads. The cells were put in the Retsch mill and ground for 2 minutes at 30 Hz. The broken cells were centrifuged at 21,000 *g* for 2 min at 4 °C. The supernatant was filtered through a 3kD molecular weight cutoff filter and the flow-through was dried under vacuum. The dried pellet was resuspended in 200 µL Buffer A (5mM ammonium acetate, pH 7.2), centrifuged, and the supernatant was injected onto an Agilent 1200 Series LC System supplied with a 250 mm x 2.1 mm diamond hydride column 4 µm from Cogent. Detection of the eluent was performed using a Thermo Scientific LTQ XL electrospray mass spec in negative mode. Separation of the nucleotide sugars was achieved using a flow rate of 0.3 mL per minute with a linear gradient from 0 to 60% Buffer A over 20 minutes with a 10 minute equilibration period. The second buffer was 90% acetonitrile with 10 mM ammonium acetate pH 6.5.

2.5.10 Measurement of acetic acid content of AIR material

AIR material (1 mg for leaf, 0.5 mg for stem) was suspended in 100 µL water. The material was saponified by addition of 100 µL 1 M sodium hydroxide and incubation at 25 °C for 1 h with shaking at 600 rpm. Saponification was halted by addition of 100 µL 1 M HCl and the samples were centrifuged at 21,000 *g* for 10 min. The supernatant (10 µL) was transferred to a well of a 96-well microtiter plate and the acetic acid content measured using the Acetic Acid Kit (K-Acet, Megazyme) with the reaction scaled down as follows. The assay was performed by addition of 30 µL of Solution 1, 12 µL of Solution 2, and 94 µL water to the 10 µL sample. The absorbance at 340 nm was read after 3 min and 12 µL of a 10X dilution of Solution 3 was added. The absorbance was measured after 4 min and followed by an addition of 12 µL of a 10X dilution of Solution 4. A final absorbance measurement was taken after 12 minutes. These absorbance values were used to calculate the acetic acid content relative to a standard curve following the formulas in the kit's manual.

2.5.11 NMR spectroscopy of the oligosaccharide

The purified oligosaccharide was dissolved in D_2O, freeze dried, and dissolved in D_2O with an internal reference standard of DSS (3-(Trimethylsilyl)-1-propanesulfonic acid) (0.01 mg/mL). A broker AVANCE 600 MHz NMR spectrometer with an inverse gradient 5-mm TXI CryoProbe was used to record the ^1H NMR spectra. The TOCSY experiment was performed using the standard Bruker pulse program "dipsi2ph" with the following parameters: 2048 x 256 data points for F2 x F1 dimensions, 48 scans, inter-scan delay of 1.5 s, 4 dummy scans, a pre-scan delay of 69.6 µs and a mixing time of 200 ms. Bruker's Topspin 3.1 software was used for data processing and analysis.

2.5.12 OREX protein purification and digestion

The oligoxyloglucan reducing-end specific xyloglucanase (OREX) enzyme (Yaoi and Mitsuishi 2002) from *Aspergillus nidulans* was purified from transgenic *P. pastoris* expressing the enzyme as described (Bauer et al. 2005). In brief, *P. pastoris* was grown in 50 mL BMGY at 28 °C to an OD600 of between 3 and 6 (250 rpm shaking). The cells were then spun down (10 min, 5000 *g*) and resuspended in MM media, and grown for 3 days with a daily addition of 0.3 mL methanol. Cells were spun down (10 min, 5000 *g*) and the supernatant was concentrated 10X with 10 kDa cutoff filter (Ultra-15, Amicon) by centrifugation at 5000 *g* at 4 °C. The protein (His-tagged) was purified using NiNTA beads (Qiagen) following the manufacturers recommendations. Digestion of the purified, non-reduced oligosaccharide was performed in 50 mM glycine buffer at pH 3.0 overnight at 37 °C. The digestion products were analyzed using MALDI-TOF MS (2.5.3).

2.5.13 Leaf extensibility

Sections (8 mm) of petiole material from 4-week old plants were used for acid-induced wall extension and stress-strain assays as previously described (Park and Cosgrove 2012a). For the acid-induced wall extension assay, the clamped segments were incubated first in neutral buffer (250 µL 20 mM HEPES, pH 6.8) for 30 min with a load of 7.5 g. The buffer was then replaced with 250 µL of 20 mM sodium acetate (pH 4.5) containing 5 mM dithiothreitol. A position transducer was used to record the displacement every 30 s. The stress-strain assays were performed by heat-inactivating the petiole sections and placing them in an extensometer. An

extension rate of 3 mm/min was used and the force was recorded and used to calculate the plastic and elastic wall compliances (Cosgrove 1989; Yuan et al. 2001).

2.5.14 Immunodetection of XyG by dot blot

The cotyledons of developing nasturtium seeds were ground and washed with 70% ethanol and 1:1 chloroform methanol (v/v) to remove soluble components. The remaining residue was extracted with 4 M KOH with 10 mM $NaBH_4$ overnight with shaking at room temperature. The supernatant was partially neutralized with acetic acid before spotting onto nitrocellulose membrane with dilutions in 0.8 M KOH. The nitrocellulose membrane was blocked with 5% dry milk in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) for 30 minutes and the immunodetection performed using the CCRC-M1 or CCRC-M100 antibodies (Pattathil et al. 2010; Puhlmann et al. 1994) as primary (1:500) and a goat anti-mouse:HRP (170-6516, Bio-Rad) antibody as the secondary in a 1:3000 dilution.

3 A forward genetic screen to identify genes involved in XyG biosynthesis and metabolism

3.1 Background

While several of the GTs and GHs involved in XyG biosynthesis and metabolism are known (1.2.3), less well understood are other classes of genes likely required including those encoding transcription factors, trafficking proteins, proteins affecting post-transcriptional regulation of XyG biosynthesis/metabolism, and proteins involved in the production and transport of activated donor substrates. Uncovering these genes is challenging due to the fact that, unlike for the GTs, it is difficult to estimate the number of distinct activities required or to select candidate genes based on gene family. A forward genetics screen was undertaken to overcome these limitations and to provide a complementary approach to the reverse genetics approaches (4, 5) presented in this thesis.

3.1.1 Efficacy of forward genetics in identifying cell wall related genes

Forward genetics has been a productive approach for identifying genes involved in cell wall biosynthesis. The *mur* screen for *A. thaliana* mutants with altered cell wall monosaccharide compositions resulted in the identification of GTs (MUR2, MUR3, MUR10) and nucleotide sugar interconversion enzymes (MUR1, MUR4) (Bosca et al. 2006; Burget and Reiter 1999; Madson et al. 2003; Reiter et al. 1997; Vanzin et al. 2002). Other screens have identified mutants by using various assays including a conditional root expansion phenotype (Hauser et al. 1995), radial root swelling (Arioli et al. 1998), examination of stem for irregular xylem (Turner and Somerville 1997), FT-IR on hypocotyl walls (Mouille et al. 2003), and morphological response to treatment with glycosyl hydrolases (Gille et al. 2009). While typical reverse genetics approaches require genomic sequence information, the availability of knockout alleles and information on putative gene family function (Brown et al. 2005), forward genetics can be used to identify genes from families of unknown function as prior information about gene function is not required (Gille et al. 2011b). An additional advantage of forward genetics is that it can demonstrate the importance of a previously characterized gene to a process not readily associated with the known function of the corresponding protein. An example of this is the XyG galactosyltransferase MUR3, which in addition to its XyG galactosyltransferase activity has been reported to have a direct role in cytoskeleton organization and an indirect function in activating plant defense response (Tamura et al. 2005; Tedman-Jones et al. 2008). The short generation time, small size (Page and Grossniklaus 2002), well annotated-genome (The Arabidopsis Initiative 2000, http://www.arabidopsis.org), and T-DNA collections (Alonso et al. 2003; Rosso et al. 2003) available for *A. thaliana* have made this a popular model plant species for forward genetic screens. T-DNA collections are useful for forward genetics because they can provide independent mutant alleles to ensure that disruption of a particular gene is responsible for the observed phenotypes. High-throughput sequencing methods such as the platforms offered by Illumina, 454, and Ion Torrent can be used to re-sequence the entire genome of *A. thaliana* mutants quickly (Loman et al. 2012), facilitating the identification of candidate mutations (Austin et al. 2011; Laitinen et al. 2010; Ossowski et al. 2010).

Forward genetics relies upon populations harboring genetic variation resulting in mutant phenotypes that can be identified in a screen. Genetic variability can be created by mutations induced by chemical mutagenesis (EMS), radiation (X-ray, fast neutron) (Koorneef 1982), T-DNA insertion (Krysan et al. 1999), or transposon activation (Long et al. 1993) as the level of spontaneous mutation is quite low (Ossowski et al. 2010). The location of insertions in T-DNA and transposon mutants can often be identified using PCR-based methods (McKinney et al. 1995), facilitating identification of the genetic basis of these mutants. The identification of mutations in EMS and X-ray-treated populations is usually more time consuming as they must often be mapped and at least partially resequenced, as the point mutations cannot serve as primers for PCR-based techniques. While EMS and X-ray treatment usually lead to point mutations, fast-neutron mutagenesis can result in either point mutation or deletions (Belfield et al. 2012). The frequency of introduced mutations by these methods is variable depending on the treatment conditions, but can result in several hundred to several thousand mutations introduced per individual (Belfield et al. 2012; Uchida et al. 2011). As well as requiring different methods for identification of the causative mutation, the different methods of mutagenesis can lead to different types of alleles recovered. T-DNA mutagenesis can be performed with an insert containing a strong promoter, allowing for either disruption of gene function if the T-DNA is inserted into the coding sequence of a gene or overexpression of the gene if the T-DNA is inserted into the promoter region (Weigel et al. 2000). EMS treatment, which typically introduces G to A point mutations, can result in missense or nonsense mutations (Drake and Baltz 1976). Alleles recovered from EMS treatment are often recessive mutations (Gille et al. 2011b; Gunl et al. 2011; Gunl and Pauly 2011), but as most of the coding sequence is intact a point mutation could result in a dominant allele as well (Bleecker et al. 1988).

One of the notable limitations of forward genetics approaches is its inability or difficulty identifying relevant genes due to genetic redundancy or gene lethality. *A. thaliana*, like many plant species, often has multiple copies of genes with similar biochemical activity (Vuttipongchaikij et al. 2012). In those cases knocking out a single copy of a gene will not result in a strong or any observable phenotype because paralogs are able to function redundantly (Cavalier et al. 2008; Jensen et al. 2011; Verhertbruggen et al. 2013). The probability of EMS-induced mutagenesis resulting in the disruption of two or more redundant genes decreases exponentially with redundant gene number, so screens are unlikely to pick up fully redundant genes. Because the wall is important for cell structure and function, it is not uncommon for the disruption of cell wall biosynthetic genes to result in decreased plant fitness or even lethality (Geshi et al. 2013; Goubet et al. 2003; Sitaraman et al. 2008). Lethal mutations are typically not identified in forward genetic screens. One way to mitigate this problem is to screen under conditions that plants are typically not propagated in and therefore identify conditionally-sensitive alleles. These conditionally-sensitive alleles may display a cell wall and growth phenotype under a stress condition but have relatively normal function and growth under regular conditions. This approach has been used to identify mutants partially defective for the cellulose synthase gene CesA1 under high temperature (Arioli et al. 1998) and low light (Desnos et al. 1996). The identification of XyG mutants seems favorable, as *A. thaliana* plants can tolerate significant genetically-induced structural perturbation of this polysaccharide (1.2.1) (Cavalier et al. 2008; Madson et al. 2003; Pena et al. 2012; Vanzin et al. 2002).

3.1.2 The AXY (Altered XYloglucan) Screen

To complement the reverse genetic approaches described in this thesis and to look for additional biosynthetic and regulatory genes, a forward genetic screen was performed for *A. thaliana* mutants with Altered XYloglucan (the AXY screen). The AXY screen was performed using OLIMP in combination with a XyG specific endoglucanase (XEG) (Pauly et al. 1999b) to screen etiolated seedlings of EMS-mutagenized *A. thaliana* plants (Lerouxel et al. 2002). This method can detect the relative abundance of XyG oligosaccharide ions but not their absolute quantities, therefore restricting the types of mutants that can be recovered to mutants with changes in the relative abundance of one or more XyG oligosaccharides. As the OLIMP method is destructive to the etiolated seedlings on which the screen was performed, plants from mutant pools (pooled M2 seeds) with mutant phenotypes had to be grown individually for seed and the progeny screened again to verify the mutant and obtain homozygous lines. After the initial identification of mutants, classical genetic mapping (Konieczny and Ausubel 1993)and high throughput sequencing (Mardis 2008) were used to identify the causative mutations.

Approximately 20,000 M2 etiolated seedlings were screened using XyG OLIMP resulting in the identification of approximately 33 distinct mutants. The mutants were placed into groups based on the observed phenotype (such as increased galactosylation or decreased fucosylation of XyG) (Table 3-1). Groups with phenotypes similar to previously identified mutants including *mur3* (decreased galactosylation and fucosylation) and *mur2*/*mur1* (decreased fucosylation), were subjected to Sanger sequencing of the corresponding suspected locus to check if these mutants were likely allelic to previously identified XyG mutants. Mutants found to be free of mutations in these known XyG biosynthetic genes and those with novel XyG OLIMP phenotypes were investigated further to identify the causative mutation. Using this approach several proteins related to XyG have previously been identified including AXY8, a XyG α-1,2-L-fucosidase, AXY3, a XyG α-1,6-D-xylosidase, and AXY4, a putative XyG *O*-acetyltransferase (Gille et al. 2011b; Gunl et al. 2011; Gunl and Pauly 2011). The forward genetics portion of this thesis consisted of confirming and characterizing thirteen additional *axy* mutants (Table 3-1).

Pool	Plant	Name	Gene / Status	Mutation
↑ Gal				
146	98		(line lost)	
196	70		*	
14	52		(line lost)	
No Fuc				
2	10	*axy1.1*	MUR1	
237	29		MUR1	
319	58		MUR2	
129		*axy2.4*	*	
↑ Gal, ↓ Fuc				
148	45	*axy3.1*	AXY3	E630K
160	55	*axy1.160*	*	
319	8	*axy1.3*	*	
236	81		*	
148	100		*	
↓ Ac				
58	79	*axy4.1*	AXY4	P126S
179	83	*axy9.1*	*	

↑ Ac				
172	83			
333	90			
8	59			
300	24			
80	36			
174	45			
233	98			
↑ Fuc				
14	99	*axy6.1*	*	
↓ Gal, ↓ Fuc				
43	45		MUR3	G190D
8	52		*	
↑ Fuc, unusual oligos				
43	133	*axy8.1*	AXY8	W481Stop
80	56	*axy8.2*	AXY8	R398Stop
160	4	*axy8.3*	AXY8	P562L
194	10	*axy8.4*	AXY8	C511Y
148	41	*axy8.7*	AXY8	W441Stop
8	64		*	
129	16		*	
142	71		*	
148	69		*	

Table 3-1. Summary of the *axy* mutants. The mutants are grouped based on observed XyG OLIMP phenotype. The genetic basis for each of the previously characterized mutant lines is listed. * indicates mutants that were selected for characterization as part of this thesis work. M1 plants were grouped into pools for seed collection and subsequent screening of M2 etiolated seedlings. To obtain mutant lines, pools with observed phenotypes were grown and M3 seeds were collected separately from each M2 plant (plant number listed in the table). For corresponding XyG oligosaccharide spectra see Figure 3-1, A5-1, A5-2, and A5-3. Abbreviations: Ac – acetylation, Gal – galactosylation, Fuc – fucosylation.

3.2 Results

Thirteen candidate *axy* mutants were subjected to OLIMP analysis to confirm their previously observed phenotype (2.5.3). Abnormal XyG OLIMP spectra were observed for eleven of the mutants, which were placed into six phenotypic classifications (Figure 3-1, Figure 13-1, Figure 13-2, Figure 13-3). Of these phenotypic classifications, four (*mur1/mur2*-like, *mur3*-like, *axy3*-like and *axy8*-like) were similar to previously published XyG mutants (Gunl et al. 2011; Gunl and Pauly 2011; Madson et al. 2003; Vanzin et al. 2002) whereas the *axy6* and *axy9* classes represented XyG phenotypes that had not been previously reported. All eleven mutants were analyzed further to determine the genetic basis of the observed XyG oligosaccharide mass profile.

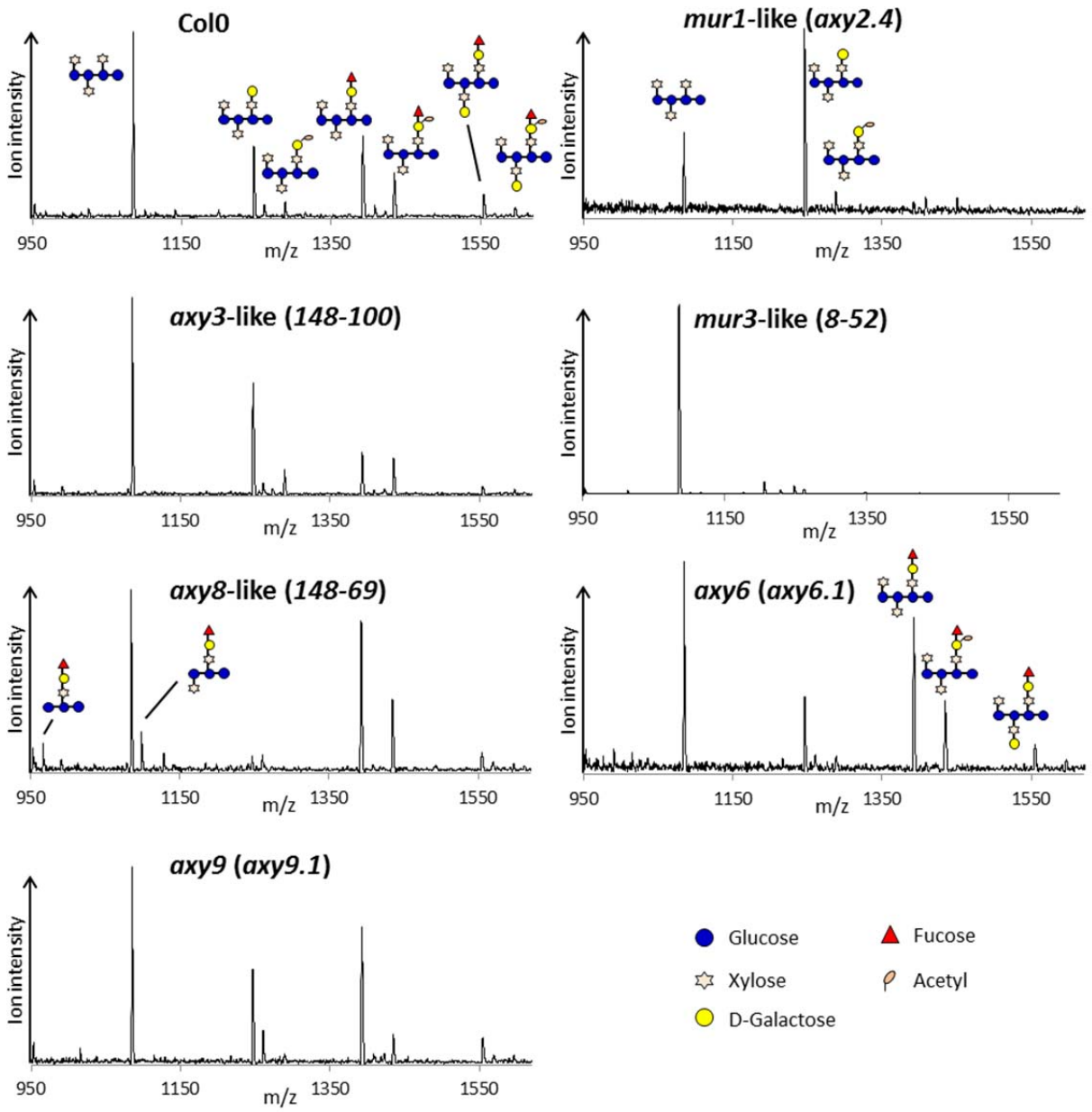


Figure 3-1. XyG OLIMP spectra *axy* mutant classes. MALDI-TOF MS analysis was performed on XEG-released oligosaccharides from cell wall material of etiolated seedlings and representative spectra are shown for the indicated mutant classes. The *mur1*-like mutants lack XyG fucosylation, whereas *axy8*-like and *axy6* mutants have increased fucosylation with novel oligosaccharides present in the *axy8* spectrum. The *axy3*-like mutants have increased relative abundance of XXXG and XXLG. *mur3*-like mutants have very little XyG substitution in etiolated seedlings. The *axy9* mutant has decreased XyG acetylation as observed by decreased abundance of XXFG and concomitant increase in XXFG.

3.2.1 Identification of the genetic basis of four mutants with an *axy3*-like phenotype

The *AXY3* gene (*At1g68560*) was previously identified from the AXY screen and found to be a xylosidase that acts on XyG oligosaccharides (Gunl and Pauly 2011; Sampedro et al. 2010; Sampedro et al. 2001). A defect in AXY3 results in an increased relative abundance of XXXG and XXLG in the OLIMP spectra. AXY3 is able to affect the polymeric form of XyG due to XyG oligosaccharides being reincorporated into the polymeric structure by the action of XTHs (1.2.4) (Gunl et al. 2011). Four of the *axy* mutants examined in this work, *236-81*, *axy1.160*, *148-100*, and *axy1.3*, had phenotypes matching that of the already characterized *axy3* alleles based on increased relative abundance of XXXG and XXLG and decreased relative abundance of fucosylated oligosaccharides (Figure 3-1, Figure 13-1). Sanger sequencing of the AXY3 locus in these mutants revealed that all four had mutations within the coding sequence of the AXY3 gene (Figure 3-2, Appendix 1, 2.2.7). For one of the mutants, *axy1.3*, an allelism test was performed by crossing the mutant to Col0 and *axy3.2*, a null allele for AXY3. The XyG oligosaccharide profiling of the F1 resulting from the cross to Col0 indicated that the *axy1.3* mutant contains a recessive mutation as the phenotype was rescued (Figure 3-3). The XyG oligosaccharide profiling of the F1 of the cross to *axy3.2* had the mutant phenotype indicating that the two mutations are allelic as *axy3.2* is also a recessive mutation. The remaining alleles were not tested by genetic complementation but likely also represent alleles of *axy3*. For *axy1.160* and 236-81, the mutations within the coding sequence result in early stop codons, W482Stop and W468Stop, respectively, which end the protein before the missense mutation in the *axy1.3* allele. The remaining allele, 148-100, has a missense mutation resulting in E630K (Figure 3-2, Appendix 1).

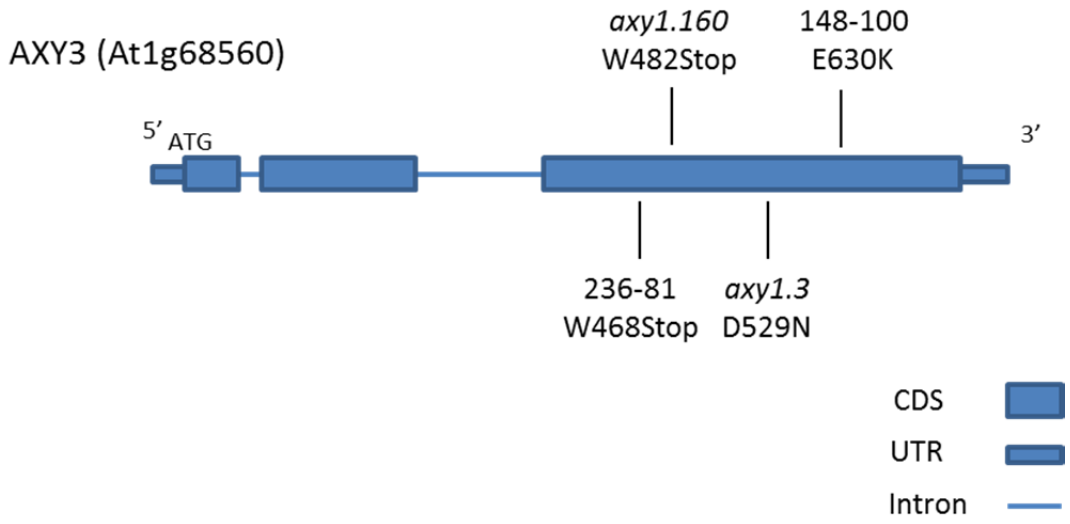


Figure 3-2. *AXY3* gene model. The *AXY3* gene (*At1g68560*) contains two introns and codes for a protein of 915 amino acids. The protein has an N-terminal signal peptide and is an apoplastic α-1,6-D-xylosidase active on XyG oligosaccharides. Four alleles of *axy3*, in addition to those previously described (Gunl and Pauly 2011), were identified in this work. CDS – coding sequence, UTR – untranslated region.

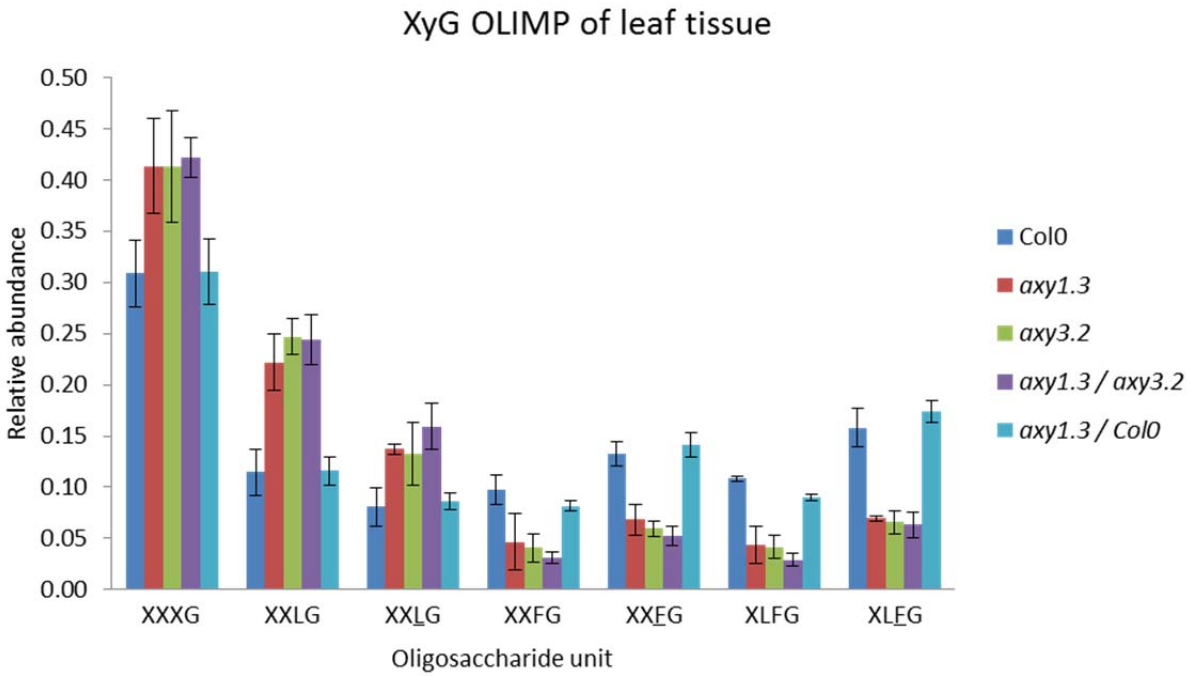


Figure 3-3. Genetic complementation of *axy1.3*. The *axy1.3* mutant, containing a mutation within the *AXY3* gene, was crossed to Col0 and the recessive *axy3.2* mutant for an allelism test. The heterozygous *axy1.3* / Col0 plants exhibited a wild type phenotype as revealed by XyG OLIMP of leaf tissue, indicating that the *axy1.3* allele is recessive. The phenotype of the heterozygous *axy1.3* / *axy3.2* line matched that of the homozygous *axy1.3* and *axy3.2* mutants, indicating *axy1.3* is allelic to *axy3.2*. n = 4. Error bars show standard deviation.

3.2.2 Characterization of three mutants with an *axy8*-like phenotype

The AXY8 protein represents a fucosidase that, like AXY3, is active on XyG oligosaccharides (Gunl et al. 2011; Leonard et al. 2008). A defect in the *AXY8* gene (*At4g34260*) results in an increase in fucosylated XyG oligosaccharides and unusual fucosylated XyG oligosaccharides as revealed by XyG OLIMP (2.5.3). These novel oligosaccharides likely include GFG and XFG, which are observed at an m/z of 967 and 1099 respectively (Gunl et al. 2011). Three mutants with an *axy8*-like phenotype, *148-69*, *142-71*, and *129-16* (Figure 13-2), were found to have mutations within the *AXY8* locus (Figure 3-4, Appendix 1, 2.2.7). In two cases, the mutation resulted in stop codons and in the third case the mutation disrupted the splice junction at an intron/exon border. All three of these mutations, occurring relatively early in the coding sequence of the gene, would result in major disruptions to the polypeptide sequence of the AXY8 protein.

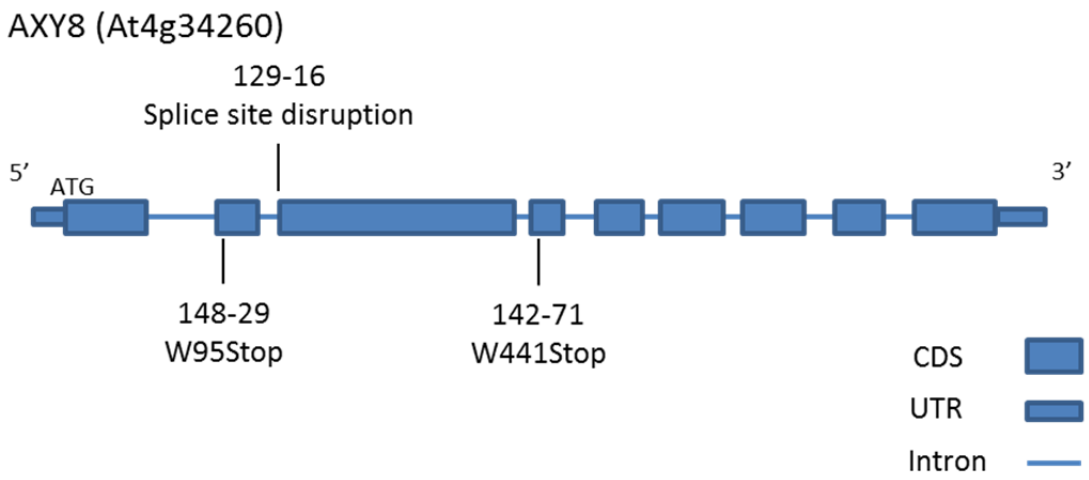


Figure 3-4. *AXY8* gene model. The *AXY8* gene (*At4g34260*) contains eight introns and codes for a protein of 843 amino acids. The protein has an N-terminal signal peptide and is an apoplastic α-1,2-L-fucosidase active on XyG oligosaccharides. Three alleles of *axy8*, in addition to those previously described (Gunl et al. 2011) were identified in this work. CDS – coding sequence, UTR – untranslated region.

3.2.3 Identification of the genetic basis of *mur1*-like and *mur3*-like *axy* mutants

One of the *axy* mutants analyzed in this work, *8-52*, had a *mur3*-like phenotype due to a strong decrease in observed galactosylation and fucosylation in XyG OLIMP (Figure 3-1, A5-1, 2.5.3). Sanger sequencing revealed a point mutation in the MUR3 coding sequence resulting in a missense mutation of C400Y, which is likely the genetic basis of the mutant phenotype (Figure 3-5, Appendix 1). A *mur1*-like phenotype was observed for the *axy2.4* mutant (Figure 3-1, Figure 3-6A). The MUR1 locus was sequenced in this mutant (performed by Markus Günl) but no mutation was found so the genomic region harboring the mutation was identified using a classical mapping approach. For the mapping *axy2.4* was crossed to Ler and mutants were selected from the F2 for genetic linkage analysis (Figure 3-6). The segregation ratio of mutants to wild type plants was consistent with a single recessive locus (Figure 3-6B) and genetic linkage analysis mapped the mutation to a 0.9 Mb region on chromosome 3 between markers 2fm7 and 2fm8 (Figure 3-7, Figure 3-8, 2.2.3). This region contains the MUR1 locus, so the mutant was crossed to the *mur1.1* mutant for an allelism test. The phenotype of the heterozygous F1 plants matched that of the *axy2.4* and *mur1.1* plants indicating that the mutants are allelic (Table 3-2). Indeed, a second round of Sanger sequencing of the MUR1 locus in *axy2.4* resulted in the identification of a mutation in *MUR1* leading to a single amino acid change S183F (Figure 3-9, Appendix 1, 2.2.7).

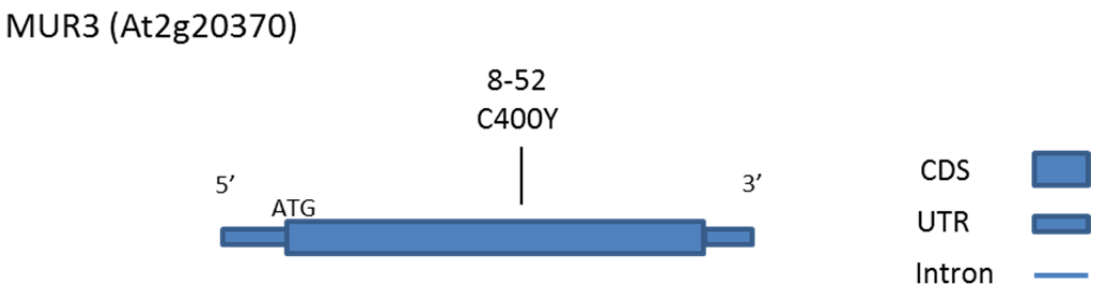
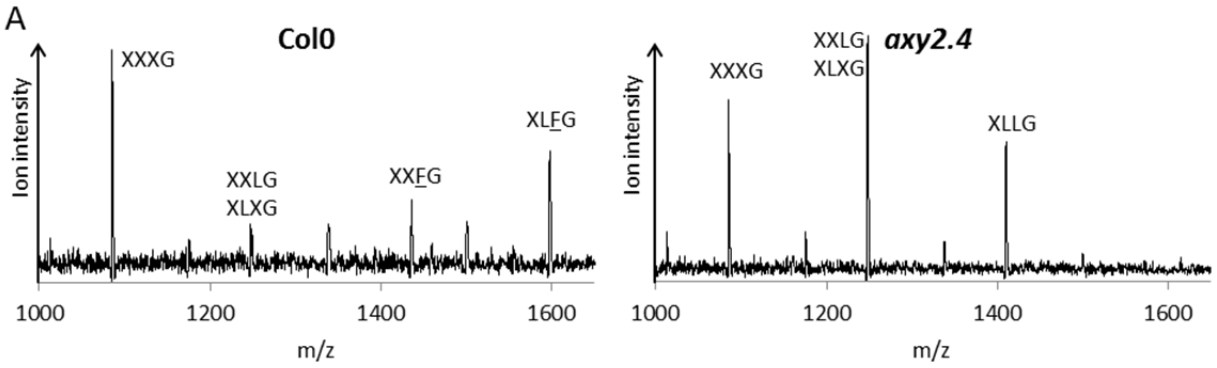


Figure 3-5. MUR3 gene model. The MUR3 gene (At2g20370) contains no introns and codes for a XyG galactosyltransferase. The enzyme consists of 619 amino acids and has an N-terminal transmembrane domain. A mutant with a mur3-like phenotype was found to have an alteration within the MUR3 coding sequence resulting in the missense mutation C400Y. CDS – coding sequence, UTR – untranslated region.



B

	Wild type	Mutant	Unknown
Count	152	57	31
Ratio	2.7	1.0	0.5

Figure 3-6. Mutant selection for mapping of *axy2.4*. A segregating F2 population of *axy2.4* crossed to the *A. thaliana* ecotype Landsberg erecta was established. The OLIMP phenotype of leaf material was used to select plants with a mutant phenotype within this population. The *axy2.4* phenotype (A) is clearly visible by the lack of the fucosylated oligosaccharides XXFG and XLFG, with m/z of 1435 and 1597 respectively, present in the OLIMP spectra of the mutant. Out of 240 plants phenotyped, 152 wild type and 57 mutant plants were identified (B). For 31 of the plants the OLIMP spectra was not of sufficient quality to confidently assign a phenotype classification. The approximately 3:1 ratio of wild type to mutant plants is consistent with a single recessive allele as the basis for the *axy2.4* phenotype based on the chi-square test (*p*-value > 0.4).

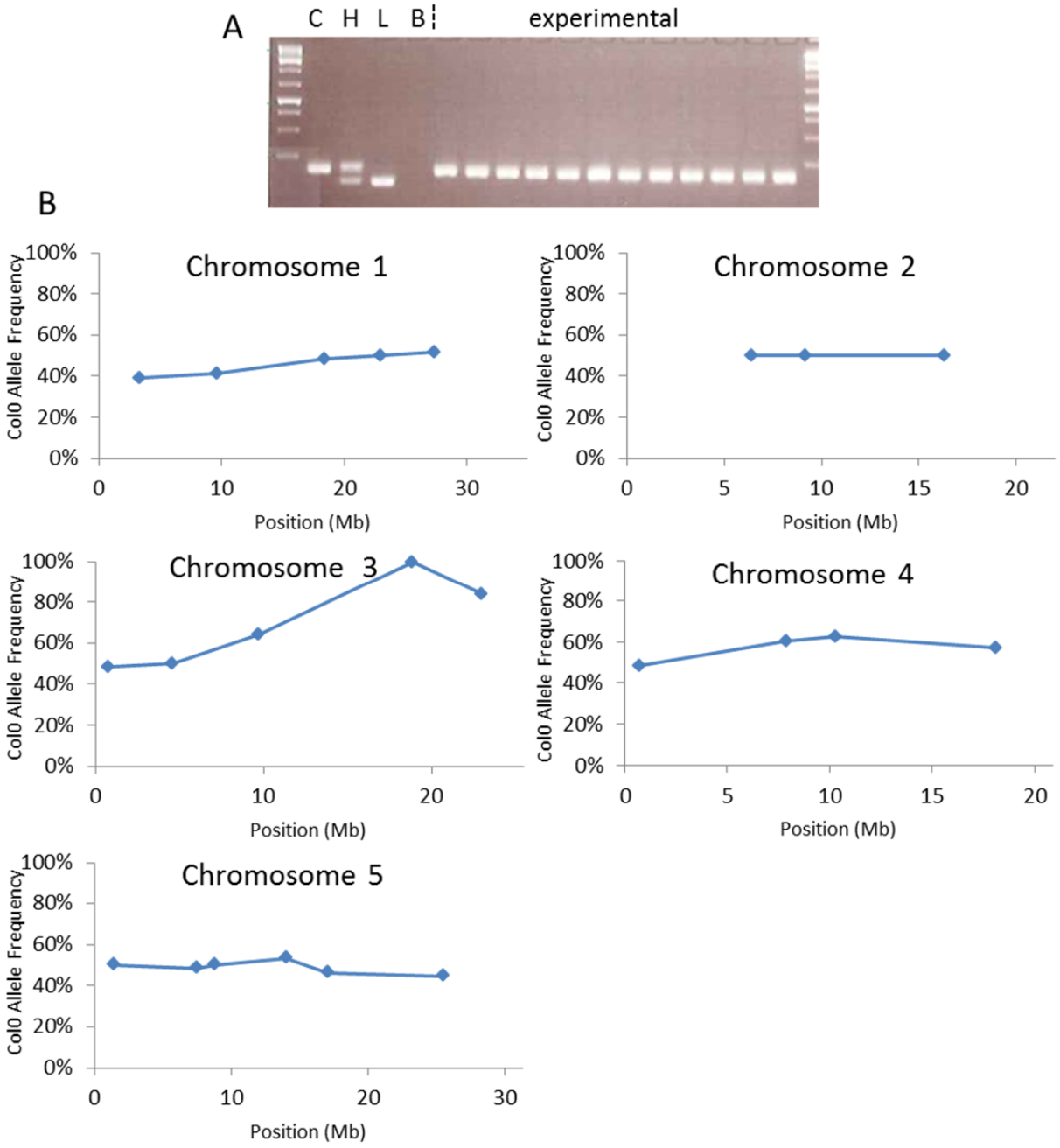


Figure 3-7. Genetic linkage of the *axy2.4* locus. Plants showing a mutant phenotype selected from the F2 of a cross between *axy2.4* and Landsberg erecta (Ler) were genotyped with SSLP markers at between three and six positions along each chromosome. (A) an example genotyping DNA gel showing the PCR products resulting from the use of a simple sequence length polymorphism (SSLP) marker. The first four lanes (excluding the ladder) are Col0 (C), heterozygous Col0/Ler (H), Ler (L), and a water blank (B). The remaining lanes are from PCRs performed with genomic DNA from mutant plants. As only Col0 alleles were detected (the parental ecotype of the *axy* mutants), this marker shows strong linkage with the mutation. (B) Graphs showing the frequency of Col0 alleles for each of the markers along the chromosomes.

Strong genetic linkage was detected for the marker on chromosome 3 at 19 Mb, with all 28 mutant plants showing homozygosity for the Col0 allele.

Marker	CIW11	2fm2	2fm4	2fm5	2fm7	CIW4	2fm8	2fm9	2fm10	2fm11	NGA6
Position (Mb)	9.8	14.2	16.0	17.0	18.7	18.9	19.6	20.9	22.0	23.0	23.0
Percent Col0	64%	80%	85%	93%	98%	100%	96%	89%	88%	84%	84%
Mutant # 1	H	H	H	C	C	C	C	C	C	C	C
2	H	C	C	C	C	C	C	C	C	C	C
3	C	C	C	C	C	C	C	C	C	H	H
4	C	C	C	C	C	C	C	C	C	C	C
5	H	H	H	H	C	C	C	C	C	C	C
6	C	C	C	C	C	C	C	C	C	C	C
7	H	H	C	C	C	C	C	C	C	C	C
8	H	H	H	C	C	C	C	C	H	H	H
9	L	L	H	H	C	C	C	C	C	C	C
10	L	C	C	C	C	C	C	C	C	H	H
11	C	C	C	C	C	C	H	H	H	H	H
12	H	H	C	C	C	C	C	C	C	C	C
13	H	C	C	C	C	C	C	C	C	C	C
14	C	C	C	C	C	C	C	H	H	H	H
15	C	C	C	C	C	C	C	C	C	C	C
16	H	C	?	C	C	C	C	C	C	C	C
17	C	C	C	C	C	C	C	C	C	C	C
18	H	H	H	C	C	C	C	C	C	C	C
19	C	C	C	C	C	C	C	H	H	H	H
20	H	C	C	C	C	C	C	C	C	C	C
21	C	C	C	C	C	C	C	H	H	H	H
22	H	H	H	H	C	C	C	C	C	C	C
23	H	H	H	H	H	C	C	C	C	C	C
24	C	C	C	C	C	C	C	C	C	C	C
25	C	C	C	C	C	C	H	H	H	H	H
26	H	C	C	C	C	C	C	H	H	H	H
27	L	H	H	C	C	C	C	C	C	C	C
28	H	C	C	C	C	C	C	C	C	C	C

Mapped Region

Colombia 0 C Heterozygous H Landsberg erecta L

Figure 3-8. Fine mapping of *axy2.4*. Additional genotyping was performed within the identified region on chromosome 3 using CAPS markers (Table 11-2). The results indicate that the axy2.4 locus is located between 18.7 and 19.6 Mb on chromosome 3.

	XXXG	XXLG	XXFG	XLLG	XXFG-Ac	XLFG-Ac
Col0	0.47[1] ± 0.01	0.10 ± 0.01	0.01 ± 0.02	0.01 ± 0.02	0.13 ± 0.01	0.28 ± 0.03
axy2.4	0.34 ± 0.02	0.45 ± 0.03	n.d.	0.22 ± 0.02	n.d.	n.d.
mur1.1	0.41 ± 0.07	0.42 ± 0.02	n.d.	0.18 ± 0.02	n.d.	n.d.
axy2.4/mur1.1	0.34 ± 0.06	0.44 ± 0.02	n.d.	0.28 ± 0.09	n.d.	n.d.

Table 3-2. Allelism test for *axy2.4* and *mur1.1*. Allelism test for *axy2.4* and *mur1.1*. Leaf XyG OLIMP was performed on homozygous *axy2.4* and *mur1.1* plants as well as F1 plants from a cross between *axy2.4* and *mur1.1 (n ≥ 2)*. Like the single mutants, the F1 (heterozygous) plants were found to lack XyG fucosylation. As both *axy2.4* and *mur1.1* are recessive mutations, this indicates that the *axy2.4* mutation is allelic to *mur1.1*. Value after ± is standard deviation. n.d. – not detected. [1]Relative abundance of ion.

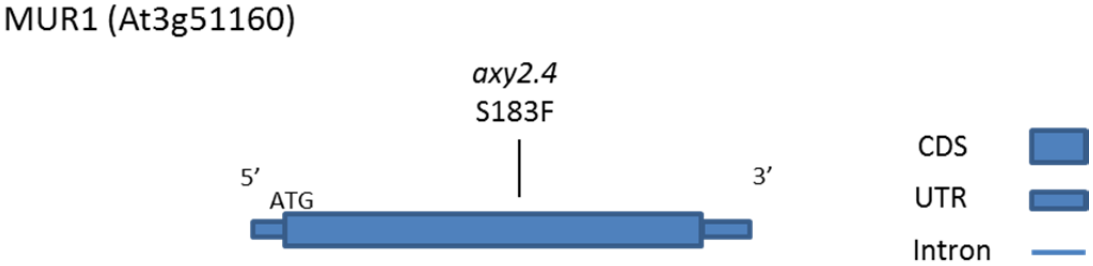


Figure 3-9. *MUR1* gene model. The *MUR1* gene (*At3g51160*) contains no introns and codes for a protein of 373 amino acids. This enzyme, predicted to localize to the cytosol, catalyzes the oxidation of GDP-Mannose, a step required for the de novo synthesis of GDP-Fucose. The *axy2.4* mutant was found to have a mutation within the coding sequence of this gene resulting in the missense mutation S183F. CDS – coding sequence, UTR – untranslated region.

3.2.4 The *axy6* mutant has reduced XyG fucosylation and is distinct from *axy8*

The *axy6* mutant has an OLIMP spectrum that reveals increased XyG fucosylation, as observed by an increase in the relative abundance of XXFG by 17%, but without the unusual oligosaccharides such as GFG or XFG observed in the *axy8* mutants (Figure 3-1, 2.5.3) (Gunl et al. 2011). The AXY8 locus of the *axy6* mutant was sequenced and found to not contain any mutations (2.2.7). Therefore, genetic mapping was pursued to identify the location of the AXY6 locus. Mutants were selected from an F2 population of *axy6* crossed to Ler by performing OLIMP on root tissue (Figure 3-10). Root tissue was used for phenotyping as etiolated seedlings do not provide sufficient material for both OLIMP and genomic DNA extraction and leaf material does not exhibit a XyG phenotype (Figure 13-5). The observed difference between wild type and *axy6* root XyG fucosylation was approximately 15% for the relative abundance of the sum of XXFG and XXFG (Figure 3-10). This is a relatively subtle phenotype given the variation that can be observed between individuals of the same genotype, although analysis of the F2 did show a biased bi-modal distribution indicative of segregation of a recessive trait (Figure 3-11). The mutant and wild type phenotypes of individual plants overlapped indicating that some wild type plants could be misclassified as mutants. Nevertheless, genetic linkage analysis was

performed and an enrichment for Col0 alleles was found on chromosome 1 (Figure 3-12, 2.2.3). Fine mapping indicated the mutation is likely present between 7.4 and 18.4 Mb on chromosome 1 (Figure 3-13). The existence of incorrect mutant picks (e.g. plants 22, 34, 80, and 127) in the mapping population made it difficult to narrow down the region. Illumina sequencing was performed on plants homozygous for the *axy6* mutation to identify candidate SNPs caused by the EMS treatment (2.2.15). Within the mapped region a total of 37 SNPs within genes (including non-coding sequences) were identified, 16 of which result in predicted amino acid changes (Table 3-3, Table 3-4, 2.4.1). One of these genes (XTR4/XTH30, At1g32170) represents a putative XyG transglycosylase/hydrolase (XTH) and, since such enzymes affect XyG structure, is a promising candidate for genetic basis of the *axy6* mutant. Future work will involve checking this and other candidate genes by phenotyping T-DNA lines and genetic complementation to determine the causative mutation for the *axy6* XyG phenotype.

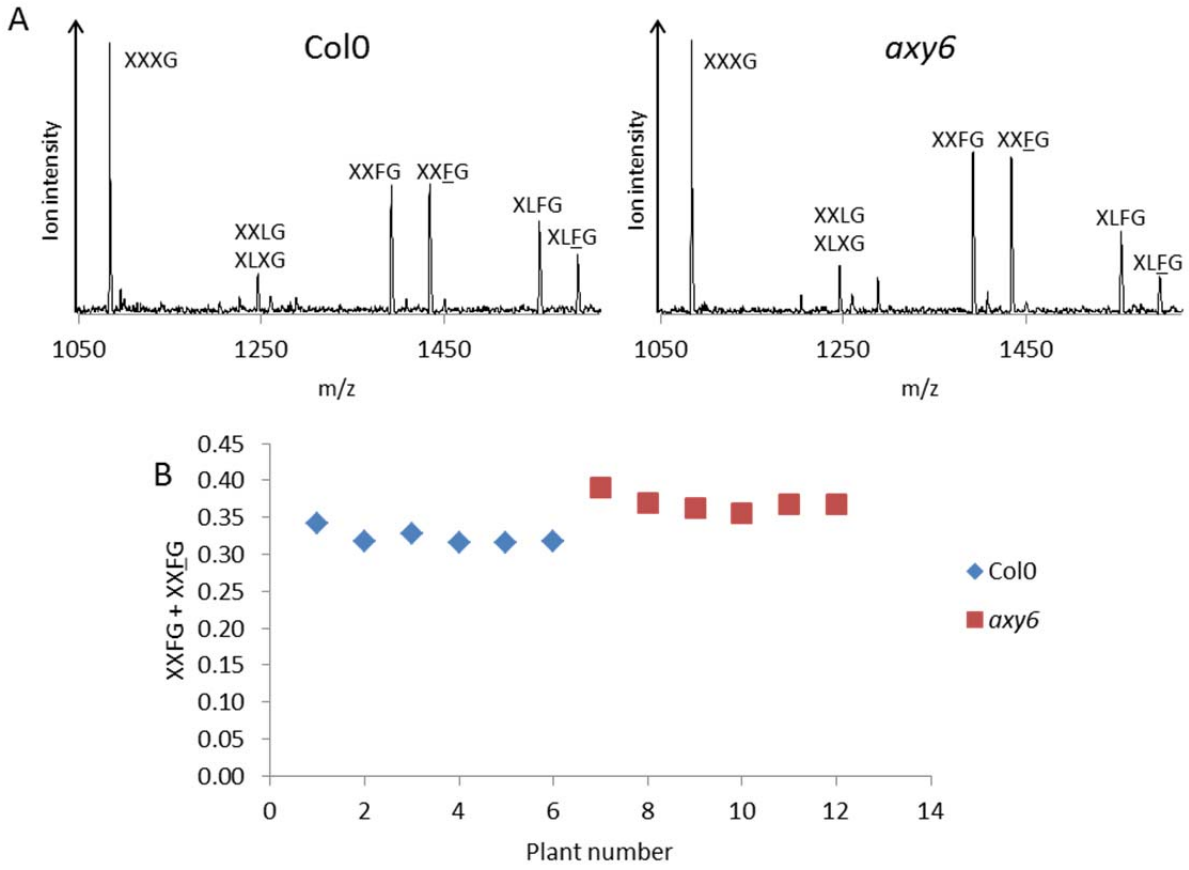


Figure 3-10. XyG OLIMP of *axy6* root tissue. (A) Representative spectra and (B) the relative sum of the abundance of XXFG and XXFG (Y-axis) for six individual plants (spread along X-axis) of each genotype are shown. The *axy6* mutant lacks a leaf OLIMP phenotype (Figure 13-5) but has a root phenotype as seen by the increased relative abundance of XXFG and XXFG in the mutant.

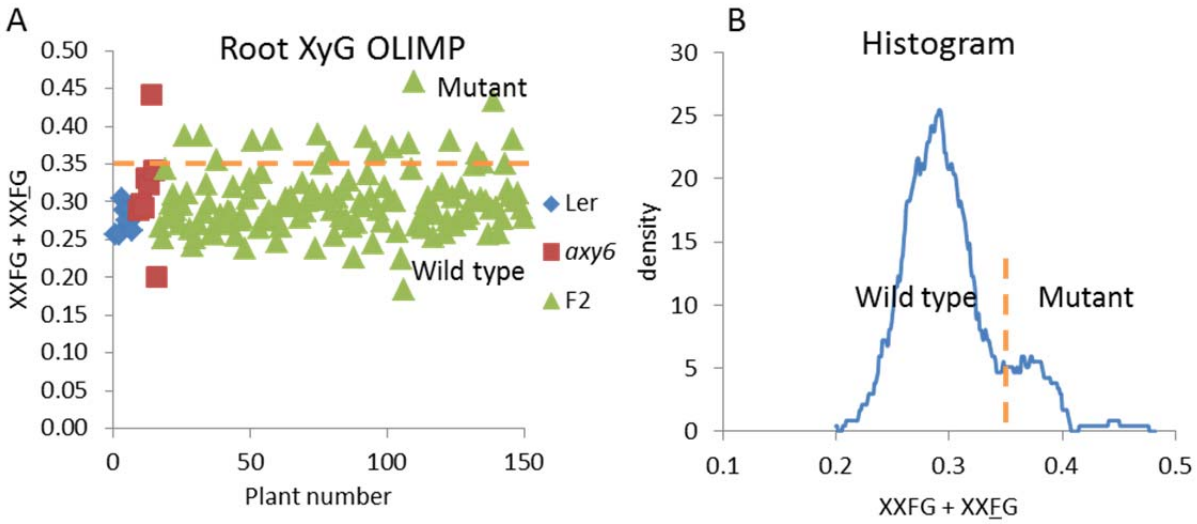


Figure 3-11. Mutant selection for genetic mapping of *axy6*. OLIMP was performed on root material from F2 plants of *axy6* crossed to Landsberg erecta (Ler). (A) Individual plants are plotted with the relative amount of fucosylation (XXFG + XXFG) shown on the Y-axis. Though the *axy6* plants average increased fucosylation, some individual spectra overlap with those from Ler wild type. (B) A histogram of the data presented in (A). Despite the overlap between mutant and wild type phenotypes, segregation was observed based on the biased bi-modal distribution observed in the histogram. Mutants were selected based on the cutoff shown by the orange dashed line.

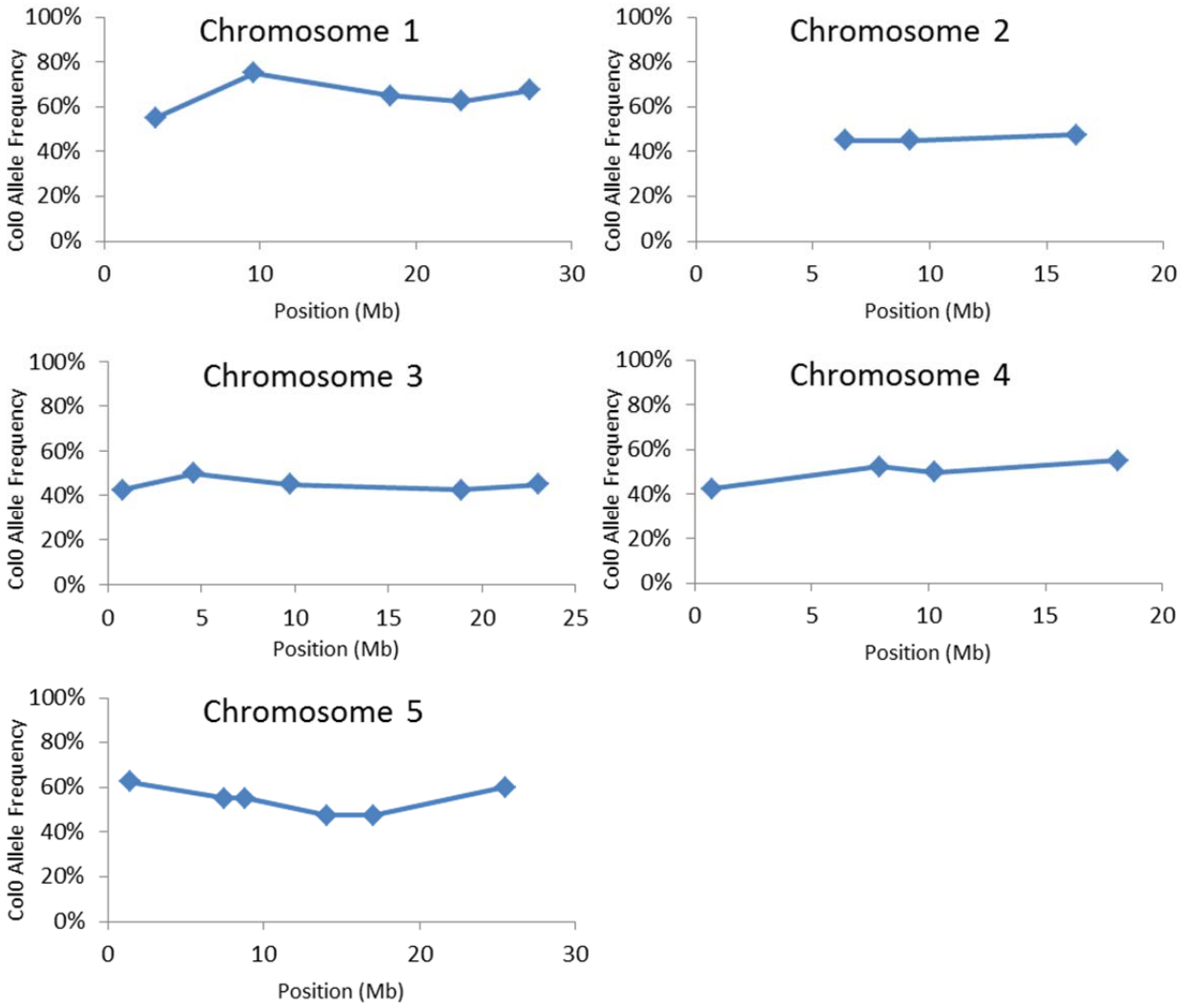


Figure 3-12. Genetic linkage analysis of the *axy6* locus. Plants showing a mutant phenotype from the F2 of a cross between *axy6* and Landsberg erecta were genotyped with SSLP markers at between three and six positions along each chromosome. An increased frequency of Col0 alleles (the parental ecotype of the AXY mutants) was found on chromosome 1 near 10 Mb (*p*-value < 0.002).

Marker		T27I1	6fm1	6fm2	6fm4	CIW12	6fm6	6fm7	6fm10	6fm12	6fm13	CIW1
Position (Mb)		3.3	6.1	7.4	9.6	9.6	11.3	12.9	13.3	15.9	16.8	18.4
Percent Col0		55%	65%	65%	75%	75%	75%	75%	75%	78%	82%	65%
Plant #	10	H	C	C	C	C	C	C	C	C	C	C
	16	C	C	C	C	C	C	C	?	?	C	L
	22	H	H	H	H	H	H	H	H	H	H	H
	34	L	L	L	L	L	L	H	H	H	H	H
	35	L	H	H	C	C	C	C	C	C	C	C
	42	H	H	H	C	C	C	C	C	C	C	C
	59	C	C	C	C	C	C	C	C	C	C	H
	60	H	H	H	C	C	C	C	C	C	C	C
	63	H	C	C	C	C	C	C	C	C	C	H
	76	C	C	C	C	C	C	H	H	H	H	H
	77	L	L	L	L	L	L	H	?	?	?	C
	80	L	L	L	L	L	L	L	L	L	L	L
	86	H	H	H	C	C	C	C	H	H	C	H
	93	H	H	H	H	H	L	L	H	C	C	C
	107	C	C	C	C	C	C	C	C	C	C	H
	116	C	C	C	C	C	C	C	C	C	C	H
	117	C	C	C	C	C	C	C	C	C	C	C
	119	H	H	H	H	H	C	H	H	H	H	H
	127	L	H	H	H	H	H	H	H	H	H	H
	130	C	C	C	C	C	C	C	C	C	C	C

Mapped Region

Colombia 0 C Heterozygous H Landsberg erecta L

Figure 3-13. Fine mapping of *axy6*. CAPS markers were identified within the mapped region and used to genotype axy6 mutant plants. The occurrence of Ler alleles throughout the mapped region in some of the plants suggests a high likelihood of incorrect mutant picks (such as plants 22, 34, 80, and 127). When eliminating these incorrect mutant picks the mutation appears to be present in a region spanning from 7.4 and 18.4 Mb on chromosome 1. Failed genotyping reactions are indicated by ?.

Criteria (additive)	# Candidate SNPs
Total SNPs	8754
Unique to *axy6*	1922
Homozygous	1038
Causes AA change	210
In mapped region (Chr1 7.4-18.4 Mb)	16

Table 3-3. Genomic sequencing of *axy6*. Illumina sequencing was performed on homozygous *axy6* mutant plants (not backcrossed) to identify candidate mutations. Additive criteria were used to filter the identified mutations to homozygous mutations unique to *axy6* located within the mapped region and resulting in an amino acid change. There were sixteen such mutations.

Position (bp)	Gene	AA Change	Gene Annotation
7,375,359	At1g21065	Trp193Stop	unknown protein
7,598,258	At1g21650	Pro354Ser	Encodes a component of the thylakoid-localized Sec system
7,608,800	At1g21660	Ala513Thr	Chaperone DnaJ-domain superfamily protein
8,143,641	At1g23000		Heavy metal transport/detoxification superfamily protein
8,281,494	At1g23330	Gly108Arg	alpha/beta-Hydrolases superfamily protein
8,462,394	At1g23935		CONTAINS InterPro DOMAIN/s: Apoptosis inhibitory 5 (InterPro:IPR008383)
8,687,155	At1g24510	Val135Ile	TCP-1/cpn60 chaperonin family protein
9,513,871	At1g27390		Form of TOM20 involved in transport of nuclear-encoded mitochondrial proteins
9,598,116	At1g27595		CONTAINS InterPro DOMAIN/s: Symplekin tight junction protein C-terminal
9,843,186	At1g28180		P-loop containing nucleoside triphosphate hydrolases superfamily protein
10,030,689	At1g28520	Pro254Ser	vascular plant one zinc finger protein (VOZ1)
10,211,374	At1g29220		transcriptional regulator family protein
10,355,777	At1g29640		Protein of unknown function, DUF584
10,399,160	At1g29724		
10,781,197	At1g30470		SIT4 phosphatase-associated family protein
11,553,095	At1g32120	Ala57Val	FUNCTIONS IN: molecular_function unknown
11,577,603	**At1g32170**	**Pro287Ser**	**xyloglucan endotransglycosylase-related protein (XTR4)**
11,847,852	At1g32750	Ser1476Asn	This gene is predicted to encode a histone acetyltransferase.
11,976,387	At1g33060		NAC 014 (NAC014)
12,521,587	At1g34320		Protein of unknown function (DUF668)
12,960,728	At1g35320		unknown protein
13,332,170	At1g35860		TOC75 pseudogene due to a retrotransposon inserted at 5' end of the gene
13,456,265	At1g36060		member of the DREB subfamily A-6 of ERF/AP2 transcription factor family.
13,616,139	At1g36240		Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein
13,917,577	At1g36756	Pro51Ser	unknown protein
16,894,573	At1g44750	Ala182Val	Member of a family of proteins related to PUP1, a purine transporter.
16,981,246	At1g44910	Ser657Phe	Functions as a scaffold for RNA processing machineries.
17,156,445	At1g45229		unknown protein
17,284,462	At1g47128	Gly203Ser	cysteine proteinase precursor-like protein/dehydration stress-responsive gene (RD21).
17,508,747	At1g47620	Arg311Lys	member of CYP96A
17,724,126	At1g48050		Required for proper maintenance of the telomeric C strand.
18,055,321	At1g48820		Terpenoid cyclases/Protein prenyltransferases superfamily protein
18,147,366	At1g49040	Asp162Asn	Encodes soluble protein involved in cytokinesis of leaf cells.
18,177,300	At1g49140		Complex I subunit NDUFS6

18,283,242	At1g49400		embryo defective 1129 (emb1129)
18,329,808	At1g49520	Pro25Ser	SWIB complex BAF60b domain-containing protein
18,382,088	At1g49670		Involved in oxidative stress tolerance.

Table 3-4. *axy6* candidate genes. A total of 37 homozygous SNPs within genes in the mapped region with the amino acid change (if present), gene, and annotated gene function listed. Of particular note is the XTR4 gene (bold), which is a member of the XyG Transglycosylase / Hydrolase gene family.

3.2.5 Identification of the genetic basis of the *axy9.1* mutant

The *axy9.1* mutant was identified in the AXY screen as having a 60% decrease in XyG acetylation (Figure 3-1, 2.5.3). As mutants defective for AXY4 also have decreased acetylation, ranging from a 20% reduction to a complete absence depending on the allele (Gille et al. 2011b), the AXY4 locus was sequenced and an allelism test was performed. The *axy9.1* mutant was found not to have any mutations in the AXY4 gene and the *axy9.1*/*axy4.3* F1 plants had XyG acetylation levels near wild type indicating they are not allelic as both *axy9* and *axy4* are recessive mutations (Figure 3-15, Figure 3-16) (Gille et al. 2011b). Another protein involved in wall polysaccharide *O*-acetylation is RWA, of which there are four paralogs in *A. thaliana* (Lee et al. 2011; Manabe et al. 2011). The strength of the *axy9.1* mutant, having a 60% decrease in XyG acetylation in etiolated seedlings, is significantly stronger than the phenotype of any of the single RWA mutants, which also have decreased XyG acetylation (Sascha Gille personal communication) (Manabe et al. 2011). For these reasons, it was believed that *axy9.1* was not allelic to any of the RWAs. The *axy9* mutant was crossed to Ler for genetic mapping and mutants were selected from the F2 population with a segregation ratio indicative of a single recessive locus being responsible for the *axy9* phenotype (Figure 3-16). Rough mapping showed a strong enrichment for Col0 alleles at the beginning of chromosome 3 (Figure 3-17, 2.2.3). The mutation was fine mapped using CAPS markers to the first 1.6 Mb of this chromosome (Figure 3-18, Table 11-2, 2.2.3).

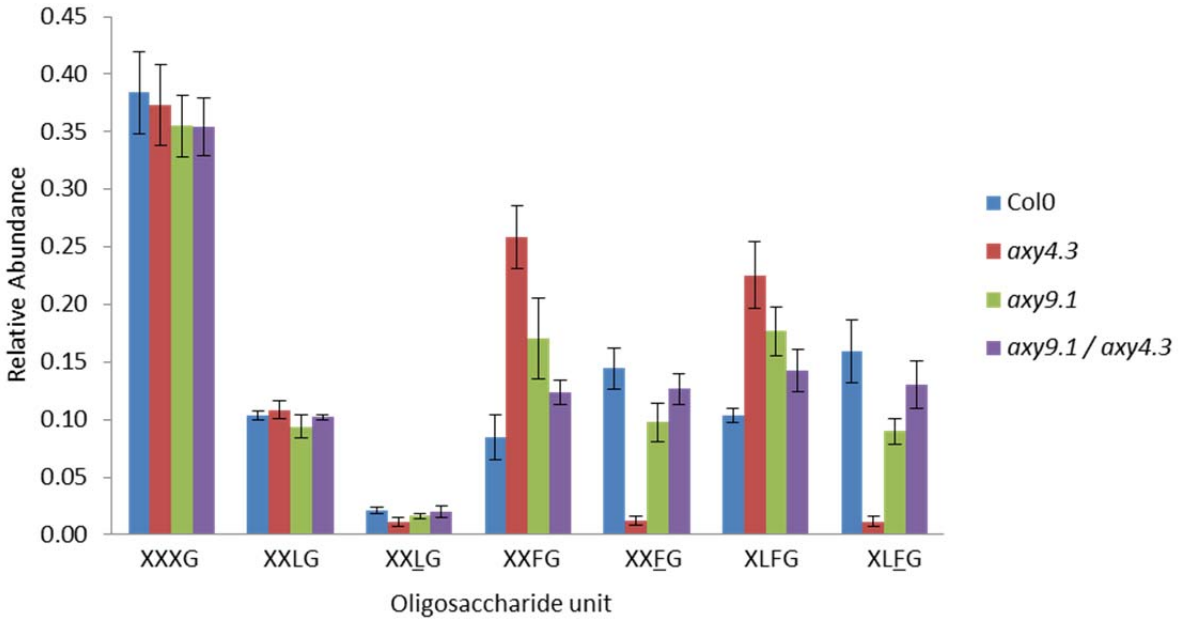
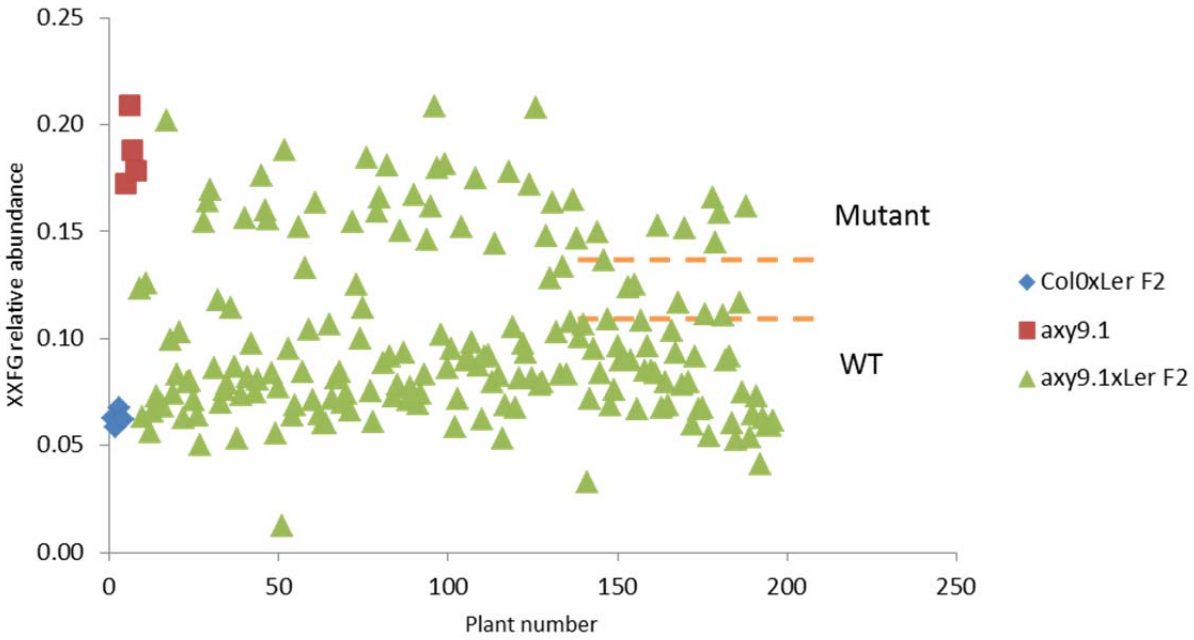


Figure 3-14. Allelism test between *axy9.1* and *axy4*. Leaf OLIMP was performed on wild type (Col0), the *axy4.3* and *axy9.1* homozygous single mutants, and the *axy9.1* / *axy4.3* heterozygous F1. The heterozygous plants were found to have a near-wild type acetylation level, indicating that *axy9.1* is not allelic to *axy4*.

A



B

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

Figure 3-15. Selection of mutants for mapping of *axy9.1*. Leaf OLIMP was used to phenotype F2 plants from a cross between axy9.1 and the ecotype Landsberg erecta (Ler). The F2 population was observed to be segregating (A) and mutants were selected based on the relative abundance of the non-acetylated oligosaccharide XXFG. Of 187 plants phenotyped, 41 mutant and 133 wild type plants were identified (B). This segregation ratio is consistent with that expected for a single recessive locus based on the chi-square test (*p*-value > 0.60).

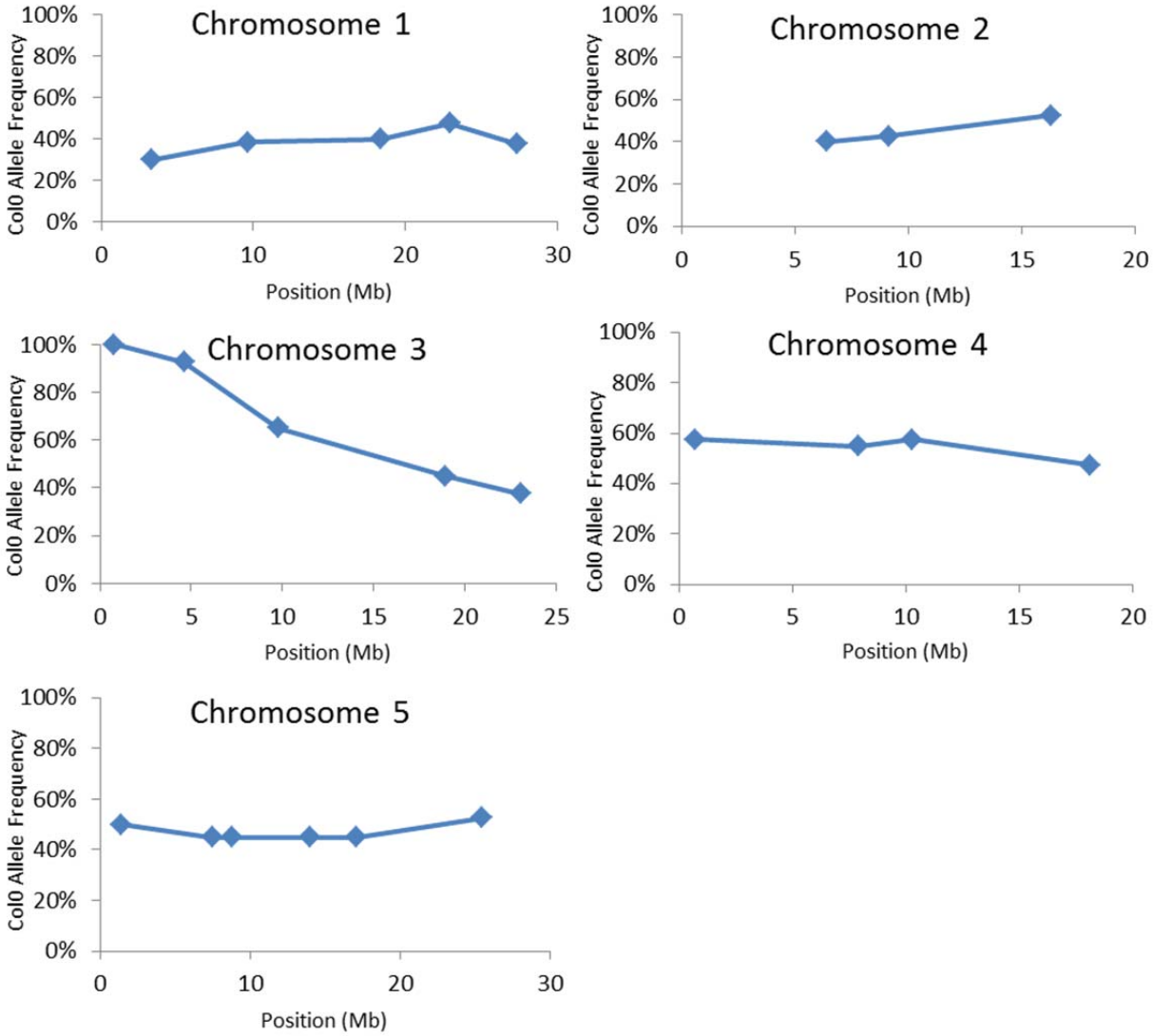


Figure 3-16. Genetic linkage analysis of the *axy9* locus. Using between three and six SSLP markers along each chromosome, the selected mutant plants from the *axy9.1* x Ler F2 population were genotyped. The frequency of Col0 alleles (the parental ecotype of the AXY mutants) revealed a strong genetic linkage near the 1 Mb position on chromosome 3.

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

Col0 C Heterozygous H Landsberg erecta L

Figure 3-17. Fine mapping of *axy9.1*. Fine mapping of *axy9.1*. CAPS markers were used to gain greater resolution of the mapped region. The genotyping result is shown for the mutant plants at each of the markers indicated. These results indicate that the *axy9.1* mutation maps to the first 1.6 Mb of chromosome 3.

Illumina sequencing (100 bp single-end reads, 1 lane HiSeq 2000) was performed on homozygous *axy9.1* mutant plants (not backcrossed) (2.2.15). The reads were assembled and an average coverage of 100X was achieved for the genomic DNA. Of nearly 9,000 total SNPs detected there were 652 homozygous ones unique to the *axy9.1* mutant (not found in other *axy* mutants) (2.4.1). 148 of these mutations resulted in amino acid changes, two of which were located within the mapped region (Table 3-5, Table 3-6). Homozygous T-DNA lines were obtained for both genes and showed normal XyG acetylation (Figure 13-4) except for a line containing an insertion in At3g03210, which showed approximately a 75% decrease in XyG acetylation in etiolated seedlings (Figure 3-18, 2.2.2). This line, SALK_090612, was named *axy9.2*. The *axy9.1* mutant has a mutation in this gene, which introduces a stop codon at amino acid position 276 (Figure 3-19, Appendix 1, 2.2.7). Hence, strong evidence is presented that the observed decrease in XyG acetylation is due to mutations in the locus At3g03210.

Criteria (additive)	# Candidate SNPs
Total SNPs	8827
Unique to axy9	1987
Homozygous	652
Causes AA change	148
In mapped region (Chr1 0-1.5 Mb)	2

Table 3-5. Genomic sequencing of *axy9.1*. Illumina sequencing was performed on the *axy9.1* mutant to identify candidate mutations for the causation of the *axy9.1* phenotype. Two homozygous amino acid-changing SNPs were identified within the mapped region.

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

Table 3-6. Candidate SNPs for the basis of the *axy9.1* phenotype. Eight mutations within genes in the mapped region were identified in the *axy9.1* mutant, two of which result in amino acid (AA) changes.

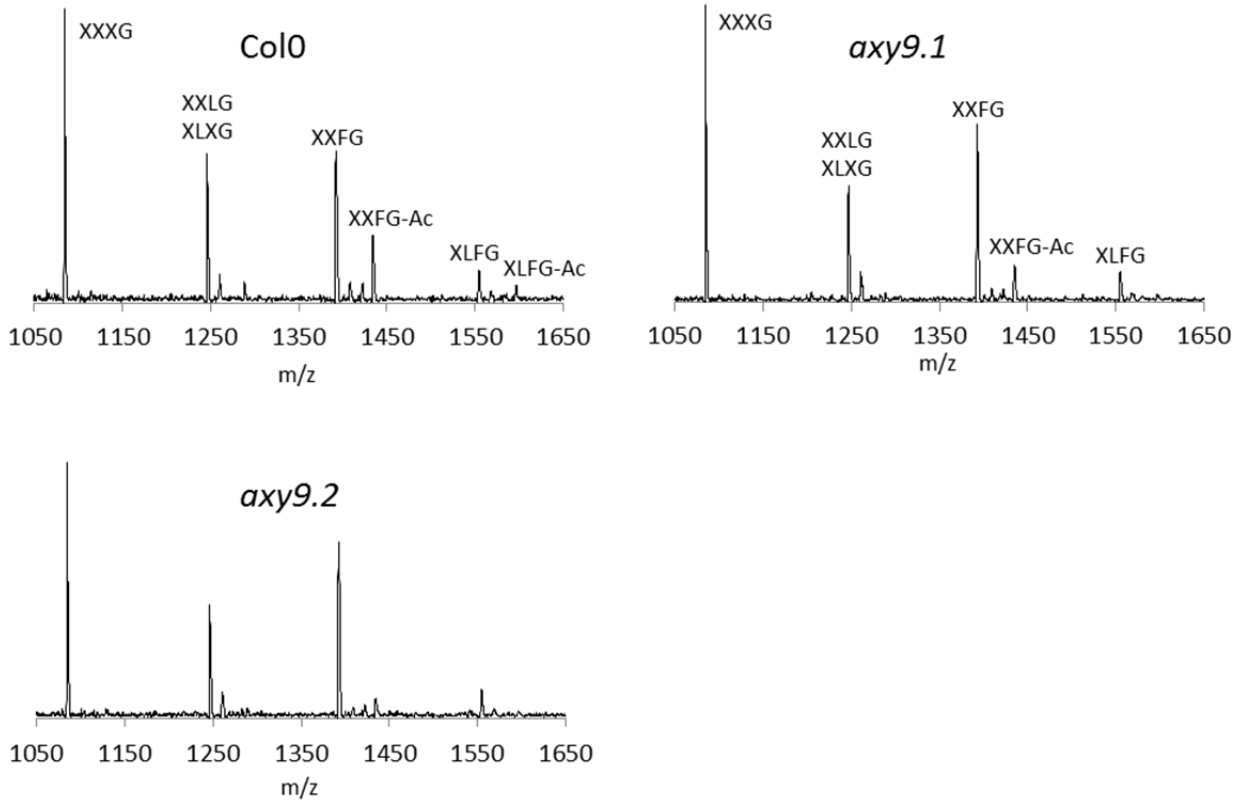


Figure 3-18. XyG OLIMP of *axy9* etiolated seedlings. A T-DNA insertional mutant line for the candidate gene At3g03210 had decreased XyG acetylation as revealed by OLIMP. This line, SALK_090612, was named *axy9.2*.

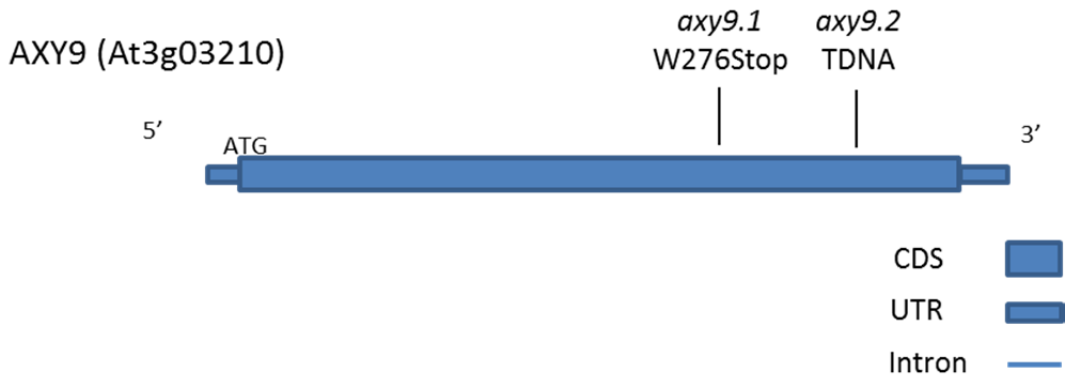


Figure 3-19. *AXY9* gene model. The *AXY9* (*At3g03210*) gene has no introns and codes for a protein with 368 amino acids. The *axy9.1* mutant has a mutation that results in an early stop codon at the 276th codon. The *axy9.2* mutant (SALK_090612) contains a T-DNA insertion at the 284th codon (Appendix 1). CDS – coding sequence, UTR – untranslated region.

A summary of all *axy* mutants describe in this work is presented in Table 3-7.

Name	Pool	Phenotype	Causative Mutation	Evidence
	236-81	*axy3*	AXY3 W468Stop	alleles[1]
axy1.160	160-55	*axy3*	AXY3 W482Stop	alleles
axy1.3		*axy3*	AXY3 D529N	alleles, complementation
	148-100	*axy3*	AXY3 E630K	alleles
	148-69	*axy8*	AXY8 W95Stop	alleles
	142-71	*axy8*	AXY8 W441Stop	alleles
	129-16	*axy8*	AXY8 splice disruption	alleles
	8-52	*mur3*	MUR3 C400Y	alleles
axy2.4		*mur1*	MUR1 S183F	alleles, complementation, mapping
axy6	14-99	↑ fucosylation	unknown	
axy9	179-83	↓ acetylation	At3g03210 W276Stop	alleles, complementation, mapping

Table 3-7. Summary of *axy* mutants characterized in this work. An altered XyG phenotype was confirmed for eleven of the *A. thaliana* mutants. Some of these mutants were found to be allelic to previously discovered mutants affecting XyG structure including *axy3*, *axy8*, *mur1* and *mur3*. The locus responsible for the *axy6* phenotype (increased fucosylation) has not yet been identified. The *axy9* phenotype was found to be caused by a mutation in the gene *At3g03210*. The evidence column indicates the experimental evidence that the mutant phenotype is caused by the specified mutation. [1]indicates that plants with independent mutations in the specified gene have the same phenotype. [2]supported by genetic or transgenic complementation analysis. [3]the mutant was mapped to a region containing the specified gene.

3.2.6 Characterization of the *axy9* mutants

3.2.6.1 The AXY9 gene

A BLAST search using the AXY9 protein sequence as a query revealed that there are no close paralogs of AXY9 within the *A. thaliana* genome (www.arabidopsis.org). The most significant match, with an e-value of 0.004, is the protein TBL1. The TBL proteins are putative polysaccharide *O*-acetyltransferases (Gille et al. 2011b; Xiong et al. 2013; Yuan et al. 2013). With such a high e-value, AXY9 is not a member of the TBL family (Bischoff 2010) but does share some sequence homology – notably both AXY9 and the TBLs have a conserved GDS motif, believed to be important for acetyl esterase/transferase activity in this superfamily of proteins (Akoh et al. 2004). Whereas the TBL proteins have a single N-terminal hydrophobic domain, the AXY9 proteins have two N-terminal hydrophobic domains which are predicted to be transmembrane domains. This region of the AXY9 proteins is highly conserved among land plants (Figure 13-6). At least one putative ortholog was identified in all land plants queried using Phytozome (www.phytozome.net), with some species having multiple copies. Homologs of AXY9 can be identified in species of green algae but these proteins lack the highly conserved N-terminal domain, suggesting that these proteins have a different activity or mode of action than AXY9 and are likely not orthologous in function.

3.2.6.2 Morphological phenotype of the *axy9.2* mutant

While etiolated seedlings of the *axy9.2* mutant appear normal compared to *axy9.1* and wild type seedlings, a severe phenotype can be observed for plants grown on soil (Figure 3-20, 2.1.4). At 12 days, the *axy9.2* plants appear smaller and are darker green in color. The phenotype becomes more apparent in older plants with delayed bolting, smaller bolts, and poor or no seed set. This phenotype is rescued by complementation of the *axy9.2* mutant with overexpression of the AXY9 protein using the enTCUP2 promoter (Figure 3-20, Figure 10-7, 2.2.5, 2.2.11) (Malik et al. 2002). In addition, the xylem tissue of the *axy9* plants was examined. While Col0 and *axy9.1* plants have large and open xylem vessels, *axy9.2* xylem vessels appear collapsed (Figure 3-21, 2.3.2).

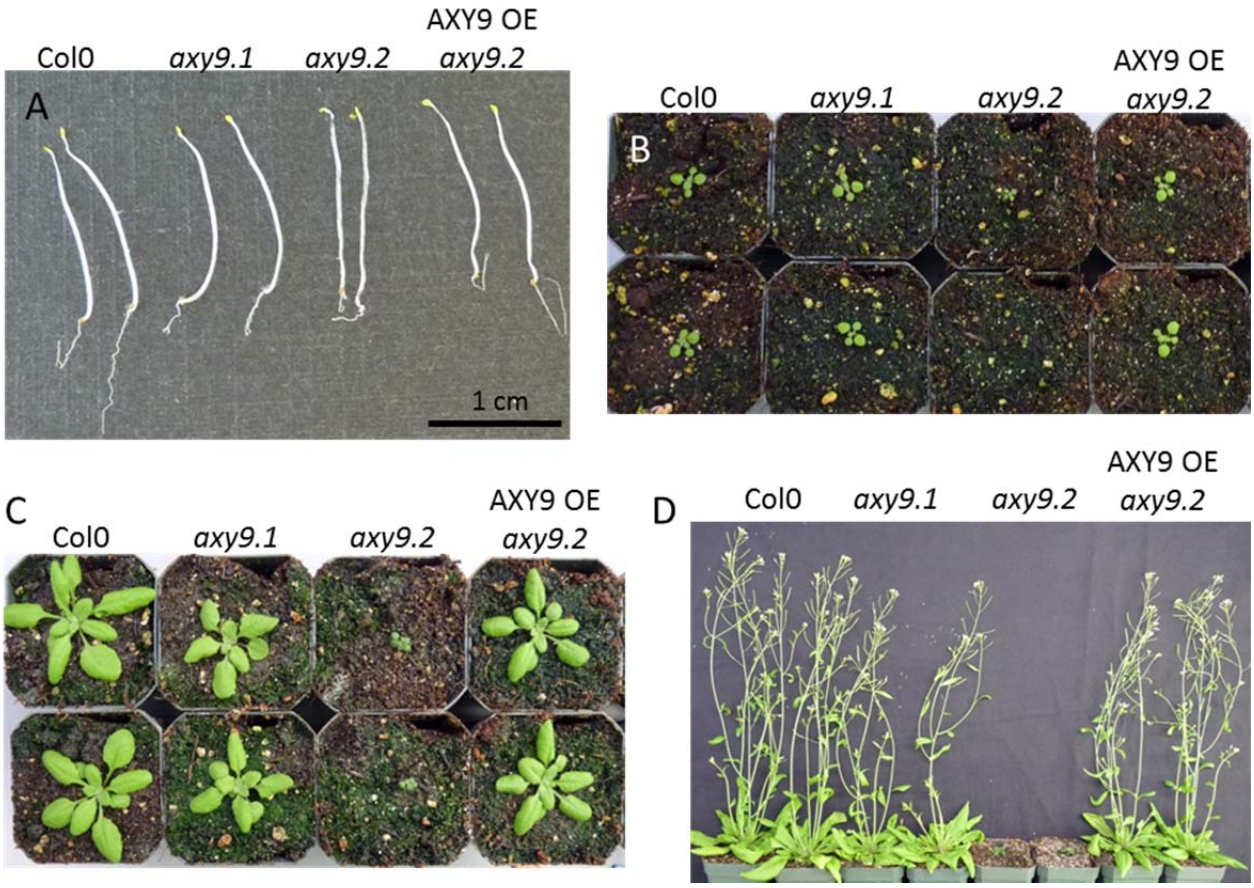


Figure 3-20. Morphological phenotype of the *axy9.2* mutant. Images of wild type (Col0), *axy9.1*, *axy9.2* and AXY9 overexpression in the *axy9.2* background (AXY9 OE) plants a various developmental stages. (A) etiolated seedlings grown for one week in the dark on plates, (B) 12-day old plants grown on soil, (C) three-week old plants grown on soil, (D) five-week old plants grown on soil. Two representative plants are shown for each genotype.

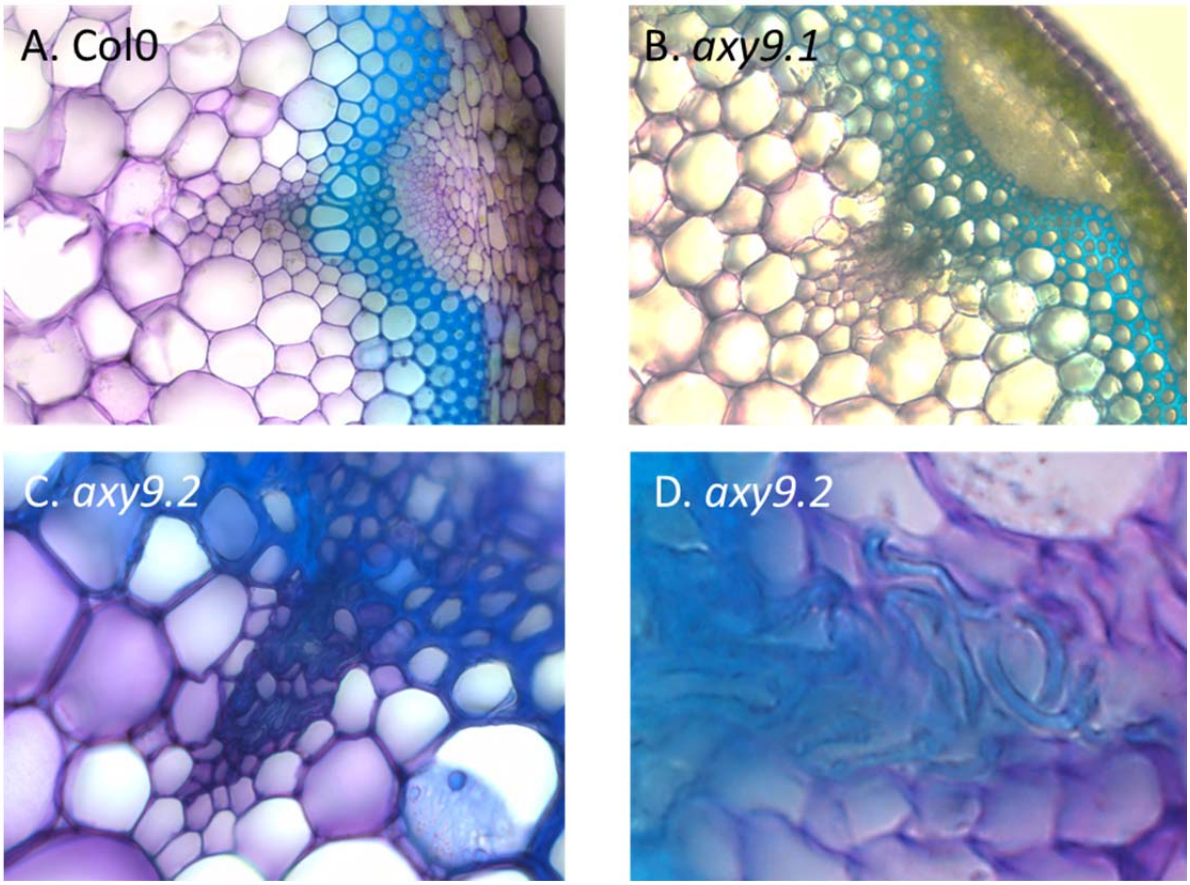


Figure 3-21. Vascular tissue of *axy9.2*. Stem cross sections from Col0 (A), *axy9.1* (B) and *axy9.2* (C, D) were stained with toluidine blue and examine with a light microscope. The xylem tissue of the wild type and *axy9.1* plants has vessels that are large and open whereas these in the mutant the xylem vessels appear collapsed.

3.2.6.3 *O*-Acetylation phenotypes of the *axy9* mutants

XyG OLIMP was performed on etiolated seedling and cotyledon tissue (Figure 3-22, 2.5.3). The *axy9.1* and *axy9.2* plants both had decreases in total percent XyG acetylation compared to wild type in all tissues examined. This phenotype was significantly stronger in *axy9.2* than in *axy9.1* in etiolated seedlings (75% and 60% decrease, respectively). To assess the impact of disruption of the AXY9 gene on the acetylation of other polysaccharides in the wall, total acetyl ester content of AIR material was measured for stem and leaf tissue (2.5.10). The *axy9.1* mutant was found to have a 35% decrease in stem acetyl ester content but no significant difference was observed in leaf material (Figure 3-23). As the acetylation status of XyG has a relatively minor contribution to the total wall acetyl ester content (Gille et al. 2011b), this suggested that other wall polysaccharides are affected in stem tissue. NMR was used to investigate the composition of the stem cell wall, particularly related to the decrease in total acetyl ester content observed (Cheng et al. 2013). NMR is a useful method because it can detect the relative amounts of many different chemical linkages and functional groups from both polysaccharides and phenolics

present in the wall (Cheng et al. 2013). This method revealed a decrease in acetylated xylan (28%) and mannan (7%) for the *axy9.1* mutant compared to wild type (Figure 3-24) (performed by Kun Cheng).

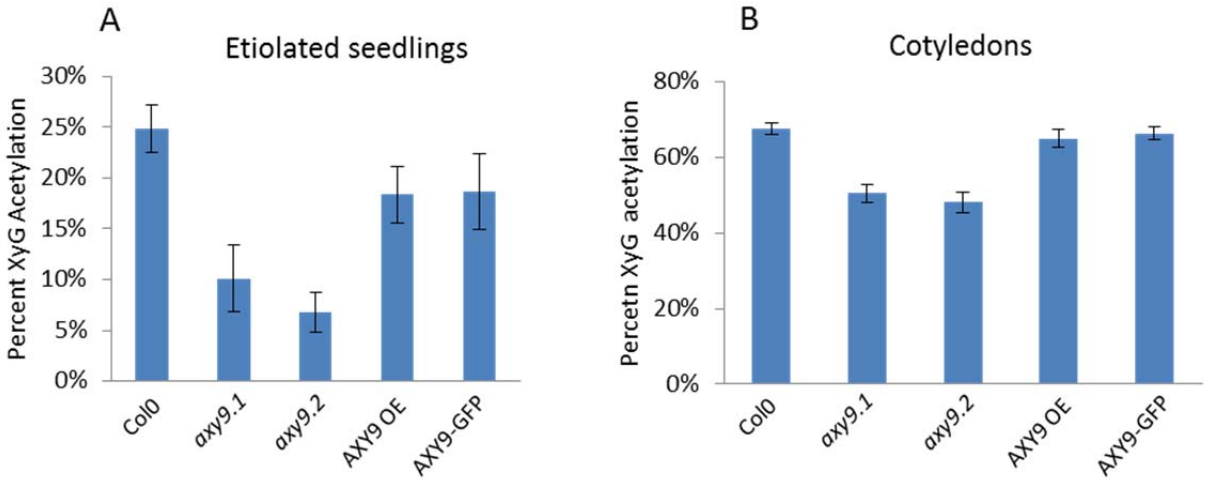


Figure 3-22. XyG OLIMP phenotypes and complementation lines of *axy9*. XyG OLIMP was performed on etiolated seedlings (A) and 12-day old cotyledons (B) for Col0, *axy9.1*, *axy9.2*, and *axy9.2* complemented with AXY9 over-expression and AXY9-GFP (Figure 10-7, Figure 10-4). The *axy9.2* mutant has a stronger phenotype than *axy9.1* in etiolated seedlings. The complementation lines show a partial (etiolated seedlings) or full (cotyledons) rescue. Error bars indicate standard deviation. n ≥ 6.

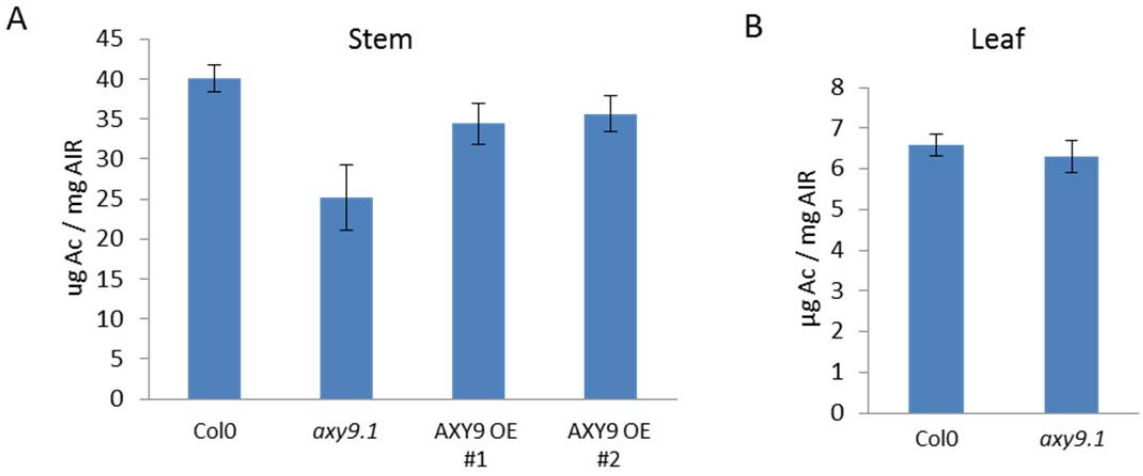
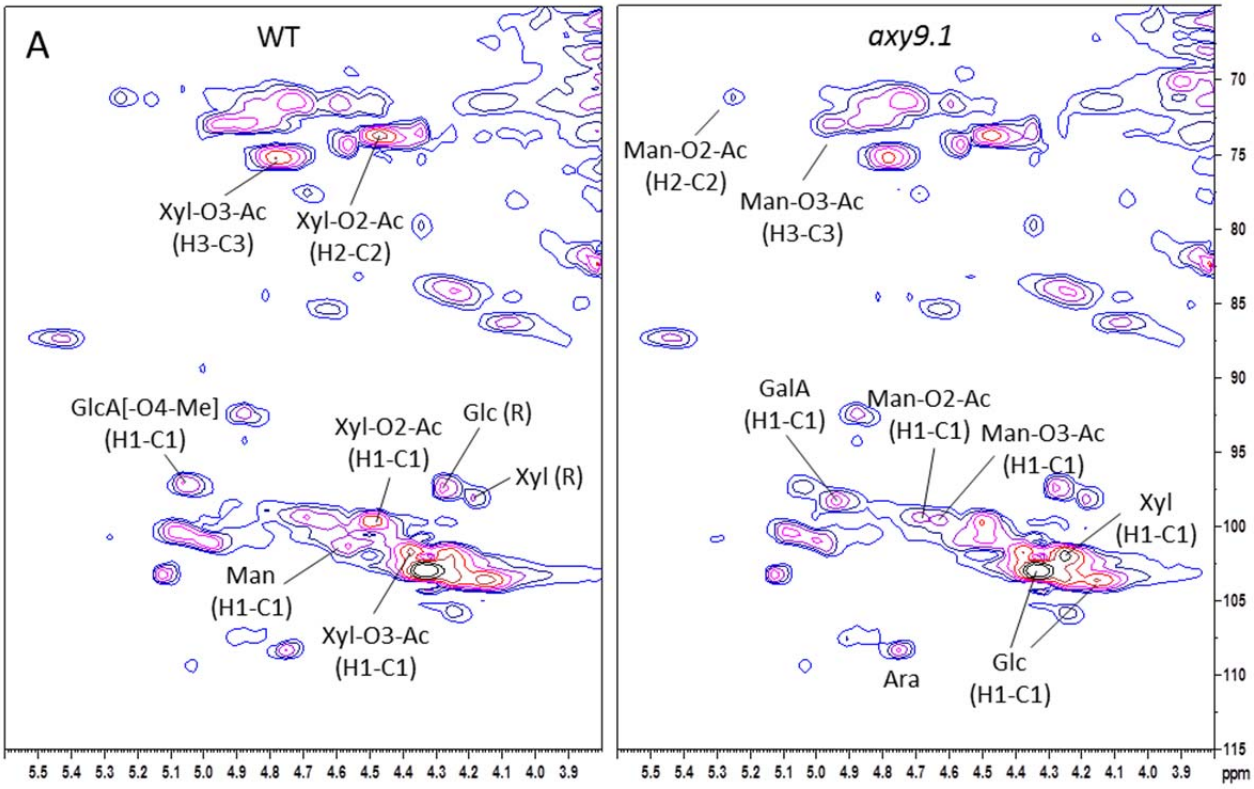


Figure 3-23. Total acetyl-ester content of *axy9* wall material. Acetyl-esters were released by base treatment and the amount of resulting acetate was measured. Values are plotted as µg acetate per mg AIR material for stem (A) and leaf (B) tissue. Two independent

complementation lines each for AXY9 over-expression showed a partial rescue of the stem phenotype. Error bars indicate standard deviation. n = 3.



B

	WT	*axy9.1*
Xylan		
O2-Ac	28%	16%
O3-Ac	31%	28%
Total	60%	43%
Mannan		
O2-Ac	27%	23%
O3-Ac	30%	29%
Total	57%	53%

Figure 3-24. 2D HSQC NMR spectrum of *axy9* stem tissue. (A) The ^{1}H (x-axis) / ^{13}C (y-axis) NMR spectra for wild type and the *axy9.1* mutant. Prominent peaks corresponding to known polysaccharide linkages are labeled. (B) quantification of xylan and mannan acetylation based on NMR peak volume. The *axy9.1* mutant showed a decrease in xylan and mannan acetylation compared to the wild type. NMR analysis performed by Kun Cheng.

3.2.6.4 Subcellular localization of the AXY9 protein

A plasmid was constructed containing the AXY9 coding sequence with a C-terminal GFP fusion for expression in *N. benthamiana* and *A. thaliana* (Figure 10-4, 2.2.5). Infiltration of the AXY9-GFP construct in *N. benthamiana* resulted in a GFP signal observed in small mobile vesicles (Figure 3-25, 2.3.1). This GFP signal colocalized with a Golgi-mCherry marker (targeted by a peptide from a soybean mannosidase, Nelson 2007), when coinfiltrated. Expression of AXY9-GFP in the *axy9.2* mutant rescued the XyG acetylation defect, indicating that the construct of the chimeric protein is functional and likely correctly localized (Figure 3-22, 2.2.11). In these transgenic *A. thaliana* plants, the GFP signal again localized to small mobile vesicles consistent with the Golgi apparatus (Figure 3-26).

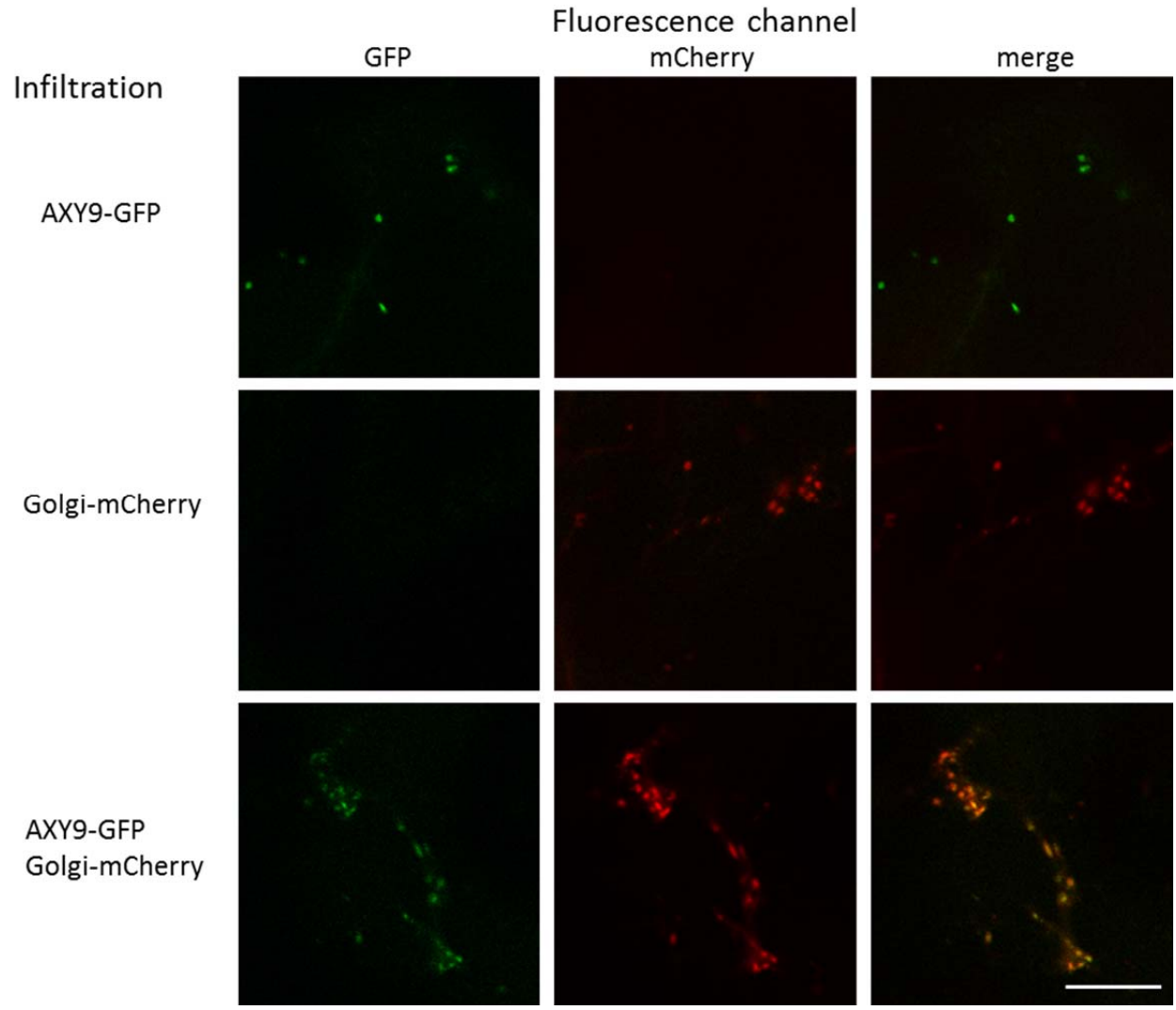


Figure 3-25. Subcellular localization of AXY9-GFP in *N. benthamiana*. A fusion construct of AXY9-GFP was infiltrated for transient expression in *N. benthamiana* with (bottom row) or without (top row) a Golgi-mCherry marker construct. A confocal microscope was used to view the GFP (left column) and mCherry (middle column) fluorescence specifically. The AXY9-GFP fluorescence signal colocalizes with that of the Golgi marker. Scale bar = 10 µm.

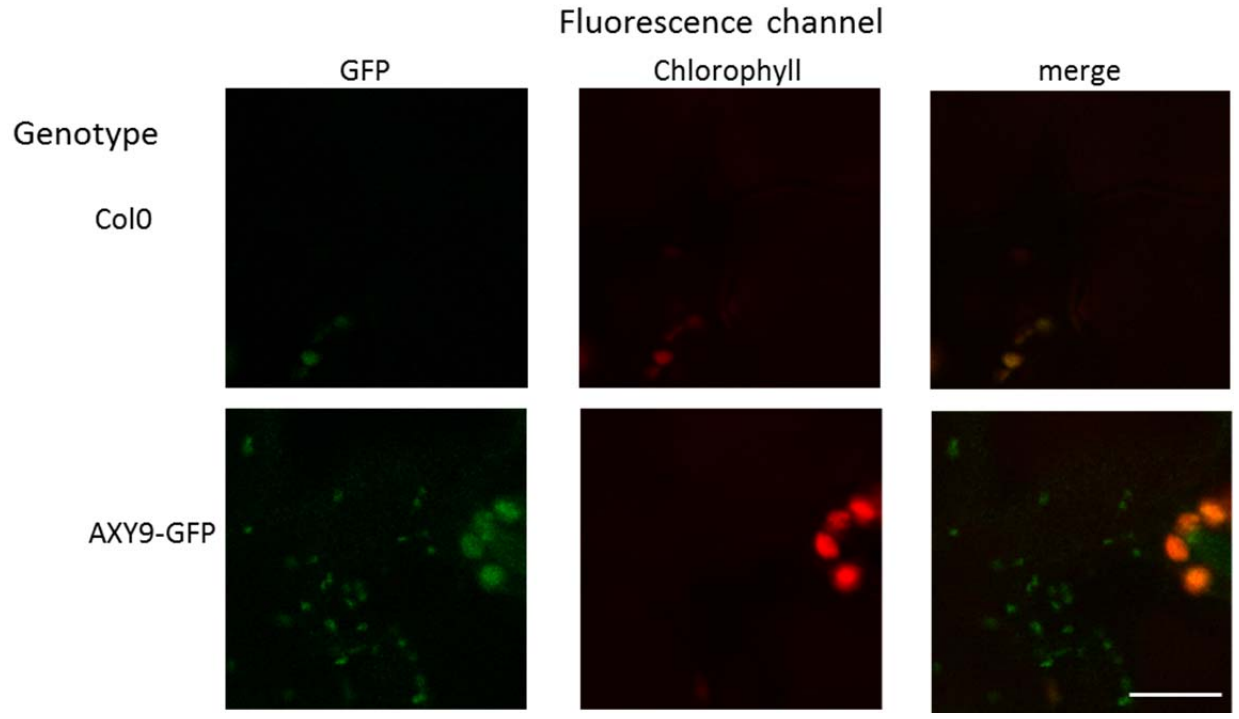


Figure 3-26. Subcellular localization of AXY9-GFP in *A. thaliana*. *A. thaliana* constitutively expressing a AXY9-GFP in the *axy9.2* background was examined using a confocal microscope to look for localization of the AXY9 protein. The GFP and chlorophyll fluorescence channels are shown along with a merged image (columns). A wild type control did not show any GFP fluorescence (top row). For the line expressing AXY9-GFP, a fluorescence signal was observed in small mobile intracellular vesicles consistent with Golgi. The larger florescent bodies visible in the chlorophyll and the GFP channels are chloroplasts. Scale bar = 10 µm.

3.3 Discussion

3.3.1 Identification of additional *axy3*, *axy8*, *mur1* and *mur3* alleles

Several of the *axy* mutants characterized as part of this work have OLIMP phenotypes similar to previously published mutants affect XyG structure including *mur1*/*mur2*, *mur3* and the recently published *axy3* and *axy8* (Figure 3-1, Table 3-1, Table 3-7) (Gunl et al. 2011; Gunl and Pauly 2011; Madson et al. 2003; Vanzin et al. 2002; Zablackis et al. 1996). As part of the work presented in this thesis, three independent alleles were identified for *axy3*, four for *axy8*, and one each for *mur1* and *mur3* (Table 3-7). The total number of alleles identified for *axy3* and *axy8* in the entire *axy* screen, five and eight respectively, is higher than that seen for the MUR genes, AXY4 or AXY9 (Table 3-1, Table 3-7). One possible explanation for this is that both AXY3 and AXY8 glycosyl hydrolases are large proteins with 915 and 843 amino acids, respectively,

compared to MUR1 (373), MUR3 (619), AXY4 (416), and AXY9 (368). The size of the coding sequence would be expected to be approximately proportional to the relative frequency at which an EMS mutation occurs within that sequence. Therefore, the larger AXY3 and AXY8 would be expected to be mutated more frequently than the smaller MUR1 and AXY9. Additional factors, such as sensitivity of the protein to sequence perturbation and growth phenotype of knockouts (in the case of AXY9) could also play a role in the number of mutants identified. As the genes associated with the AXY3 and AXY8 loci have been characterized previously, mutants harboring mutations in them were not followed up beyond identification of the causative mutation.

3.3.2 Genetic basis of the *axy6* mutant

Like *axy8* mutants, the *axy6* mutant has an increase in fucosylated XyG oligosaccharides (Figure 3-1). However, *axy6* appears distinct as it lacks the unusual oligosaccharides observed in *axy8* mutants (Figure 3-1) (Gunl et al. 2011), a mutation was not detected within the AXY8 locus in the *axy6* mutant, and the *axy6* mutation was mapped to a different chromosome from where AXY8 is located (Figure 3-12, Figure 3-13). Of 16 candidate genes for the genetic basis of the *axy6* mutation (Table 3-4) the only one previously implicated in XyG metabolism is XTR4, a member of the XyG endotransglycosylase/hydrolase family (Eklof and Brumer 2010). However, if and which of these genes harbors a mutation that causes the *axy6* XyG phenotype could not be ascertained here.

Future work will involve testing T-DNA knockout lines for these and other candidate genes for the *axy6* XyG OLIMP phenotype. As the mapped region was quite large due to uncertainties in mutant selection, it would be advantageous to narrow down the region either by repeating the mapping procedure with a stricter cutoff for the mutant phenotype or by phenotyping M3 seedlings collected separately from M2 plants.

3.3.3 Effect of AXY9 deficiency on plant morphology and growth

3.3.3.1 Correlation between decreased polysaccharide *O*-acetylation and irregular / collapsed xylem

While the xylem tissue of wild type *A. thaliana* stems has large and open vessels, the xylem vessels of the *axy9.2* mutant are severely collapsed (Figure 3-21). Similar but less severe phenotypes were reported in mutants defective for TBL29 (Xiong et al. 2013) and an RWA quadruple mutant (Lee et al. 2011). TBL29 is a putative xylan *O*-acetyltransferase (Xiong et al. 2013; Yuan et al. 2013) and the RWA proteins affect the acetylation of multiple polysaccharides including xylan, XyG, and pectin (Lee et al. 2011; Manabe et al. 2011). As xylan acetylation is reduced in the *axy9.2*, *tbl29* and the RWA quadruple mutant and xylan is an important structural polymer for proper xylem tissue development and function (Brown et al. 2007; Keppler and Showalter 2010), it appears that a defect in xylan acetylation may be the cause for the collapsed xylem phenotype in the *axy9.2* mutant. However, a phenotype was not observed in *A. thaliana* expressing a fungal xylan acetyl esterase that resulted in ~50% reduction in stem acetyl ester content (Pogorelko et al. 2013), a similar reduction observed in the *tbl29* mutant, which does have a strong phenotype (Xiong et al. 2013). This suggests that perhaps the timing of polysaccharide acetylation (transfer during synthesis in the Golgi versus removal in the

apoplast) or the specific positions of the acetyl groups, which could be different between *tbl29* and lines expressing the fungal xylan acetyl esterases, have a significant effect on xylan function ultimately resulting in the collapsed xylem phenotype.

3.3.3.2 Possible mechanisms by which reduced xylan acetylation may lead to collapsed xylem

While decreasing xylan acetylation appears to result in a collapsed xylem phenotype in some cases, eliminating glucuronic acid substituents has little or no effect (Mortimer et al. 2010). Acetylation of xylan, by interfering with hydrogen bonding between xylan chains, likely contributes to maintaining solubility of the polysaccharide. Glucuronic acid could likely function in a similar way. But as the remaining glucuronic acid is not able to substitute for the loss of the acetyl groups and a loss of the glucuronic acid does not impact plant growth (Mortimer et al. 2010) it appears that acetylation has a unique function. As mutants deficient in xylan biosynthesis have similar collapsed xylem phenotypes to those mutants deficient for xylan acetylation (Brown et al. 2007; Keppler and Showalter 2010), an alternative hypothesis is that xylan acetylation is somehow required for xylan biosynthesis. However, a decrease in the total amount of xylan was not observed in *axy9.1*, *tbl29* or the RWA quadruple mutant (Figure 3-23) (Lee et al. 2011; Xiong et al. 2013). Yet another hypothesis is that xylan acetylation can act to protect xylan polysaccharides from endogenous glycosyl hydrolases, such as endoxylanases or xylosidases, as has been shown for fungal xylanases (Biely et al. 1986). The role of endogenous plant xylan degrading enzymes is not fully understood but they seem to be involved in development (Goujon et al. 2003; Minic et al. 2004; Simpson et al. 2003). Some endoxylanases, notably At1g58370, are strongly expressed specifically in stem tissue and localize to the apoplast in *A. thaliana* (Suzuki et al. 2002; Winters et al. 2007). A reduction of xylan acetylation could lead to enhanced xylan hydrolysis at an important developmental stage, thereby resulting in a weakened wall. Alternatively, oligosaccharides released by xylanase activity could have a signaling role ultimately affecting wall biosynthesis as has been proposed for mannan oligosaccharides (Zhao et al. 2013). The identification of suppressor mutants of the collapsed xylem phenotype of the *axy9.2* mutant line may provide insight into the mechanism by which xylan acetylation affects the integrity of xylem vessels. Another approach to test the xylanase hypothesis would be to overexpress xylanase inhibitor peptides (Juge et al. 2004) in the *axy9.2* mutant to see if these could rescue the collapsed xylem phenotype.

3.3.3.3 Irregular / collapsed xylem vessels may lead to decreased growth in *axy9.2*

While *axy9.1* plants appear only slightly smaller than wild type plants, *axy9.2* plants are severely deficient in growth, have dark-green leaves, and are delayed in flowering. This phenotype can be rescued by overexpression of AXY9 (Figure 3-20). A growth phenotype is not observed in etiolated seedlings grown on plates suggesting that a wall elongation defect is not responsible for the growth phenotype. Rather, it appears that the collapsed xylem phenotype may be at least partially causal for the dwarf phenotype of the *axy9.2* mutant. There are several examples of mutants with defects in the wall structure that have aberrant xylem vessels and subsequently diminished growth (Brown et al. 2007; Turner and Somerville 1997). Because the xylem vessels must withstand negative pressure to conduct water up the stem, a general defect in cell wall strength could lead to partial or full collapse of xylem vessels (Amrhein et al. 1983;

Turner and Somerville 1997). A lack of functional xylem would likely impede the flow of water and nutrients from the roots to the leaves, thereby affecting plant growth. As the *axy9.2* plants do not appear wilted or seem to be suffering from water stress (Figure 3-20), also observed for other *irx* mutants (Chen et al. 2005; Turner and Somerville 1997), it may be that a regulatory mechanism exists to sense a lack of appropriate water transport and prevents the plants from growing beyond the water conducting capacity of the xylem tissue. This hypothesis is supported by observation that the expression of drought response genes and abscisic acid (ABA) levels are increased in the *lew2* (CesA8/IRX1) mutant and that growing this mutant in high humidity conditions can partially rescue the growth and dark-green leaf phenotypes (Chen et al. 2005). Notably, mutants defective for gibberellin (GA) biosynthesis or perception are dwarfed, have dark-green leaves, and are delayed in flowering (Koornneef et al. 1990) reminiscent to the phenotype observed for the *axy9.2* mutant (Figure 3-20). This suggests that repression of GA signaling may be involved in the phenotype observed in *axy9.2* and to a lesser extent in *lew2*. ABA and GA often have antagonistic roles (Razem et al. 2006) suggesting that the production of ABA in response to water stress perceived by *axy9.2* / *lew2* mutants might reduce the production / perception of GA resulting in dwarfed plants with dark-green leaves and delayed flowering.

3.3.3.4 *axy9.1* and possibly *axy9.2* represent weak alleles of *axy9*

The *axy9.2* mutant, containing a T-DNA insertion in the AXY9 gene, has a stronger acetylation defect in etiolated seedlings than the point mutant *axy9.1,* suggesting that *axy9.1* is a weak allele (Figure 3-18, Figure 3-22). The *axy9.1* mutation causes a stop codon to be introduced prior to the location of the T-DNA insertion in the *axy9.2* mutant (Appendix 1, Figure 3-19). The T-DNA insertion results in a coding sequence containing approximately ¾ of the original AXY9 CDS with an approximately 25 codons introduced by the T-DNA insertion before a stop codon (Appendix 1). The effect of these extra amino acids is unknown, but as the truncated AXY9 protein in *axy9.1* seems to be at least partially functional there is the possibility that the aberrant AXY9 protein produced in *axy9.2*, which contains the entire peptide present in *axy9.1*, is also partially functional. This is one possible explanation as to why XyG acetylation is still present in the *axy9.2* as there are no other copies of this gene in *A. thaliana*. The extra 25 amino acids present in the protein coded for by the *axy9.2* allele could destabilize the folding of the protein or interfere with substrate binding / protein activity. Alternatively, as the wild type 3' UTR is disrupted in this allele the mRNA could be less stable than that in *axy9.1*. Either of these possibilities could result in the *axy9.1* allele having a less severe phenotype than of *axy9.2*, despite *axy9.2* having a greater amount of intact AXY9 polypeptide sequence. If *axy9.2* represents indeed a weak allele, the severe growth phenotype of this mutant suggests that a full knockout may be lethal.

The total acetic acid content of leaf material was not significantly affected in the *axy9.1* mutant compared to wild type (Figure 3-23B). As a large portion of leaf acetic acid content is derived from pectin (Gille et al. 2011b), it seems likely that AXY9 is not involved in pectin acetylation, that an alternative mechanism regulates the amount of pectin acetyl esterification such as pectin acetyl esterases (Williamson 1991, Gou 2012), or that the truncated AXY9 protein forms are still able to function sufficiently for acetylation of pectin to occur.

3.3.4 Model of AXY9 involvement in plant cell wall polysaccharide acetylation

3.3.4.1 Other players in polysaccharide acetylation

The AXY9 gene does not have any close homologs in *A. thaliana* but the gene does share some similarity with other gene families. Notably, AXY9 shares distant homology with the TBL proteins (Bischoff et al. 2010). The TBL proteins are hypothesized to be polysaccharide acetyltransferases, with AXY4 believed to be an *O*-acetyltransferase specific for XyG (Gille et al. 2011b; Xiong et al. 2013; Yuan et al. 2013). That AXY9 shares homology with the TBLs suggests AXY9 harbors a similar catalytic activity, possibly transferring an acetyl group between a donor and an acceptor substrate.

In addition to the TBL proteins, the RWA proteins have been shown to be involved in polysaccharide acetylation in plants (Lee et al. 2011; Manabe et al. 2011). The RWA proteins, of which there are four in *A. thaliana*, have multiple transmembrane domains and have been hypothesized to act as transporters for an activated form of acetate upstream in the pathway from the TBLs (Manabe et al. 2011). The RWA proteins have been reported to be in the Golgi (Lee et al. 2011) or ER (Manabe et al. 2011) whereas AXY4 is believed to be localized only in the Golgi (Gille et al. 2011b) as is AXY9 (Figure 3-25, Figure 3-26). While TBLs contain a single transmembrane domain, AXY9 has two putative transmembrane domains at the N terminus. With two transmembrane domains, the protein is predicted to have its active site on the cytoplasmic side of Golgi membrane. If this topology prediction is correct, and the RWA proteins are indeed integral membrane proteins acting as transporters, it would suggest that AXY9 acts upstream of the RWAs, possible affecting their substrate. While acetyl-CoA has been demonstrated to act as at least an upstream donor of activated acetate for acetylation of wall polysaccharides when fed to plant microsomes (Pauly and Scheller 2000), it has not been shown that acetyl-CoA is the direct donor for the polysaccharide acetyl transfers. A possible role of AXY9 is therefore to generate an acetylated intermediate that can be transported across the Golgi membrane by the RWAs and utilized by the polysaccharide acetyltransferases.

3.3.4.2 Acetyl-CoA shuttling systems

There are several types of shuttling systems that work to maintain and regulate separate pools of acetyl-CoA in the cytoplasm, mitochondria, peroxisome and plastid. In plants, acetyl-CoA equivalents can be transported out of the mitochondria in the form of citrate via conversion of acetyl-CoA to citrate by the enzyme citrate synthase (Wiegand and Remington 1986). Once in the cytosol, the enzyme citrate lyase can act on citrate to form acetyl-CoA (Fatland et al. 2002). This pathway is believed to be the major source of cytosolic acetyl-CoA and silencing it via RNAi can lead to severe growth phenotypes in *A. thaliana* (Fatland et al. 2005; Oliver et al. 2009). Interestingly, some of the growth phenotypes observed in the ATP citrate lyase knockdown lines, such as dark-green leaves, dwarfism, and delayed senescence, resemble those observed with the *axy9.2* mutant (Figure 3-20) (Fatland et al. 2005). A second acetyl-CoA shuttling system known to exist utilizes the amino acid carnitine as a carrier molecule. The enzyme carnitine acetyltransferase can reversibly catalyze the transfer of the acetyl group from acetyl-CoA onto carnitine (Kispal et al. 1993; Schmalix and Bandlow 1993), which can then be transported between intracellular compartments by acetyl-carnitine/carnitine antiporters (Indiveri et al. 1997; Pande and Parvin 1980). Carnitine acetyltransferase activity has been detected in plant

mitochondrial and chloroplast membranes and acetylcarnitine can act as a donor for the formation of fatty acids in chloroplast suspensions (Budde et al. 1991; Masterson and Wood 2000; Thomas and Mcneil 1976; Thomas and Wood 1982; Wood et al. 1983). The carnitine acetyltransferase enzymes in both the chloroplast and the mitochondria seem to be localized in the stroma or matrix respectively, suggesting that acetylcarnitine may move out of the mitochondria and into the chloroplast without entering the cytoplasmic acetyl-CoA pool. The RWA proteins do not share homology with known acetylcarnitine transporters from mammalian systems, but plants do contain such homologs (Lawand et al. 2002). Additionally, AXY9 does not share homology to known carnitine acetyltransferases. This suggests that AXY9 and the RWA proteins may be part of a novel pathway for shuttling activated acetate across the Golgi membrane for polysaccharide acetylation. A protein (AT-1 / SLC33A1) that can transport acetyl-CoA into the ER has been identified in mammalian cells (Jonas et al. 2010), however it does not share homology to the RWA proteins.

3.3.4.3 AXY9 as part of a system for moving activated acetate into the Golgi for polysaccharide acetylation

A system for shuttling acetyl groups into the plant Golgi has yet to be discovered. Based on sequence homology to TBLs, topology prediction, and effect on the acetylation of multiple polysaccharides it appears that AXY9 may be a part of such a system. In this hypothetical model, AXY9 would act in the cytoplasm to transfer activated acetate from acetyl-CoA to an unknown intermediate (Figure 3-27). This yet unknown acetyl carrier molecule would then be transported into the Golgi by the RWA proteins where it would be used as a substrate by the TBLs. Future work is needed to test this hypothesis and, if correct, to identify the activated acetyl intermediate.

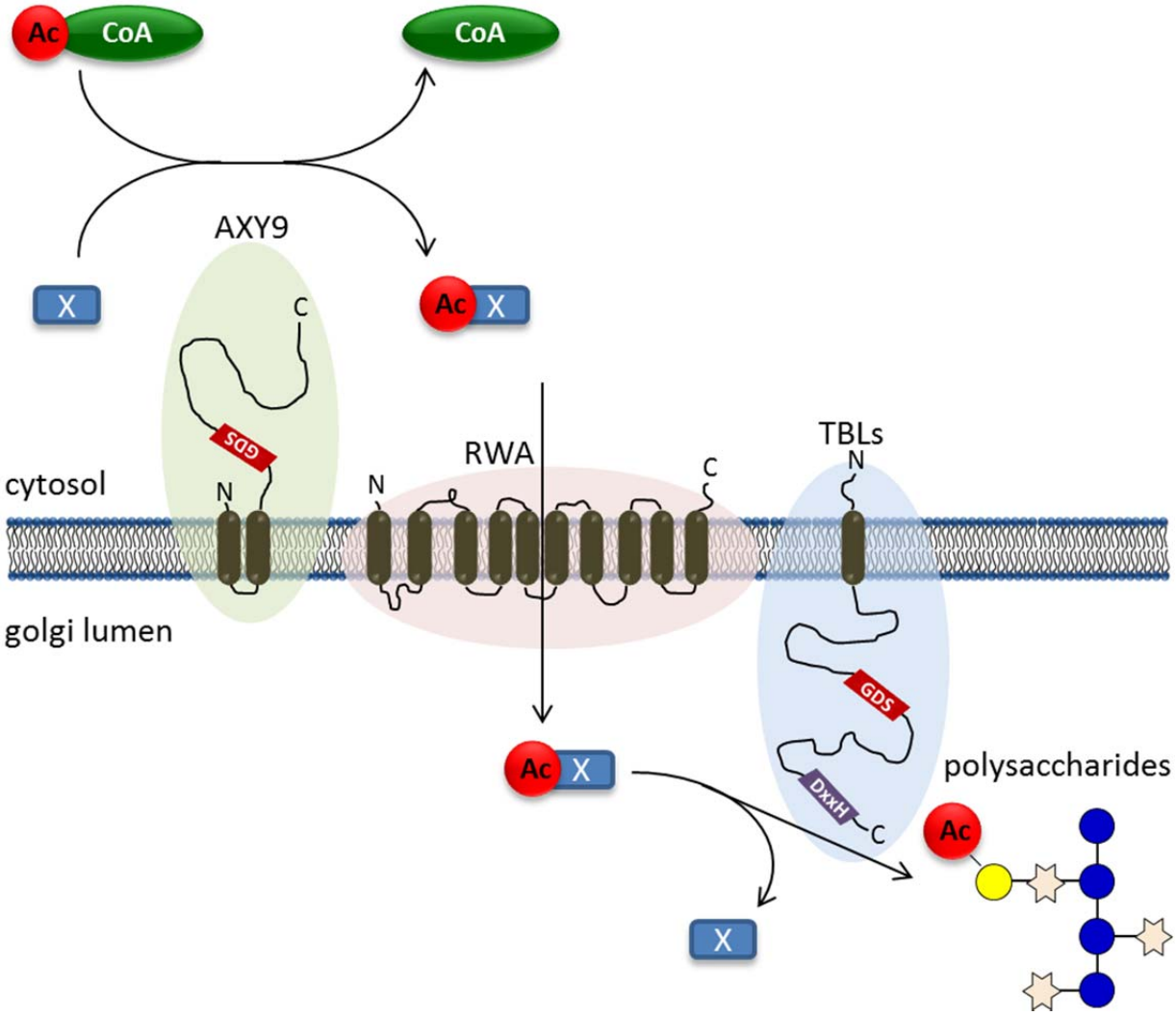


Figure 3-27. Model of plant cell wall polysaccharide acetylation. The TBL gene family is believed to act as polysaccharide *O*-acetyltransferases, whereas the RWAs are believed to function as transporters for an activated acetyl donor substrate. The AXY9 protein may act upstream of RWA to produce this hypothetical activated acetyl donor substrate (depicted as Ac-X) from acetyl-CoA. AXY9 is not a member of the TBL family, but shares sequence similarity including the GDS motif.

4 A transcriptional profiling approach to identify XyG biosynthetic genes

4.1 Background

The biosynthesis and metabolism of XyG is a complex process requiring the coordinated action of many genes. The best understood of these are the proteins that directly affect XyG structure including glycosyltransferases, glycosyl hydrolases, and proteins involved in the production of nucleotide sugars (1.1.3, 1.2.3, 1.2.4) (Gunl et al. 2011; Seifert 2004; Zabotina 2012). Genes that still need to be identified in XyG metabolism in *A. thaliana* are a galactosyltransferase with specificity for the second position of the XXXG motif, nucleotide sugar transporters, a glucosidase and plasma membrane localized monosaccharide transporters for recycling monosaccharides released by XyG turn-over in the apoplast. Based on the variation in the quantity and structure of XyG produced in different tissues and developmental stages, regulatory proteins such as kinases, phosphatases, and transcription factors may also play a role in XyG metabolism by controlling the activity and abundance of the XyG metabolic machinery. Such regulatory proteins have yet to be identified for XyG, but a complex regulatory network of transcription factors has been shown for lignin biosynthesis (Zhao and Dixon 2011), at least one of member of which can be modified post-translationally (Morse et al. 2009).

Expression analysis can be an effective approach for identifying genes involved in a specific pathway and has been used previously to identify genes involved in cell wall biosynthesis such as *IRX6-13*, *MYB52* and *PRR1* (Brown et al. 2005; Cassan-Wang et al. 2013; Persson et al. 2005; Ruprecht et al. 2011). One of the limitations of expression analysis to identify the biosynthetic genes for specific wall components is that multiple wall components are often made concurrently. Therefore transcriptional analysis approaches, such as co-expression analysis with known biosynthetic genes, are not always able to identify a compelling set of candidate genes. However, by identifying a tissue and developmental time in which the production of one polymer dominates, the corresponding set of biosynthetic genes will likely be expressed at higher levels compared to those for the other polymers. Such an approach has been used to identify genes involved in the synthesis of mannan in *Cyamopsis tetragonoloba* (guar) endosperm (Dhugga et al. 2004), mannan in *Amorphophallus konjac* (voodoo lily) tuber (Gille et al. 2011a), xylan in *Plantago ovata* (psyllium) seed coat (Jensen et al. 2011), callose from *Nicotiana alata* (tobacco) pollen tube (Doblin et al. 2001), cellulose in *Gossypium hirsutum* (cotton) fiber (Pear et al. 1996) and XyG in *Tropaeolum majus* (nasturtium) cotyledon (Cocuron et al. 2007).

A variety of storage molecules can be used by plants including starch, oils, proteins and cell wall polysaccharides such as mannans, galactans, and/or XyG (Buckeridge 2010; Buckeridge et al. 2000; Quettier and Eastmond 2009; Shewry et al. 1995). Nasturtium uses XyG as a seed storage polysaccharide in its cotyledons (Edwards et al. 1985). Developing nasturtium seeds produce large amounts of this polymer (up to 20% dry weight), which is degraded during seed germination for energy (Hoth et al. 1986). Unlike structural XyG in the primary wall nasturtium, the seed storage XyG in this species is non-fucosylated and non-acetylated (Desveaux et al. 1998; Fanutti et al. 1996; Le Dizet 1972). Because XyG deposition occurs over a period of

several days (Figure 4-1) (Desveaux et al. 1998), the transcriptome of developing nasturtium seeds should show an induction of relevant XyG biosynthetic genes during this time. This hypothesis was validated by the identification of a XyG glucan synthase gene by sequencing clones from a cDNA library derived from developing nasturtium seeds (Cocuron et al. 2007). The advent of high-throughput sequencing methods such as Illumina and 454 (Kircher and Kelso 2010) allows for much deeper sampling of the transcriptome and quantitative analysis of expression (Gille et al. 2011a; Morozova et al. 2009; Wang et al. 2010). In the work presented here, such an approach was used to facilitate the identification of additional genes involved in XyG biosynthesis. This work was published in Molecular Plant with the title "RNA-Seq analysis of developing nasturtium seeds (*Tropaeolum majus*): identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis" by the authors Jensen JK, Schultink A, Keegstra K, Wilkerson CG, and Pauly M in April 2012 (Jensen et al. 2012).

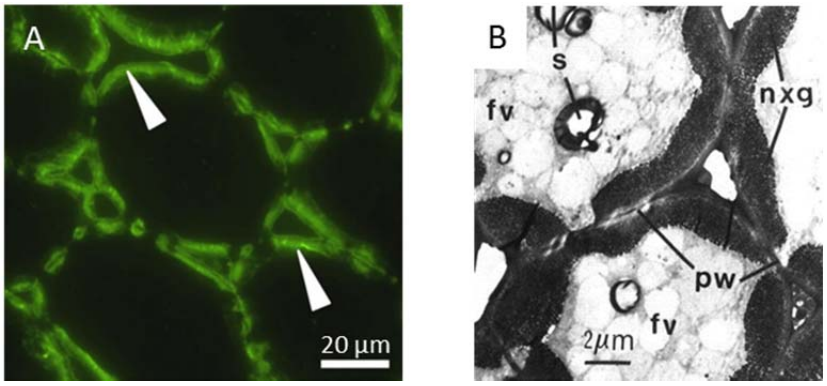


Figure 4-1. Accumulation of storage XyG in developing nasturtium seeds. *Tropaeolum majus* (nasturtium) cotyledons accumulate large amounts of non-fucosylated XyG during development for use as an energy source during seed germination. The seed storage XyG can be visualized by XyG-specific antibody labeling (A) or as a dark-staining material in TEM (nxg, B). Images reprinted from Marcus et al. 2008 and Desveaux et al. 1998 (Desveaux et al. 1998; Marcus et al. 2008).

4.2 Results

4.2.1 Quantification and timing of XyG deposition in developing nasturtium seeds

To correlate gene transcription with the production of seed storage XyG it was necessary to determine the timing of seed storage XyG deposition. The abundance of seed storage XyG can be differentiated from structural XyG in the primary wall by the lack of fucosyl groups (Desveaux et al. 1998). The CCRC-M1 and CCRC-M100 antibodies are specific to fucosylated and non-fucosylated XyG, respectively (Pattathil et al. 2010; Puhlmann et al. 1994) and can therefore be used to determine the relative abundance of structural and storage XyG. A dot blot was performed using AIR material from developing nasturtium seeds at various stages of development from 15 to 25 days post anthesis (dpa) (Figure 4-2, courtesy of Jacob Jensen, 2.5.14). The ratio of the signal from these antibodies revealed that seed storage XyG accumulates from 16 to 25 dpa. The relative amount of seed storage XyG increased in a linear and approximately constant rate during this time based on the assay used.

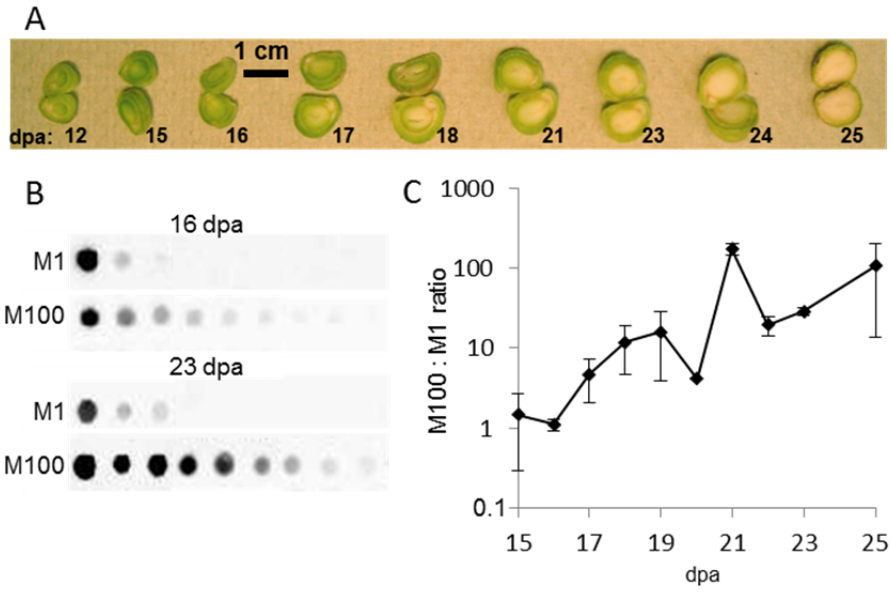


Figure 4-2. Immunodetection of peak XyG production in developing nasturtium seeds. (A) Developing nasturtium seeds (cut in half) from 12 to 25 days post anthesis (dpa). (B) Dot blot of XyG extracted from 16 and 23 dpa using the CCRC-M1 and CCRC-M100 antibodies, which are specific for fucosylated (structural) and non-fucosylated (seed storage) XyG respectively. (C) The ratio of the dot-blot response to determine the peak time of seed storage XyG accumulation. Figure courtesy of Jacob Jensen.

4.2.2 Transcriptional profiling of developing seeds

A reference genome was not available for nasturtium so a dual pronged sequencing approach was used. 454 sequencing was performed on cDNA constructed from RNA isolated from nasturtium seeds at 16, 18, 22 and 25 dpa (sample preparation by Jacob Jensen, sequencing performed by the Research Technology Support Facility at Michigan State University). The resulting sequencing reads were used to assemble contigs (performed by Nick Thrower and Curtis Wilkerson, Michigan State University), many of which corresponded to partial or complete transcripts. Illumina sequencing was then performed on cDNA from 13, 16, 17, 18, 21 and 25 dpa to obtain greater temporal resolution and deeper sampling (sample preparation by Jacob Jensen, sequencing performed by the Research Technology Support Facility at Michigan State University). The Illumina sequencing reads, shorter but in greater number than those produced by 454 sequencing (Table 4-1), were mapped back onto the contig scaffolds from the 454 assembly. Approximately 72% of the Illumina reads were mapped onto the 454 contigs (Table 4-2). The reads per contig kilobase per million reads (RPKM) was used as a measure of relative abundance for each contig at each time point.

	Developmental stage (dpa)							
	13	16	16	17	18	21	22	25
454 FLX								
Total reads		$3.79x10^5$	$5.11x10^5$		$3.37x10^5$	$7.20x10^5$	$5.18x10^5$	$1.30x10^5$
Normalized		no	yes		no	yes	no	no
Avg read length (bp)		217	350		245	362	220	215
Illumina								
Total reads	$6.78x10^6$	$7.08x10^6$	$7.06x10^6$	$7.69x10^6$	$7.79x10^6$	$7.76x10^6$		$7.45x10^6$

Table 4-1. Sequencing of cDNA from developing nasturtium seeds. 454 FLX sequencing was performed to obtained longer reads for contig assembly. In two cases (16 dpa and 21 dpa) normalization of the cDNA was performed to enrich lowly abundant cDNAs prior to sequencing. Illumina sequencing was subsequently conducted to obtain greater depth of coverage and higher resolution. Illumina sequencing was performed with a read length of 36 base pairs. Sample preparation by Jacob Jensen, sequencing performed by the Research Technology Support Facility at Michigan State University.

Assembly of 454 reads	
Total 454 reads	$2.60x10^6$
Total assembled contigs	38,885
Mapping of Illumina reads to contigs	
Total Illumina reads	$5.16x10^7$
Illumina reads mapped	72%
Contigs with at least one mapped read	90%
Total contigs with at least five mapped reads	36,120

Table 4-2. Contig assembly and read mapping of nasturtium transcriptome. The 454 reads were assembled into contigs. The Illumina reads were subsequently mapped onto these contigs to obtain a transcriptional profile of each contig throughout the developmental time course. Assembly of sequencing reads performed by Nick Thrower and Curtis Wilkerson.

4.2.3 Selection of a candidate XyG galactosyltransferase

Several control genes known to be involved in XyG biosynthesis were present in the dataset including *TmCSLC4*, *TmXXT2*, and *TmMUR3*, which were identified based on homology to the corresponding *A. thaliana* genes. The expression of these genes increased throughout the time course (Figure 4-3). To find genes with a similar patterned of induction, a linear regression line was calculated for each contig and the contigs were ranked by the slope of this line. Rather than being related to XyG biosynthesis, many of the most highly induced genes were related to other processes in developing seeds including oil accumulation (oleosins), seed storage proteins (albumins, cupins, cruciferins), and antiherbivory (protease inhibitors) (Table 4-3) (Crouch and Sussex 1981; Dunwell et al. 2004; Frandsen et al. 2001; Habib and Fazili 2007; Shewry et al.

1995). To obtain a smaller set of candidate genes involved in XyG biosynthesis, only contigs with homology to known GTs (CAzy database, www.cazy.og) were included. Of the identified glycosyltransferases, *TmCSLC4*, *TmXXT2*, and *TmMUR3* were among the four most highly induced, demonstrating the validity of this approach (Figure 4-4). The most highly induced GT, a putative ortholog of the *A. thaliana* gene *At5g62220*, had not been described previously. At5g62220 is a member of GT family 47, the same GT family as MUR3. As MUR3 specifically galactosylates XyG at the third position of the XXXG motif, it was hypothesized that At5g62220 might represent a XyG galactosyltransferase with specificity for the second position of this motif. Several other GTs were also found to be induced to lesser degrees including a putative galactinol synthase, sucrose synthases, and glycoside synthases (Figure 4-4B). Galactinol synthase is involved in the formation raffinose family oligosaccharide commonly found in seeds (Liu et al. 1998; Min Kuo et al. 1997). Indeed, the most similar gene from *A. thaliana*, which does not use XyG as a seed storage polysaccharide, is expressed specifically in seeds (Winters et al. 2007). Glycosides are found in a variety of tissues including and can serve many different functions including as pigment molecules, as storage molecules, and in herbivory defense. Based these putative functions, it is not expected that these genes are related to XyG production.

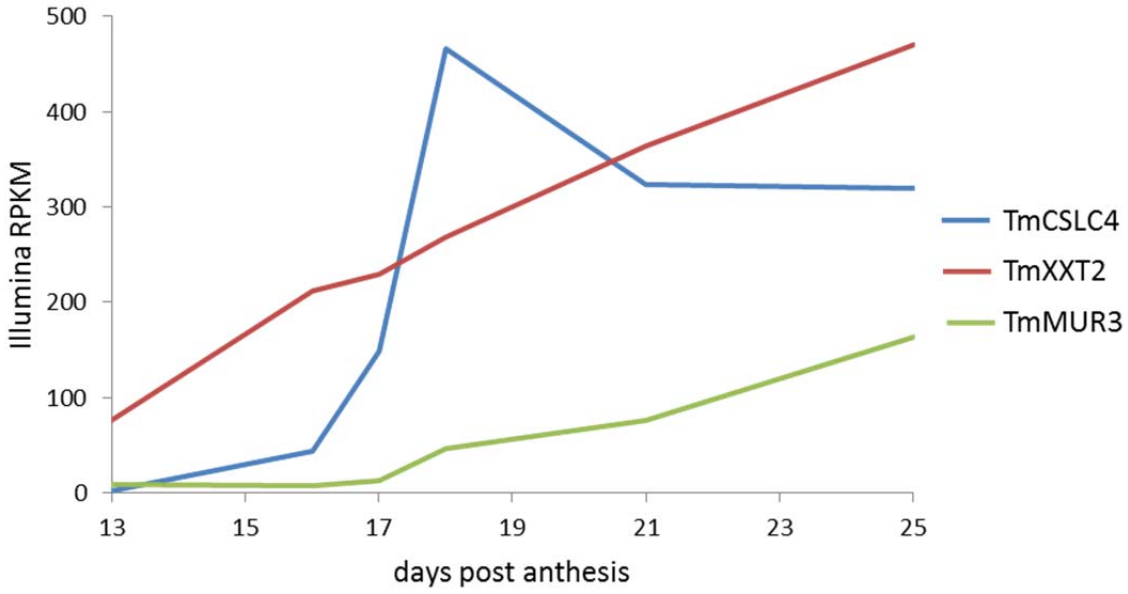
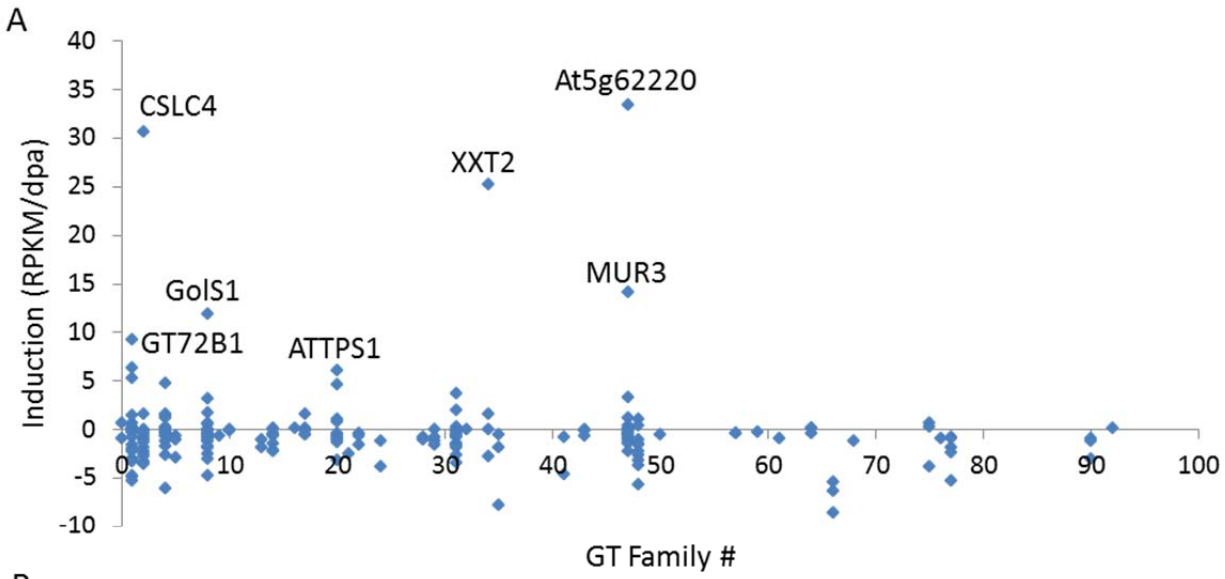


Figure 4-3. Transcript abundance of *TmCSLC4*, *TmXXT2* and *TmMUR3*. The reads per kilobase per million reads (RPKM, from Illumina data) are shown for the three nasturtium contigs, which are putative orthologs of the *A. thaliana* genes CSLC4, XXT2 and MUR3. The expression of these three contigs increased throughout the time course, with the exception of a peak for TmCSLC4 at 18 dpa.

Contig	Closest At Hit	Annotation	Induction
M01000032754	At2g17580	polynucleotide adenyltransferase family protein	6479
M01000029378	At2g25890	oleosin family protein	2208
M01000031461	At5g59320	lipid transfer protein 3	1943
M01000038093	At3g22640	cupin family protein	1725
M01000029243	At3g46230	heat shock protein 17.4	1462
M01000028700	At1g63980	D111/G-patch domain-containing protein	1380
M01000019967	At5g54740	seed storage albumin 5	1230
M01000020735	At2g01690	ARM repeat superfamily protein	1170
M01000029491	At1g62500	bifunctional inhibitor/lipid-transfer protein/seed storage protein	1086
M01000019523	At1g47710	serine protease inhibitor (SERPIN) family protein	1019

Table 4-3. Top ten most-induced contigs annotated with the closest *A. thaliana* homolog. Contigs for which a corresponding *A. thaliana* homolog was not identified are omitted. The slope of a linear regression line of the relative contig abundance (RPKM) was used as a measure of the induction (RPKM / dpa).

A

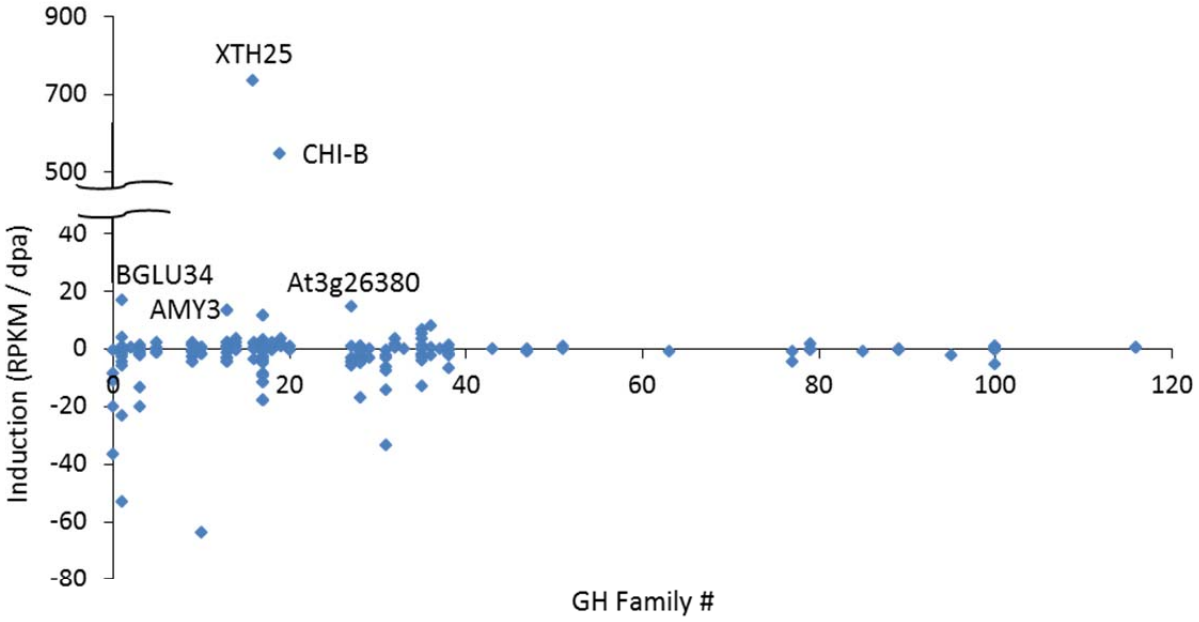


B

	Induction	Closest homolog from *A. thaliana*			
GT Family	(RPKM/dpa)	Name	AGI	Putative function	Expression pattern
47	33		At5g62220		widespread, dry seeds
2	31	CSLC4	At3g28180	XyG Biosynthesis	seedling, leaves
34	25	XXT2	At4g02500	XyG Biosynthesis	widespread
47	14	MUR3	At2g20370	XyG Biosynthesis	widespread
8	12	GolS1	At2g47180	galactinol synthase	seeds
1	9	GT72B1	At4g01070	glycosides	senescing leaves, internode, sepals
1	6		At5g49690	glycosides	seeds
20	6	ATTPS1	At1g78580	trehalose-P synthase	meristem, widespread
1	5		At2g28080	glycosides	widespread
4	5	SUS3	At4g02280	sucrose synthase	seeds

Figure 4-4. Glycosyltransferase transcipt induction. (A) The slope of a linear regression line for the transcription profile of each contig was used as a measure of the induction (Y-axis). Contigs were grouped by CAZY GT family (X-axis) (http://www.cazy.org). The four most induced contigs are named based on the closest *A. thaliana* homolog. Several known XyG GTs were identified including CSLC4, XXT2, and MUR3. The most highly induced contig corresponds to the *A. thaliana* protein At5g62220 and is in GT47, the same family as MUR3. (B) The top ten most-induced GTs are listed with the closest homolog from *A. thaliana* with the putative function/activity and expression pattern of that homolog. The expression pattern was obtained from the Arabidopsis eFP Browser (Winters et al. 2007).

Contigs with homology to known glycosyl hydrolases were also identified. Two contigs, representing a putative XyG transglycosylase and chitinase, were found to have high levels of transcriptional induction (Figure 4-5). A XyG transglycosylase could have a role in apoplastic remodeling of storage XyG polysaccharides. However, the closest *A. thaliana* homolog to the detected nasturtium transcript is strongly expressed in seed tissue as well (Figure 4-5) (Winters et al. 2007). As *A. thaliana* has not been reported to use XyG as a seed storage polymer the abundance of this transcript in seed tissue suggests it may have a function independent from the biosynthesis of seed storage XyG. The putative chitinase transcript identified shows high similarity to a chitinase from maize, which was purified and shown to act as a chitinase *in vitro* (Huynh et al. 1992). Putative glucosidases and galactosidases detected could also have a role in XyG biosynthesis, although they had much lower levels of transcriptional induction (Figure 4-5).



B

GH Family	Induction (RPKM/dpa)	Closest homolog from *A. thaliana* Name	AGI	Putative activity	Expression pattern
16	735.2	XTR3,XTH25	At5g57550	xyloglucan transglycosylase	seeds
19	547.5	CHI-B, PR3	At3g12500	chitinase	roots
1	16.7	BGLU34	At1g47600	beta glucosidase	roots
27	14.3		At3g26380	melibiase	seeds
13	12.9	AMY3	At1g69830	amylase	leaves
17	11.3		At3g07320		shoot apical meristem
36	7.5	SIP1	At5g40390	raffinose synthase	hypocotyl
35	6.0	BGAL3	At4g36360	beta galactosidase	seeds
35	5.1	BGAL9	At2g32810	beta galactosidase	widespread
1	3.5	BGLU41	At5g54570	beta glucosidase	seeds, stamens

Figure 4-5. Glycosyl hydrolase transcipt induction. (A) The slope of a linear regression line for the transcription profile of each contig was used as a measure of the induction (Y-axis). Contigs were grouped by CAZY GH family (X-axis) (http://www.cazy.org). The two most induced contigs are named based on the closest *A. thaliana* homolog. The most highly induced contig was most similar to the *A. thaliana* gene XTH25, a member of the XyG transglycosylase/hydrolase family. (B) The top ten most-induced GHs are listed with the closest homolog from *A. thaliana* with the putative activity and expression pattern of that homolog. The expression pattern was obtained from the Arabidopsis eFP Browser (Winters et al. 2007).

4.2.4 At5g62220 affects XyG galactosylation

To investigate the putative role of At5g62220 in XyG biosynthesis, an *A. thaliana* line (GK-552C10-021687) containing a T-DNA insertion within the coding sequence of the corresponding gene was obtained (Figure 4-6A, 2.2.2). A full-length transcript for this gene was not detected by RT-PCR in plants homozygous for this T-DNA insertion, indicating the protein function is likely disrupted (Figure 4-6B). Analysis of the XyG oligosaccharide profile of this mutant by HPAEC-PAD and MALDI-TOF analysis of XEG-released oligosaccharides (Figure 4-7, 2.5.3, 2.5.4) revealed that the oligosaccharides lack galactosylation at the second position of the XXXG-motif. The gene was therefore named *XLT2* for *Xyloglcuan L-side-chain Transferase at position 2*. A complementation line, in which the wild type *XLT2* gene was transformed into the *xlt2* mutant (Figure 10-5), showed a restoration of galactosylation of XyG at the second position (Figure 4-7). The structural XyG phenotype of *xlt2* is thus distinct from the phenotype of *mur3* mutants, which are deficient for galactosylation at the third position (Madson et al. 2003). A double mutant between *xlt2* and *mur3.1* created by crossing was demonstrated to have XyG that was nearly devoid of galactosylation and consisted mainly of XXXG units (Figure 4-7).

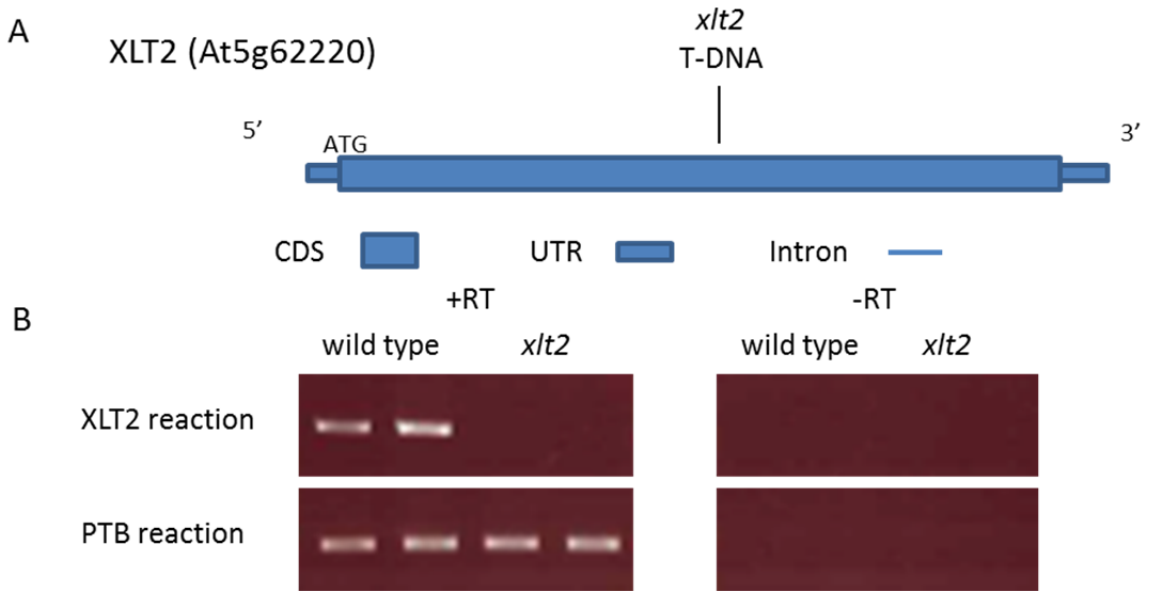


Figure 4-6. RT-PCR of an *A. thaliana* insertional mutant for At5g62220. (A) the *xlt2* mutant contains a T-DNA insertion 835 bp into the coding sequence . (B) RT-PCR (for primers see appendix 3, 4). Primers spanned the insertion site. No transcript was detected for the xlt2 T-DNA line, demonstrating a knockout-out allele. Expression of the PTB gene was used as a control for RNA extraction and cDNA synthesis. The reactions were done in duplicate for each genotype, with (+RT) and without (-RT) reverse transcriptase.

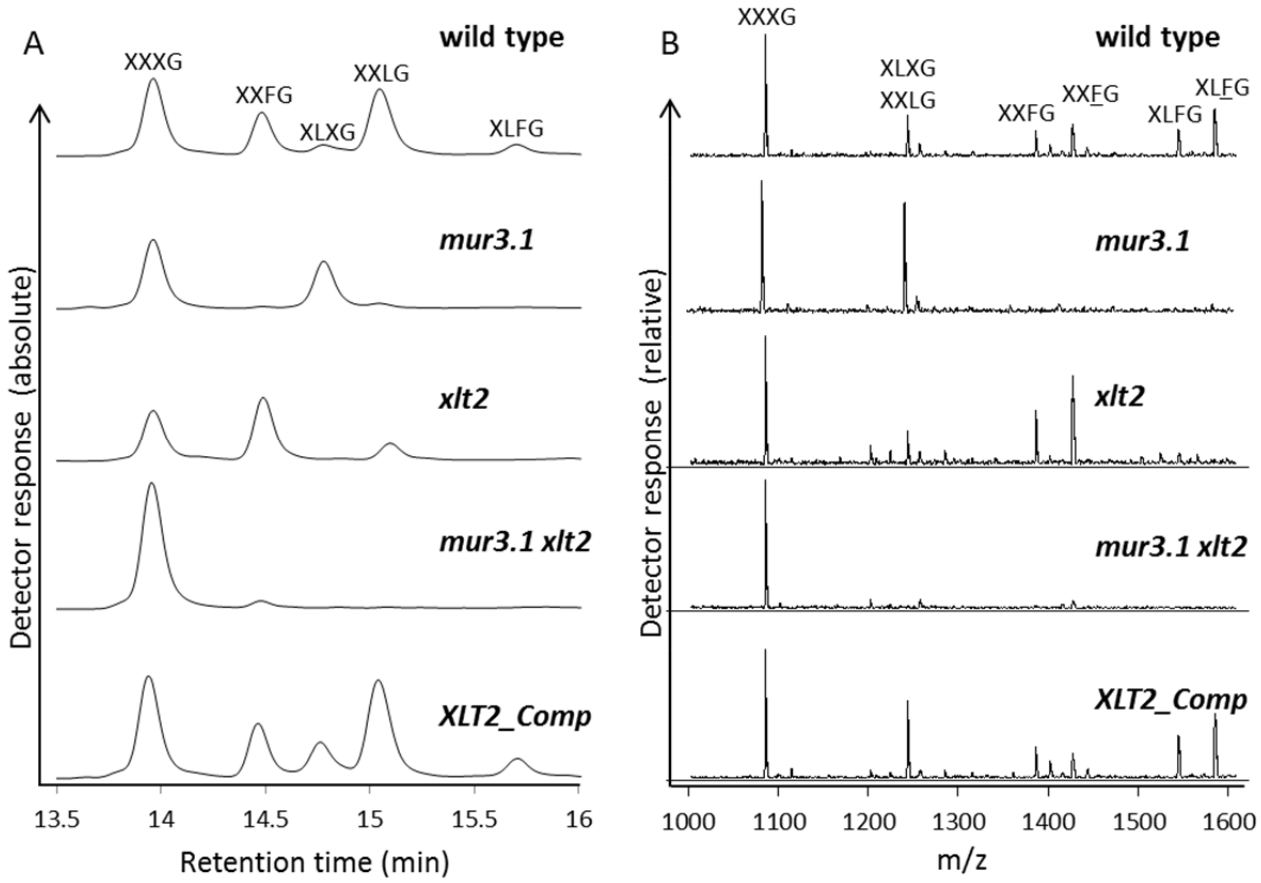


Figure 4-7. XyG structural analysis of *A. thaliana* mutants. (A) HPAEC-PAD and (B) MALDI-TOF analysis of oligomers released from leaf wall material by XEG. The *XLT2_Comp* line is the *xlt2* mutant transformed with the *XLT2* gene (including promoter, coding sequence, and terminator). The *xlt2* mutant was found to lack galactosylation at the second position of the XXXG motif, whereas the *mur3.1 xlt2* double mutant was nearly devoid of XyG substitution. Peak assignments were based on m/z, retention time, and comparison to standards (Gunl et al. 2011; Madson et al. 2003) (XXXG from Megazyme). For XyG oligosaccharide nomenclature see Figure 1-4.

4.2.5 Growth phenotypes

To assess the impact of XyG substitution pattern on plant growth, plant height and branch number of the XyG mutants were quantified. The growth habit of the *xlt2* mutant was found to be the same as the wild type plants (Figure 4-8). The *mur3.1* mutant was shorter than wild type but did not have as strong of phenotype as the double mutant. The *mur3.1 xlt2* line was found to have a decrease of approximately 30% in height and an increase of 80% in branch number, indicating that XyG substitution pattern is important for normal plant growth and development (Figure 4-8).

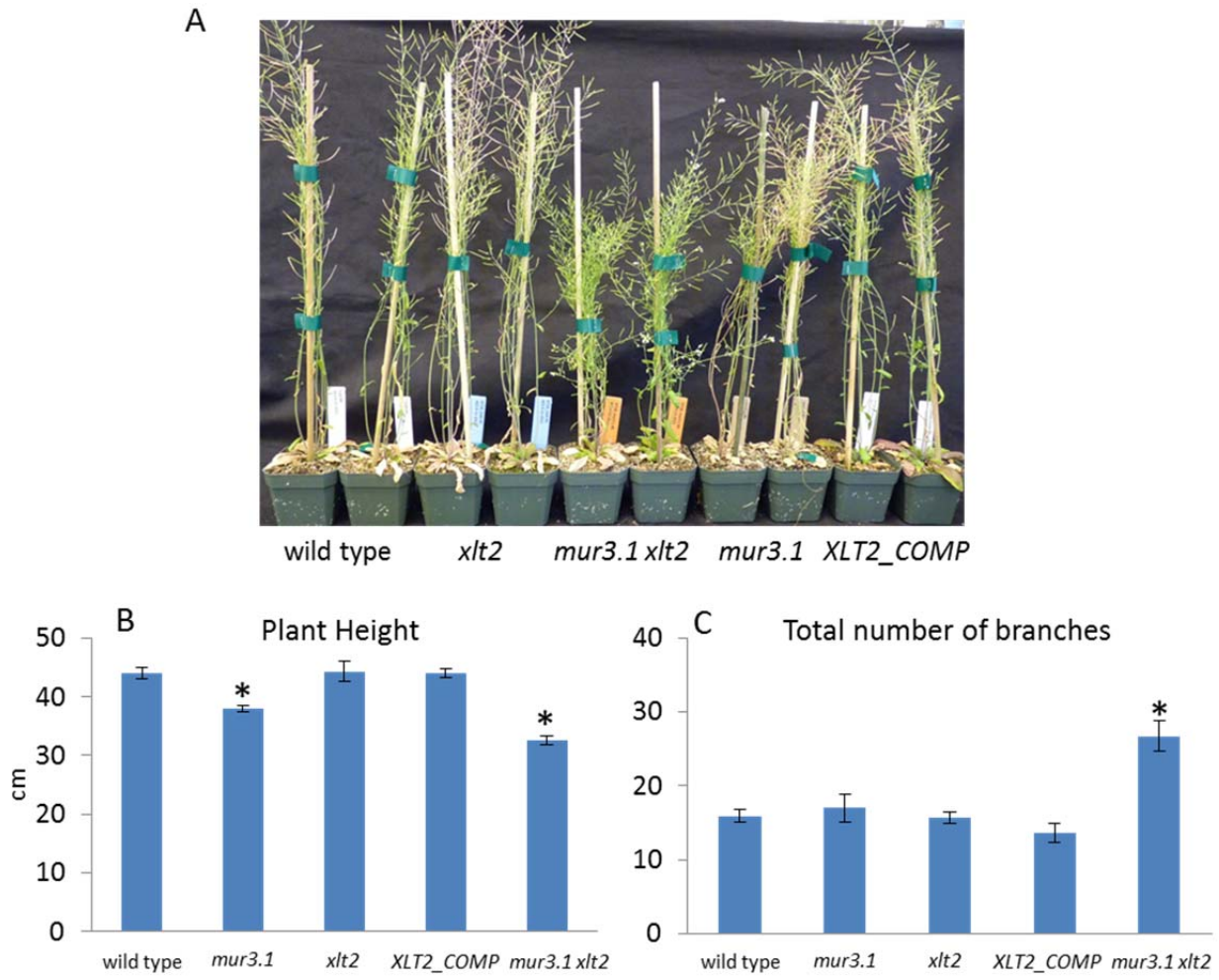


Figure 4-8. Growth phenotype of *A. thaliana* mutants. (A) Picture showing decreased plant height of the *mur3.1 xlt2* double mutant at senescence (seven weeks of growth). (B) Quantification of plant height and (C) total number of branches. Error bars indicated standard deviation. Asterisks statistically significant difference compared to wild type (Col0) plants by Student's t-test (p-value < 0.01). n > 6.

4.2.6 XLT2 subcellular localization

To determine the subcellular localization of the XLT2 protein, a C-terminal GFP gene fusion was constructed (Figure 10-4, 2.2.5) and transiently expressed in *N. benthamiana*. The GFP signal colocalized with a CFP Golgi marker, indicated that the XLT2 protein is localized to the Golgi apparatus (Figure 4-9, 2.3.1).

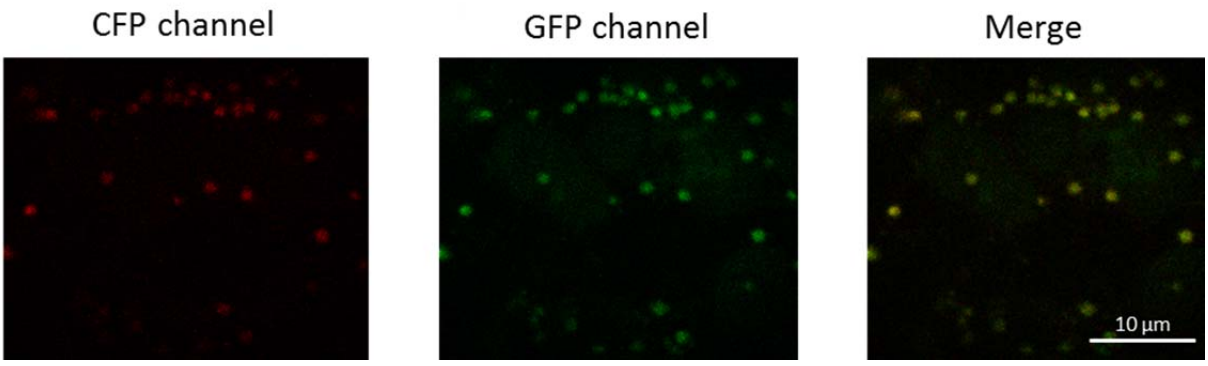


Figure 4-9. Subcellular localization of XLT2. XLT2 with a C-terminal GFP tag was transiently expressed in *N. benthamiana* with a Mannosidase-CFP Golgi-marker. CFP channel representing the Golgi marker and the GFP channel representing XLT2 localization plus the merged image are shown.

4.3 Discussion

4.3.1 Efficacy of the transcriptional profiling approach

Transcriptional analysis of developing nasturtium seeds provided a set of contigs with the relative transcript abundance at several time points of seed development. Among contigs with homology to known GTs several putative orthologs to XyG biosynthetic genes (*CSLC4*, *XXT2*, and *MUR3*) were found to be induced and an additional gene involved in XyG galactosylation (*XLT2*) was identified (Figure 4-4, Figure 4-7). However as these genes were induced at much lower levels than many genes not related to XyG biosynthesis (Table 4-3, Figure 4-4), the approach was limited at identifying XyG biosynthetic genes based on additional data on gene homology. Therefore, genes from families not previously implicated in polysaccharide biosynthesis or metabolism could probably not be identified using this approach. As many genes not related to XyG biosynthesis are strongly expressed in developing seeds, it would be advantageous to perform the same transcriptional profiling experiment on multiple closely related species, some of which accumulate XyG as a seed storage polymer and some of which do not. This would allow for filtering out some of the genes not related to XyG biosynthesis and may allow from the identification of genes involved in XyG biosynthesis from gene families not previously implicated in this process such as transcription factors or kinases.

4.3.2 XLT2 (At5g62220) represents a putative XyG galactosyltranferase

XLT2 has high sequence similarity (59% over a stretch of 390 amino acids) to the XyG galactosyltransferase MUR3 and both belong to CAZY-family 47 clade A (Li et al. 2004). The enzyme activity of MUR3 has been demonstrated *in vitro* with a purified protein (Madson et al. 2003). MUR3 specifically adds a galactosyl residue to the *O*-2 of the xylosyl unit at the third position of the XXXG motif. As the *xlt2* mutant is specifically lacking galactosylation at the second position of the XXXG motif (Figure 4-7), it is likely that XLT2 represents the hitherto elusive galactosyltransferase with specificity for the *O*-2 position of the second xylosyl group of XXXG. The amount of XXLG also decreased in the *xlt2* mutant, but as this mutant has a

concomitant increase in XXFG the total amount of oligosaccharides containing a galactosyl group at the third position remained approximately constant. This suggests that the presence or absence of a galactosyl group at the second position does not affect the activity of MUR3. It is not clear if the increase in the ratio of XXFG to XXLG in the *xlt2* mutant is the result of an increase in fucosyltransferase activity, increased substrate availability, decreased fucosidase activity, or some other mechanism. If the function of XyG substitution is to regulate polymer solubility (Chambat et al. 2005), the increased fucosylation in this mutant may be the result of a compensatory mechanism for the loss of the galactosyl groups at the second position. Localization of an XLT2-GFP fusion to the Golgi apparatus (Figure 4-9), the site of XyG biosynthesis (Moore and Staehelin 1988), is consistent with XLT2 being involved in this process.

4.3.3 XyG structure and growth phenotype of the *mur3.1 xlt2* double mutant

The *mur3.1 xlt2* double mutant plants have significantly decreased plant height and a greater number of branches at maturity than either the wild type or the single mutant plants (Figure 4-8). The *mur3.1* plants show a less severe but significant decrease in plant height compared to the wild type (p-value < 0.001). The mechanism responsible for these growth phenotypes is not clear. One possible explanation is that a lack of XyG substitution could result in an aberrant organization of the extracellular matrix with decreased capacity for being remodeled during cell elongation, therefore resulting in decreased plant growth. This hypothesis is supported by the observation that the organization of cellulose microfibrils is disrupted in the *xxt1 xxt2 A. thaliana* double mutant lacking detectable XyG (Anderson et al. 2010). An alternative explanation is that a lack of XyG galactosylation could lead to aggregation of the polysaccharide during biosynthesis, potentially disrupting the structure of the endomembrane system and leading to pleotropic effects on plant growth. Disrupted Golgi structure has been observed in *mur3* T-DNA lines, which also have a dwarf phenotype (Tamura et al. 2005). However, the phenotype of the *mur3* T-DNA lines was reported to be the result of a function of the MUR3 protein in cytoskeleton organization independent from its role in XyG biosynthesis (Tamura et al. 2005). This was supported by the fact that MUR3 function was completely disrupted in the *mur3.1* mutant (a missense allele) but its Golgi structure was not affected. However, careful analysis of the *mur3.1* mutant reveals that there is still galactosylation at the third position of the XXXG motif that is not observed in the *mur3* T-DNA line (Figure 13-7). The disruption of Golgi structure present in *mur3* T-DNA but not in *mur3.1* could therefore be the result of a lack of XyG galactosylation. Disrupting the structure of the Golgi, which is the site of hemicellulose, pectin, glycoprotein, and extensin biosynthesis and trafficking (Basu et al. 2013; Lerouxel et al. 2006; Porchia et al. 2002; Sterling et al. 2001; Velasquez et al. 2012), could have a negative impact on plant growth by disrupting these other processes. Additional experiments to investigate the relationship between XyG substitution and plant growth were performed in the second part of this chapter but future work is needed to understand this phenotype at a mechanistic level.

5 Identification of a XyG arabinofuranosyltransferase from tomato

5.1 Background

XyG found in *A. thaliana* has an XXXG-type backbone xylosylation pattern and additional galactosyl and fucosyl substitutions (Figure 1-2). This form of XyG appears to be the type most widely distributed among land plant lineages (Hoffman et al. 2005). However, some plant families or orders have XyG with alternative substitution patterns. In *A. thaliana*, the galactosyl units are linked with a β linkage to the *O*-2 position of the xylosyl residues. In other plant species, arabinofuranose, arabinopyranose, galacturonic acid, and xylose have all been detected at this position (Figure 1-4). This structural diversity is presumably due to specific GTs present in these species. However, the GTs responsible for XyG structural diversity in species other than *A. thaliana* have not been identified. The identification of such genes would allow for experiments to better understand the evolution of this GT family, the function of XyG substitution pattern *in vivo* by genetically manipulating XyG structure including impact on plant growth and wall mechanics, and for tailoring polysaccharide structure in ways suited for particular applications.

A comparative genomics approach was used to identify candidate GTs responsible for XyG structural diversity. To facilitate this approach, relatively closely related organisms that have genome sequences available and different XyG substitution patterns were needed. *A. thaliana* and tomato were selected based on these criteria. Compared to *A. thaliana*, the dominant form of XyG in tomato has a reduction in the degree of xylosylation leading to an XXGG-type repeating motif (Figure 5-1). The xylosyl residues of tomato XyG can be substituted at the *O*-2 position with arabinofuranose, a substitution pattern abbreviated as “S”, or galactopyranose “L” (Jia et al. 2003; York et al. 1996). Arabinofuranosyl moieties on the *O*-2 of a xylosyl residuce have also been found in the XyG from *Olea europaea* (Vierhuis et al. 2001), *Nerium oleander* (Hoffman et al. 2005), *Ceratoperis richardii* (Pena et al. 2008) and possibly *Tamarindus indica* (Niemann et al. 1997). While *Ceratoperis richardii* is a monilophyte that is far from the Solanales lineage, *Olea europaea* and *Nerium oleander,* like tomato, are both Asterids and therefore there may be a common origin for XyG arabinosylation in these species and their respective orders.

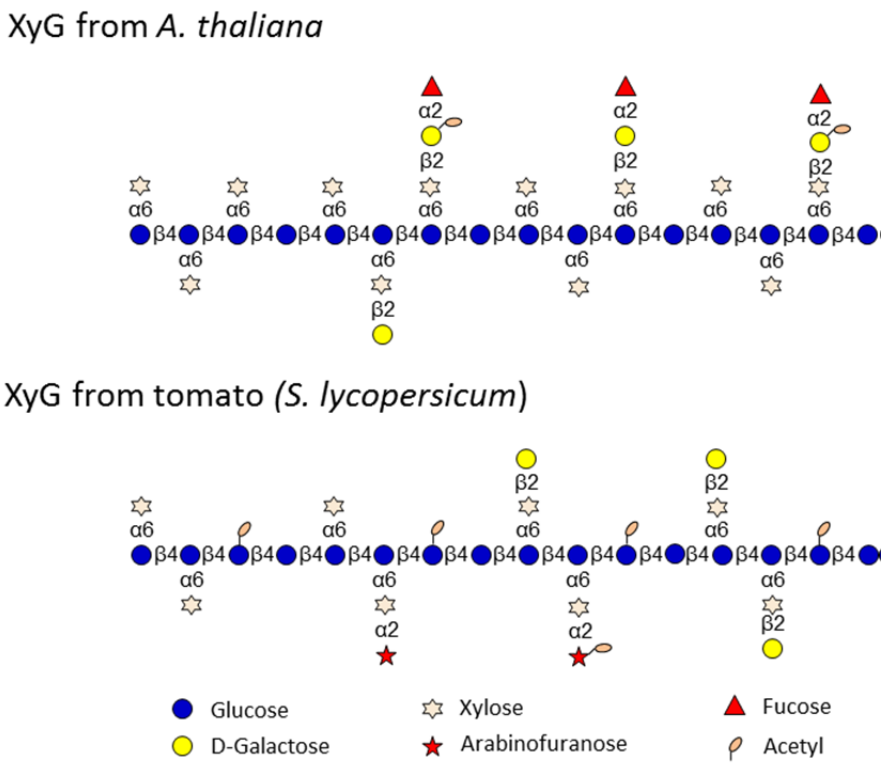


Figure 5-1. XyG structures of *A. thaliana* and *S. lycopersicum*. Compared to *A. thaliana*, tomato contains XyG with reduced xylosylation (XXGG motif), *O*-acetyl-substituents on the glucan-backbone, and arabinofuranosyl moieties in place of some of the galactosyl units.

Like the galactosyl and galacturonosyl residues of the L and Y side-chains, the arabinofuranosyl residue of the S side-chain resides at the *O*-2 position of the xylosyl residue. The genes responsible for generating the L (*XLT2* and *MUR3*) and Y (*XUT1*) side-chains are from the same subclade of GT family 47 (Li et al. 2004; Pena et al. 2012) (Figure 5-2). Several GT47s from outside of this subclade have been studied previously including IRX7, IRX10, and IRX10L, which are involved in xylan biosynthesis and XGD1, ARAD1 and NpGUT1, which are involved in pectin biosynthesis. IRX7 is believed to transfer a xylosyl group onto a rhamnose moiety as part of the biosynthesis of the reducing end group of xylan (Brown et al. 2005; Scheller and Ulvskov 2010). IRX10 and IRX10L are both involved in transferring a xylosyl group to elongate the xylan backbone (Brown et al. 2009; Wu et al. 2009). XGD1 is believed to be a β-1,3 xylosyltransferase onto oligogalacturonides (Jensen et al. 2008) and ARAD1 is a putative α-1,5 arabinosyltransferase for arabinan biosynthesis (Harholt et al. 2006). NpGUT1, a putative ortholog of IRX10, was initially believed to be a putative glucuronosyltransferase for RGII (Iwai et al. 2002), however it is able to functionally complement the *A. thaliana irx10 irx10-L* double mutant suggesting it actually has a role in xylan biosynthesis (Wu et al. 2009). The diversity of donor substrate specificity present in this family of GTs suggests that particular members of the GT47 subclade containing MUR3, XLT2 and XUT1 may be good candidate genes for XyG GTs with novel donor specificities acting at the *O*-2 position of the xylosyl residue.

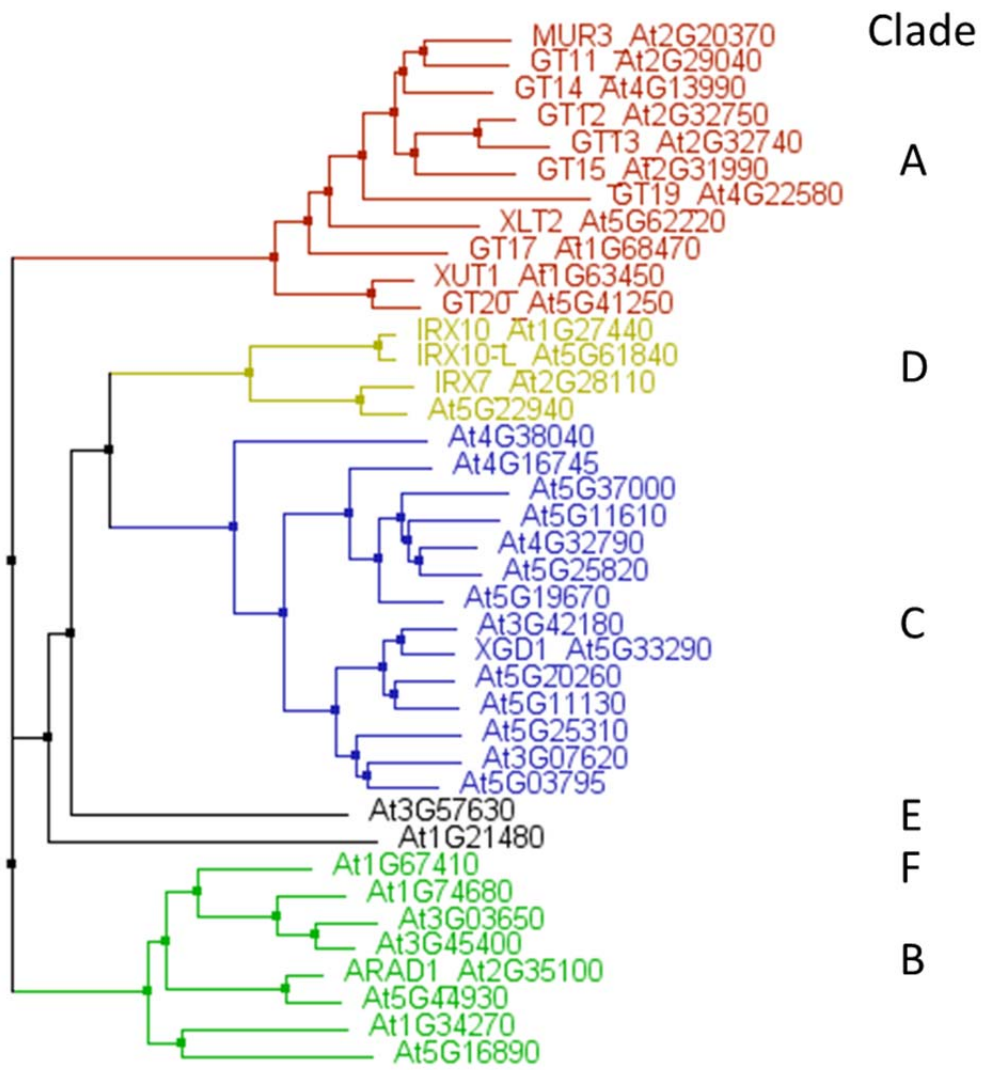


Figure 5-2. Phylogenetic tree of GT47 sequences from *A. thaliana*. There are 39 GT47 proteins in *A. thaliana*. These were aligned and a phylogenetic tree was constructed. The previously established clade names (A-F) for this family are shown (Li et al. 2004). MUR3 (Madson et al. 2003), XUT1 (Pena et al. 2012) and XLT2 (4) have all been shown to be involved in XyG substitution and are in the same clade (clade A) of GT47.

In order to identify a hitherto elusive XyG arabinosyltransferase a functional complementation approach was pursued by expressing tomato candidate genes into the *A. thaliana mur3.1 xlt2* double mutant. This work was published in the journal Plant Physiology with the title "The identification of two arabionsyltransferases from tomato reveals function equivalency of xyloglucan side-chain substituents" by the authors Schultink A, Cheng K, Park YB, Cosgrove DJ, and Pauly M in July 2013 (Schultink et al. 2013).

5.2 Results

5.2.1 Identification of candidate XyG arabinofuranosyltransferases from tomato

A BLAST search identified ten tomato proteins from the XLT2/MUR3 subclade of GT family 47 (Figure 5-3). One of the proteins, Sl09064470, is most similar to the *A. thaliana* MUR3 (AtMUR3) and three proteins (Sl02g092840, Sl07g044960, and Sl07g049610) are most similar to *A. thaliana* XLT2 (AtXLT2) (Figure 5-3). The four corresponding genes were cloned (Figure 10-6, 2.2.5) and transformed into the *A. thaliana mur3.1 xlt2* double mutant (2.2.11). Quantitative RT-PCR was used to verify transgene expression in the resulting T2 lines (Figure 5-4, 2.2.13). All the transgenes were expressed with the level of expression varying up to 20 fold between the lines.

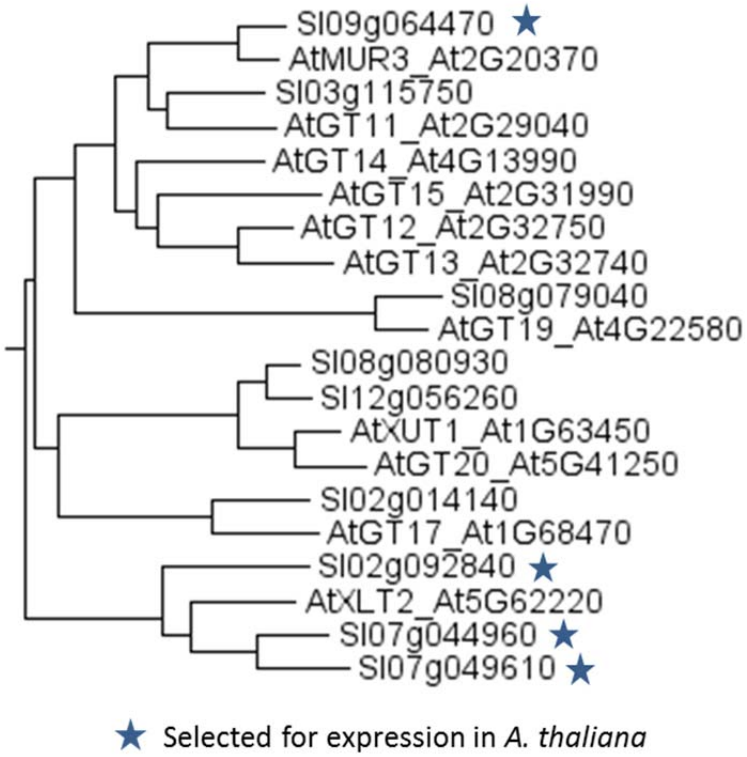


Figure 5-3. Identification of candidate XyG arabinofuranosyltransferases from tomato. A BLAST search was performed to identify GT47 proteins from tomato. A phylogenetic tree was constructed with the *A. thaliana* (At) and resulting tomato proteins (Sl) from the MUR3/XLT2 subclade. One proteins was found to be most similar to MUR3 (Sl09g064470) and three most similar to XLT2 (Sl02g092840, Sl07g044960, and Sl07g049610). The four corresponding genes were selected for expression in *A. thaliana*.

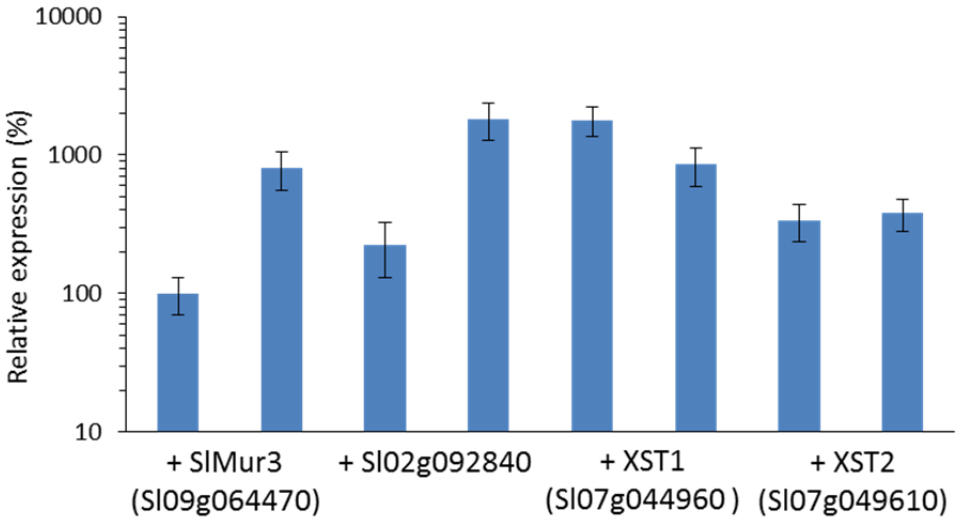


Figure 5-4. Relative quantification of tomato GT expression in transgenic *A. thaliana*. Candidate genes were expressed in the *A. thaliana* double mutant *mur3.1 xlt2*. Two transformed lines were selected for each and qRT-PCR was performed on T2 plants to check the level of transgene expression. All transgenes were detected in the transformed *A. thaliana* lines, with a twenty-fold variation in transcript abundance detected between the most highly and lowly expressed transgenes. Sl09g064470, Sl07g044960, Sl07g049610 were named SlMUR3, XST1 and XST2 respectively based on their effect on XyG structure (described below).

5.2.2 Effect of expression of tomato genes on XyG structure

The XyG structure of the transformed plants was analyzed by XEG digestion of cell wall material followed by MALDI-TOF MS and HPAEC-PAD (2.5.3, 2.5.4). While the untransformed double mutant contains primarily of XXXG units, expression of *Sl09g064470* in *A. thaliana* resulted in the presence of XXLG, XXFG and XXFG as well as minor amounts of XLFG and XLFG (Figure 5-5, Figure 5-6). This XyG structure resembles that of the *xlt2* single mutant, demonstrating that Sl09g064470 harbors MUR3 activity when expressed in *A. thaliana*. This protein, being a putative ortholog to AtMUR3 based on sequence similarity and function, was therefore named SlMUR3.

Expression of *Sl02g092840* in the *A. thaliana mur3.1 xlt2* double mutant did not result in any observed changes to XyG structure compared to the untransformed double mutant (Figure 5-5, Figure 5-6). This suggests that *Sl02g092840* is either not involved in XyG biosynthesis or that the appropriate donor or acceptor substrates are not present in *A. thaliana*. Expression of either *Sl07g044960* or *Sl07g049610* resulted in the production of a novel XEG-releasable oligosaccharide as revealed by MALDI-TOF and HPAEC-PAD (Figure 5-5, Figure 5-6). This oligosaccharide had an m/z of 1217, consistent with an oligosaccharide with four hexoses and four pentoses, and was more abundant in lines expressing *Sl07g044960* than *Sl07g049610*. As the XyG in the *mur3.1 xlt2* double mutant consists almost entirely of repeating XXXG units, which have three pentoses and four hexoses, this data suggests that the novel oligosaccharide

consists of XXXG plus one additional pentosyl substituent. To elucidate the precise structure of the novel oligosaccharide, it was purified using HPAEC-PAD for further analysis (Figure 5-6, 2.5.6).

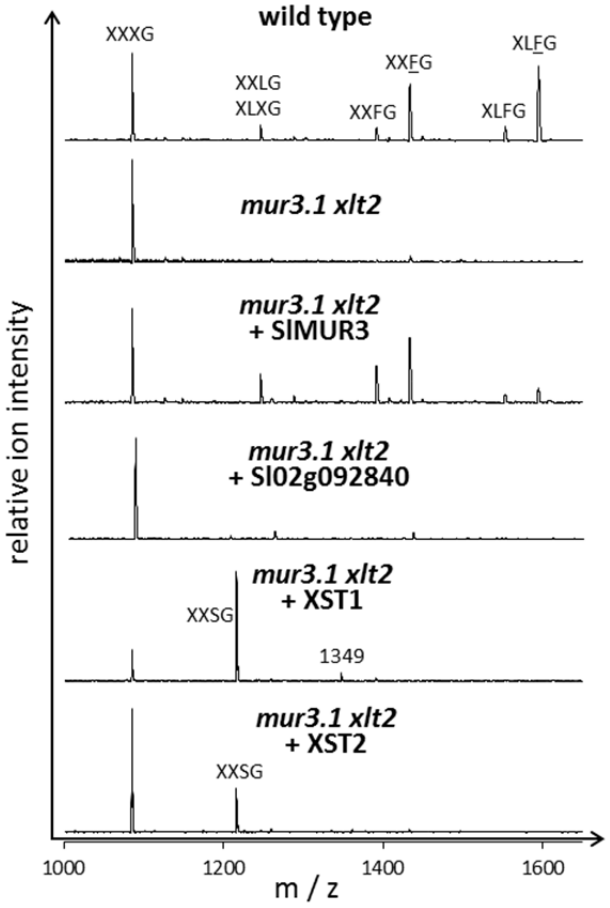
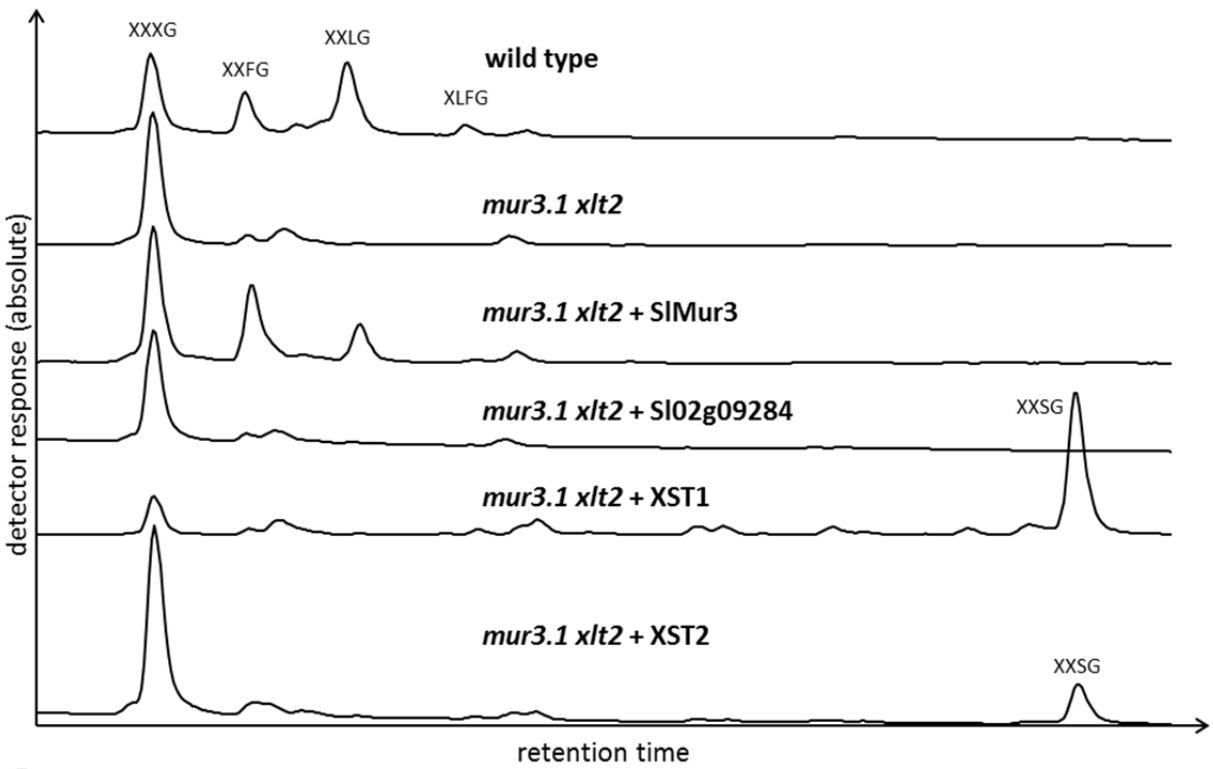


Figure 5-5. XyG OLIMP of *A. thaliana* expressing tomato GTs. The expression of *SlMUR3* resulted in the production of XXFG and XXFG, thereby complementing the *mur3* phenotype of the mutant. Small amounts of XLFG and XLFG were also observed in this line. Two genes, *XST1* and *XST2*, resulted in the production of a novel oligosaccharide with an m/z of 1217. This m/z is consistent with an oligosaccharide containing four hexoses and four pentoses. A small peak at 1349 could correspond to an oligosaccharide containing four hexoses and five pentoses. The expression of *Sl02g092840* did not result in an observed change in the XyG structure using this method of analysis.

A



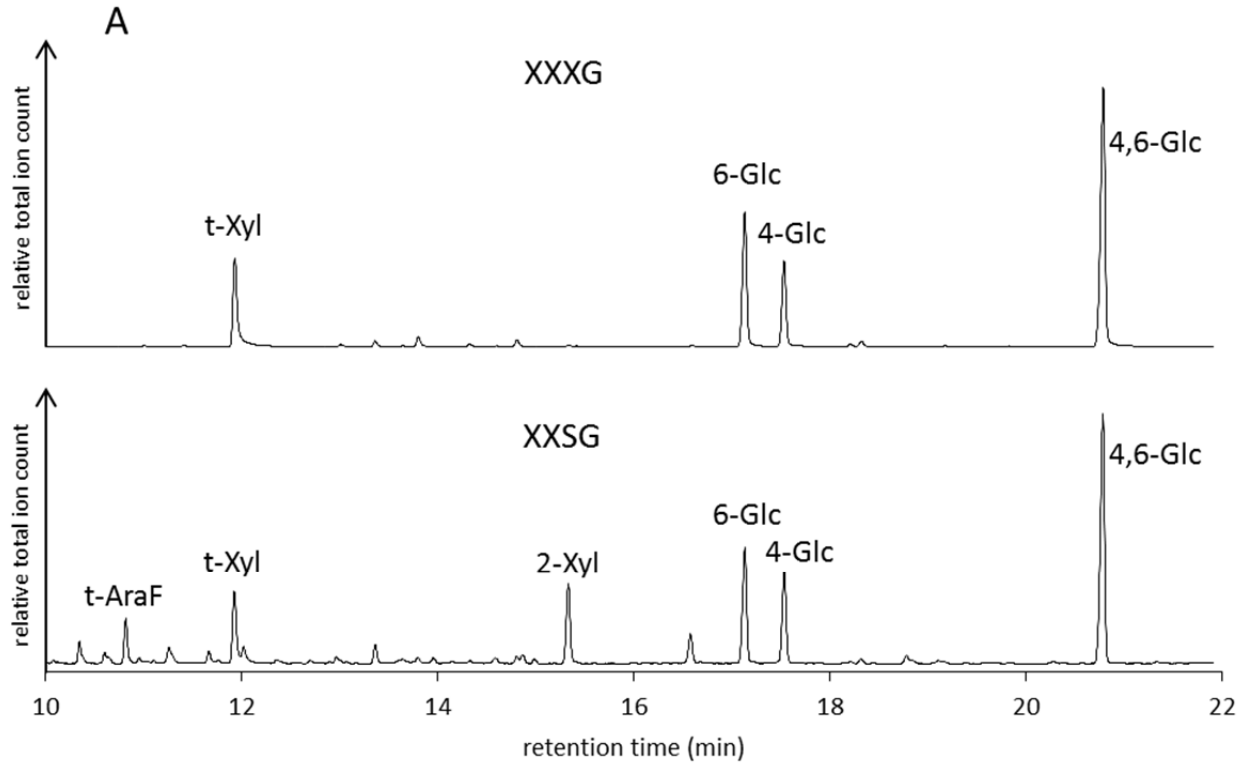
B

	XXXG	XXFG	XLXG	XXLG	XLFG	XXSG
wild type	184 ± 25	81 ± 4	9 ± 1	161 ± 24	21 ± 3	nd
mur3.1 xlt2	246 ± 33	12 ± 2	nd	nd	nd	nd
+ SlMur3	271 ± 17	143 ± 22	nd	68 ± 6	3 ± 1	nd
+ Sl02g09284	274 ± 35	18 ± 5	nd	nd	nd	nd
+ SlXST1	70 ± 12	8 ± 2	nd	nd	nd	269 ± 48
+ SlXST2	499 ± 107	26 ± 2	nd	nd	nd	107 ± 26

Figure 5-6. HPAEC-PAD of XyG oligosaccharides from *A. thaliana* expressing tomato GTs. The novel oligosaccharide, assigned as XXSG based on further characterization described below, in the *XST1* and *XST2* lines eluted several minutes after the other XyG oligosaccharides. The fraction containing the novel oligosaccharide was collected for further analysis. (B) quantification of the XyG oligosaccharides (ng / mg AIR) was performed by comparing the detector response to that for a standard of XXXG. nd – not detected.

5.2.3 Structural characterization of the novel XyG oligosaccharide

Glycosidic linkage analysis was performed on the purified oligosaccharide and an XXXG standard using the partially methylated alditol acetates procedure (2.5.7). For XXXG, the linkages detected were 4-glucose, 4,6-glucose, 6-glucose and terminal-xylose which is consistent with the known structure of this oligosaccharide (Figure 5-7). The ratio observed for the total ion counts for each partially methylated alditol acetate derivative (1.0:3.1:1.5:1.1, respectively) did not precisely match the expected molar ratio for these linkages (1:2:1:3). This deviation could be the result of differential efficiencies in ionization between the various derivative species or degradation of sugars during the derivatization process. For the unknown oligosaccharide, the same linkages were detected along with 2-xylose and terminal-arabinofuranose. This result is qualitatively consistent with the novel oligosaccharide containing an arabinofuranosyl residue sitting at the *O*-2 position of one of the xylosyl residues. Thus, the structures XXSG, XSXG, and SXXG are all consistent with this linkage pattern and the m/z of 1217.



B

	XXXG	XXSG
t-Ara	nd	5%
t-Xyl	18%[1]	8%
2-Xyl	nd	9%
6-Glc	21%	13%
4-Glc	14%	11%
4,6-Glc	47%	31%
total	100%	76%[2]

Figure 5-7. Glycosidic linkage analysis of the novel XyG oligosaccharide. (A) Chromatograms (total ion count) of the partially-methylated alditol acetates separated and detected using GC-MS. The novel oligosaccharide was found to contain t-arabinofuranose and 2-xylose in addition to the t-xylose, 4-glucose, 6-glucose and 4,6-glucose present in XXXG. (B) quantification of the peak areas from (A). [1]Percent of total peak area. [2]The remaining peaks were not assignable to known glycosidic linkages. nd – not detected.

Oligoxyloglucan reducing end-specific xyloglucanase (OREX) is an enzyme that specifically hydrolyzes XyG oligosaccharides into cellobiose backbone units, but only if the second position is not substituted (Bauer et al. 2005; Yaoi and Mitsuishi 2002). A purified OREX expressed in *P. pastoris* (2.5.12) (Bauer et al. 2005) was used to digest the unreduced purified novel

oligosaccharide. The digestion resulted in the disappearance of the 1217 ion and the appearance of an ion with an m/z of 629 (Figure 5-8). This ion has an m/z consistent with that expected for an oligosaccharide two pentoses and two hexoses, suggesting that the novel oligosaccharide structure was digested into two oligosaccharides with each oligosaccharide product having this composition. Consistent with this would be the two fragments XX and SG resulting from the digestion of XXSG. If the arabinosyl-residue were on the first or second position, the resulting cellobiosyl units would have been SX / XS and XG. Ions representing these oligosaccharides were not observed, and digestion of the oligosaccharide XSXG would not be expected as OREX is unable to hydrolyze the similar XLXG (Bauer et al. 2005).

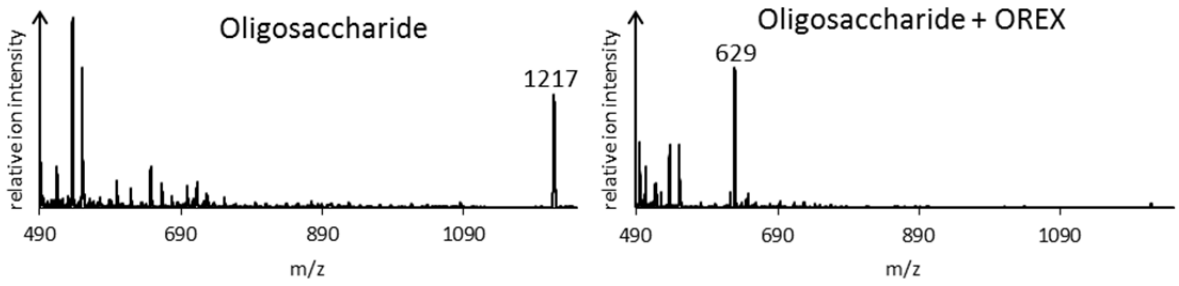
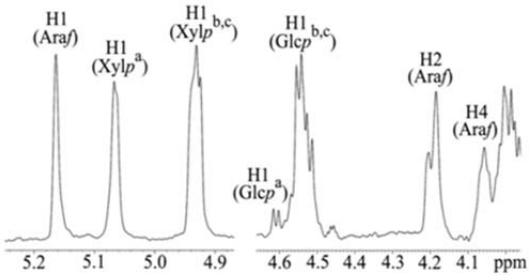


Figure 5-8. OREX digestion of the novel oligosaccharide. The ion from the oligosaccharide (m/z 1217) was converted to an ion(s) with m/z 629 after OREX digestion. An m/z of 629 is consistent with an oligosaccharide having two pentoses and two hexoses. This result in consistent with the novel oligosaccharide having a structure of XXSG and being digested to XX and SG (both with m/z of 629).

NMR analysis was used to independently determine the structure of the novel XyG oligosaccharide (performed by Kun Cheng, Pauly lab, 2.5.11). The proton NMR spectra matched that of XXSG previously characterized from olive pulp (Figure 5-9) (Vierhuis et al. 2001). The genes *Sl07g044960* and *Sl07g049610* were therefore named *XyG "S"-side-chain Transferase* (*XST1*) and *XST2*, respectively. To further confirm the structure of the novel oligosaccharide as XXSG, an XEG digest of cell wall material of *XST1*-expressing *A. thaliana* was compared to an XEG digest of XyG extracted from olive fruit (Figure 5-10). The two most abundant XyG oligosaccharides in the olive digest were found to have m/z's of 1379 and 1217, consistent with the structures XLSG and XXSG previously reported from this tissue (Vierhuis et al. 2001). The peak corresponding to XXSG from olive eluted at the same time as the oligosaccharide from *XST1*-expressing *A. thaliana* (Figure 5-10, 2.5.8).

A



B

	H1	H2	H3	H4	H5
Ara	5.159	4.193	3.938	4.062	3.843/ 3.706
Xyl[a]	5.082	3.572	3.770	3.672	3.570/ 3.728
Xyl[b]	4.946 (3.2)	3.546	3.732	3.627	3.560/ 3.710
Xyl[c]	4.932	3.541	3.710	3.618	3.540/ 3.690
Glc[a]	4.610 (8.0)	3.412 (n.d.)	n.d.	n.d.	n.d.
Glc[b]	4.547 (7.9)	3.381 (10)	3.659	3.683	3.802
Glc[c]	4.520 (8.0)	3.327 (10)	3.51	3.52	3.619

C

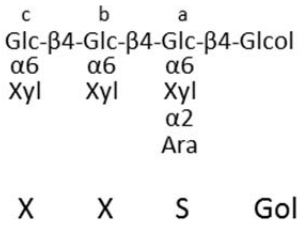


Figure 5-9. NMR of the oligosaccharide. The oligosaccharide was purified by HPAEC-PAD (see Fig. 14) and subjected to analysis by NMR. (A) anomeric region of NMR spectra indicated the presence of arabinofuranosyl-residue on the xylose at the third position of the oligosaccharide. (B) Assignments of the proton chemical shifts match a previously published spectra for the XXSG oligosaccharide derived from olive pulp (Vierhuis et al. 2001). The 3J for scalar interaction with the next higher numbered proton is shown in parentheses. Superscripts refer to the proposed structure in (C). n.d. – not determined. Monosaccharide abbreviations Glc: D-Glucopyranose; Xyl:D-Xylopyranose; Gal:D-Galactopyranose; Ara: L-Arabinofuranose.

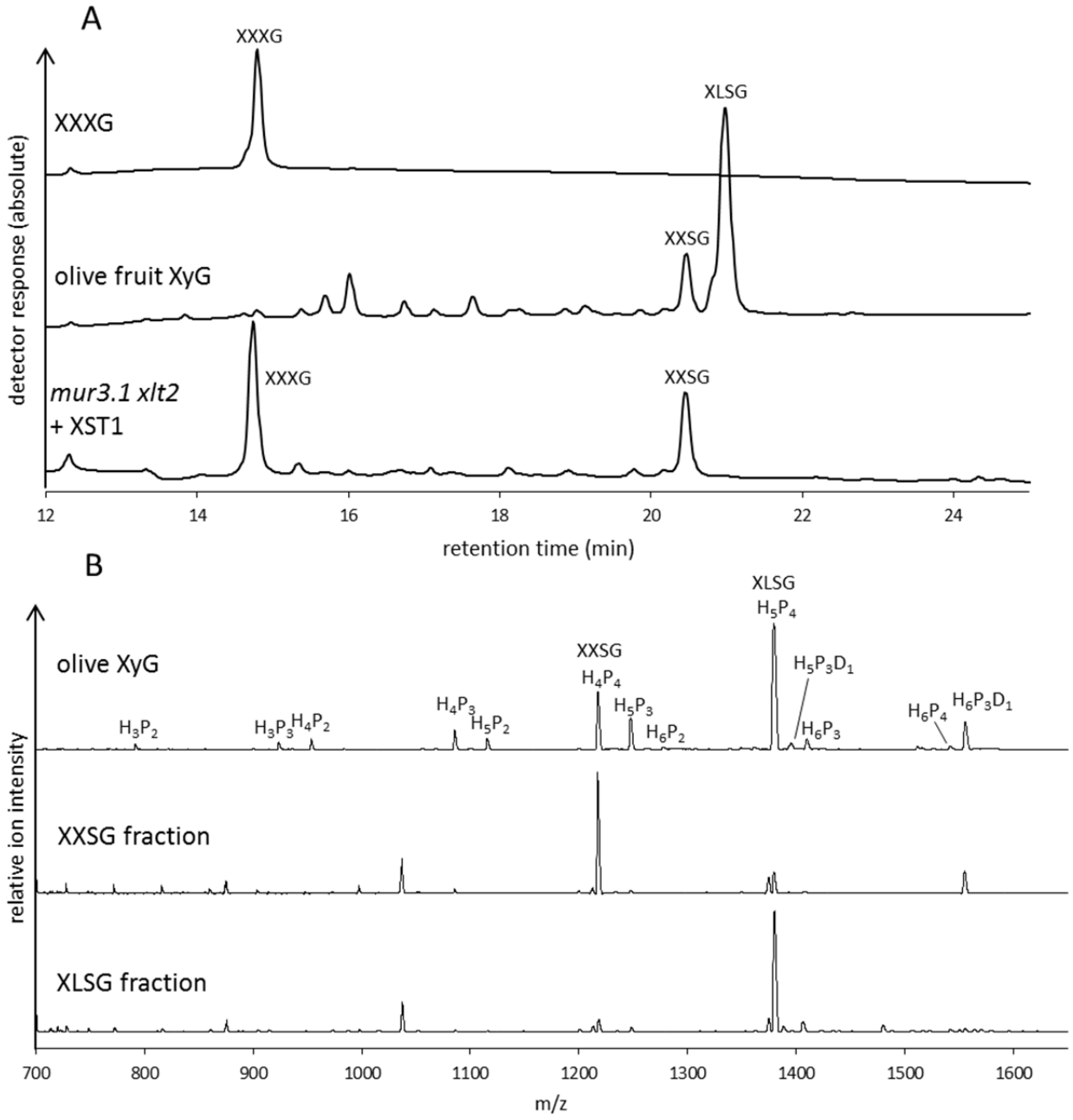


Figure 5-10. Comparison of the novel oligosaccharide to XXSG isolated from olive fruit. (A) HPAEC-PAD of an XXXG standard (Megazyme), an XEG digest of XyG extracted from olive fruit, and an XEG digest of wall material of the *mur3.1 xlt2 A. thaliana* double mutant expressing XST1. The two major peaks visible in the olive chromatogram were collected and by MALDI-TOF MS were observed to have m/z ratios of 1217 and 1379 (B), consistent with XXSG and XLSG respectively – the dominant components of olive fruit XyG (Vierhuis et al. 2001). The two major oligosaccharides present in digest of *A. thaliana* expressing *XST1* matched the retention times of the XXXG standard and the oligosaccharide assigned as XXSG from the digest of olive XyG. (B) From MALDI-TOF MS of the XEG digest of XyG extracted from olive fruit (with 4 M KOH) (B), the

dominant ions have m/z consistent with XXSG and XLSG. Eleven additional observed ions have an m/z consistent with possible XyG oligosaccharides (labeled with possible glycosyl unit composition). H – hexose; P – pentose; D – deoxyhexose.

5.2.4 Expression of XST rescues the growth phenotype of *mur3.1 xlt2*

The *mur3.1 xlt2* double mutant has a dwarfed phenotype compared to wild type or single mutants plants (Figure 4-8). To investigate how the addition of arabinofuranose to XyG affects plant growth, the heights of the *A. thaliana* mutants and lines expressing each of the transgenes from tomato were measured. Expression of *Sl02g092840* resulted in the retention of the dwarfism phenotype, consistent with this gene not resulting in a change of XyG structure (Figure 5-11). Expression of *SlMUR3* resulted in a full rescue of the dwarf phenotype, consistent with the *xlt2* single mutant also having a normal growth habit. The lines expressing *XST1* or *XST2* had a significant increase in plant height, however were significantly shorter than wild type plants for three of the four lines examined (Figure 5-11).

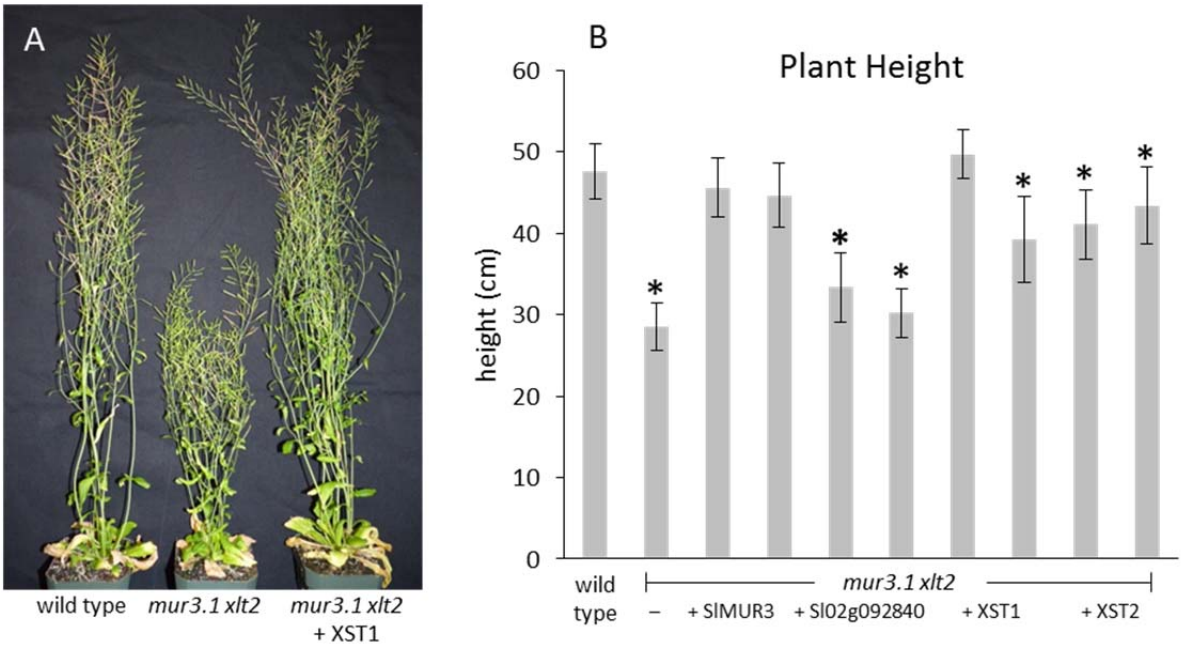


Figure 5-11. Growth phenotypes of *A. thaliana* plants expressing tomato GTs. (A) Picture of representative plants from the wild type, *mur3.1 xlt2*, and *mur3.1 xlt2* + XST1 lines. (B) Quantitative heights of additional lines (two independent for each construct). Expression of *SlMUR3*, *XST1*, or *XST2* resulted in partial or full rescue of the growth phenotype of the *mur3.1 xlt2* double mutant whereas expression of *Sl02g092840* did not. n > 4. * indicates statistical significance compared to wild type by Student's t-test (*p*-value < 0.05). Error bars show standard deviation.

5.2.5 XyG substitution pattern affects cell wall biomechanics

Biomechanical tests on leaf petioles were performed to measure the cell wall properties for the mutant and transgenic lines (by collaborator Yong Bum Park, Dan Cosgrove lab, Penn State, 2.5.13). The *mur3.1 xlt2* double mutant was found to be less extensible in acidic conditions (Figure 5-12), a process mediated by alpha-expansins and possibly XTHs, which also show pH-dependent activity (Cosgrove 2000; Maris et al. 2011). Acid-induced extension of the wall is believed to be an important part of cell elongation and therefore significant for overall plant growth (Cho and Cosgrove 2000; Perrot-Rechenmann 2010). Expression of *XST1* restored acid-induced creep to the level of the wild type. The *mur3.1 xlt2* mutant also had alterations in mechanical compliance, which was measured by stress-strain analyses. The compliance is ability of the wall material to extend as the result of a force being applied. A portion of the extension is reversible (elastic compliance) with the remainder being irreversible (plastic compliance) (Park and Cosgrove 2012a). For both the elastic and plastic compliance, lines lacking a functional *MUR3* gene had a defect whereas those with a wild type *MUR3* gene had wild type levels of compliance (Figure 5-13). Expression of *XST1* was unable to rescue the elastic compliance phenotypes of the *mur3.1 xlt2* mutant.

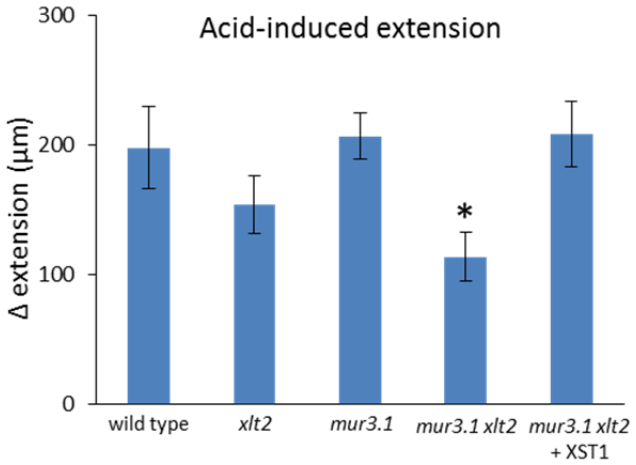


Figure 5-12. Acid-induced creep of leaf petioles. Petioles were placed under constant load first under neutral pH then with acidic buffer. The acid-induced extension was measured. The *mur3.1 xlt2* double mutant had a defect in acid-induced wall creep which was rescued by the expression of *XST1*. * indicates statistically significant difference from wild type using Student's t-test (*p*-value < 0.05). n ≥ 9. Error bars show standard error.

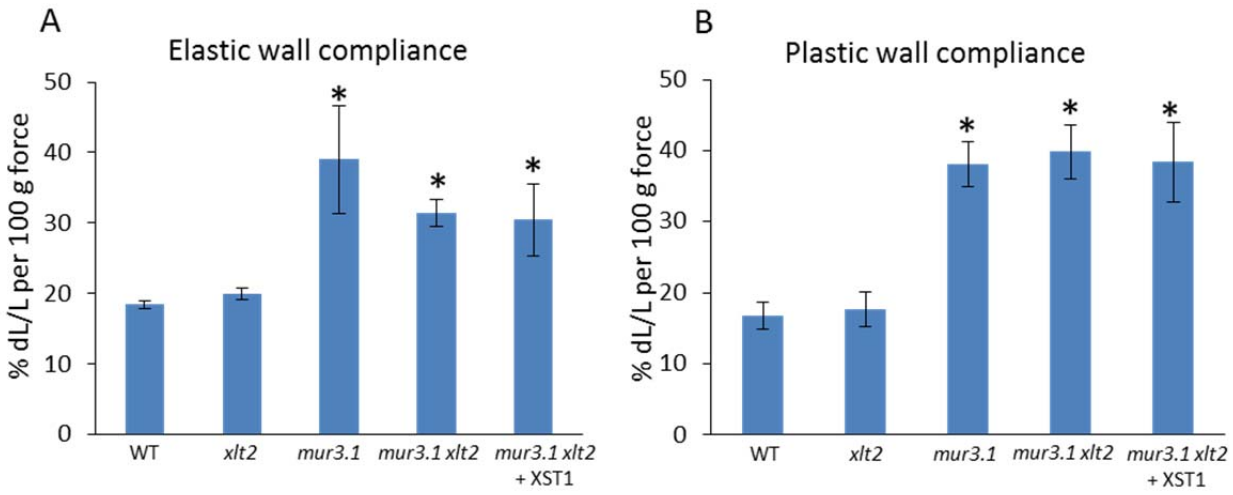


Figure 5-13. Wall compliance of leaf petioles of *A. thaliana* plants with altered XyG. The elastic (A) and plastic (B) wall compliances are shown for several of the mutants and the *XST1*-expressing line. Plants defective for MUR3 showed increased wall compliance. This phenotype was not rescued by expression of *XST1*. n ≥ 8. * indicates statistically significant difference from wild type (*p*-value < 0.05). Error bars indicate standard error.

5.2.6 Susceptibility to powdery mildew

One way in which a plant could potentially gain resistance to a pathogen is to evolve altered, diverse cell wall polymer substitution patterns that are resistant to enzymatic degradation. Fragments of plant cell polysaccharides have also been shown to be inducers of pathogen defense responses by the plant (Hematy et al. 2009). Therefore, altering polysaccharide substitution patterns, which will alter the structure of the released fragments, could affect the defense response of the plant. For these reasons, it was hypothesized that XyG structure may affect the susceptibility of *A. thaliana* to pathogens. *A. thaliana* plants with various XyG structures were exposed to the model pathogen *Eryisphe cichoracearum* (a powdery mildew) by collaborator Heidi Szemenyei from Shauna Somerville's lab. All genotypes tested showed qualitative susceptibility to this pathogen, demonstrating that arabinofuranosyl groups present on the XyG polysaccharide, as well as the loss of XyG galactosylation and fucosylation, are not sufficient to change resistance/susceptibility of this pathogen (Figure 5-14).



Figure 5-14. Images of leaves showing signs of powdery mildew infection. *A. thaliana* Leaves infected with *Erysiphe cichoracearum* (powdery mildew) were specifically selected for imaging, as most leaves for all genotypes showed no sign of infection. The infection rate was too low for reliable quantification. The leaf size is not representative of the leaves for the genotype.

5.3 Discussion

5.3.1 SlMUR3 – a XyG galactosyltransferase from tomato

Expression of *SlMUR3* in the *mur3.1 xlt2* double mutant resulted in the production of XyG with galactosyl and subsequent fucosyl groups at the third position of the XXXG motif (Figure 5-5, Figure 5-6). This is consistent with the activity expected for a protein with MUR3 activity and is similar to the phenotype of the *xlt2* single mutant (Figure 4-7). Based on phylogenetic analysis, SlMUR3 is the most similar protein from tomato to, and a putative ortholog of, AtMUR3 (Figure 5-3). SlMUR3 is therefore likely to be a XyG galactosyltransferase with specificity similar to AtMUR3. Small amounts of doubly galactosylated XyG oligosaccharides were detected in the SlMUR3 expressing lines (XLFG and XLFG, Figure 5-5) suggesting that the activity of SlMUR3 is not completely restricted to the third position. AtMUR3 has not been observed to catalyze the transfer of a galactosyl residue to a position other than the third position, indicating that the specificities of the two enzymes might be slightly different (Madson et al. 2003). The fact that tomato has retained a functional *MUR3* gene despite apparently lacking XyG with the XXXG motif suggests that perhaps SlMUR3 is able to act at the second position of the XXGG motif to form XLGG, a structure that is present in tomato XyG (Jia et al. 2003). Alternatively, while the dominant form of XyG in tomato consists of the XXGG type, there may be tissues or developmental time points in which XXXG-type XyG is produced and the function of a MUR3 enzyme is required for XyG galactosylation.

5.3.2 XST1 and XST2 – putative XyG arabinofuranosyltransferases from tomato

Expression of either *XST1* or *XST2* in the *mur3.1 xlt2 A. thaliana* double mutant resulted in the production of arabinosylated XyG. The arabinosyl group was added to the third position of the XXXG motif, resulting in XXSG oligosaccharides being released upon XEG treatment (Figure 5-5, Figure 5-6). As activity of XST1 and XST2 was not demonstrated *in vitro,* the possibility remains that they may somehow mediate the activity of an endogenous *A. thaliana* XyG arabinosyltransferase. However, because this XyG side-chain has not been previously reported in *A. thaliana* (or any Brassicaceae) and because XST1 and XST2 share significant similarity with

AtMUR3, which has been demonstrated to have GT activity *in vitro* (Madson et al. 2003), it is likely that both XST1 and XST2 are XyG arabinofuranosyltransferases. The amount of XyG arabinosylation present in *A. thaliana* lines expressing *XST1* was typically greater than that in lines expressing *XST2* (Figure 5-5, Figure 5-6). It is unclear if this is a true representation of the relative activities of these proteins or if they are translated / localized / formed into a hypothetical complex at different efficiencies. The qRT-PCR results indicate that these genes were transcribed at similar levels, though the transcript abundance varied between independent lines (Figure 5-4).

XLT2 is the *A. thaliana* protein most similar to XST1 and XST2. However, whereas XLT2 acts at the second position of the XXXG motif, both XST1 and XST2 act at the third position of the motif. As XXXG has not been reported in tomato XyG, it seems likely that both XST1 and XST2 act on XXGG to form XSGG, a structure that is commonly found in tomato XyG (Jia et al. 2003). This suggests that XST1 and XST2 recognize the XG motif, regardless of the number of X residues present. An ion with a mass to charge ratio of 1349 was detected using MALDI-TOF analysis of the XEG-released oligosaccharides in the lines expressing *XST1*, which could correspond to XXXG with two additional pentosyl-units (Figure 5-5). This suggests that XST1 may be able to act on multiple xylosyl units despite having a strong preference for the third position.

5.3.3 Origin of XST1 and XST2

To investigate the evolutionary origin of *XST1* and *XST2*, an extended phylogenetic tree was constructed using available MUR3/XLT2 subclade protein sequences from Phytozome and the MSU Medicinal Plants Database. The XLT clade of this tree is shown in Figure 5-15. The XST clade is defined by a basal branching of XST1 and XST2. It contains members of the Solanales, Gentianales, and Lamiales orders, all of which contain species that have been shown to have arabinosylated XyG (Hoffman et al. 2005; Vierhuis et al. 2001). Arabinosylated XyG oligosaccharides characterized from olive (Lamiales) and *Nerium oleander* (Gentianales) had arabinosylation at the third position of the XXXG motif consistent with the observed structure, when *XST1* and *XST2* are expressed in *A. thaliana*. The Sl02g092840 protein falls in a different clade and as expression of the corresponding gene did not result in any XyG substitution, a putative function cannot be assigned. The XLT2 clade contains one or more representatives from most of the species analyzed with the notable exception by species from the Solanales order, which are not represented. This analysis demonstrates a common origin of XyG arabinosylation for the Solanales, Gentianales and Lamiales orders and reveals that this occurred through a gene duplication and subsequent neofunctionalization of an ancestral *XLT2* gene. As Gentianales and Lamiales species both have XXXG type XyG and have not been reported to have XyG backbone acetylation, the reduction in the degree of xylosation and gain of XyG backbone acetylation present in the Solanales lineage likely occurred after the emergence of XyG arabinosylation.

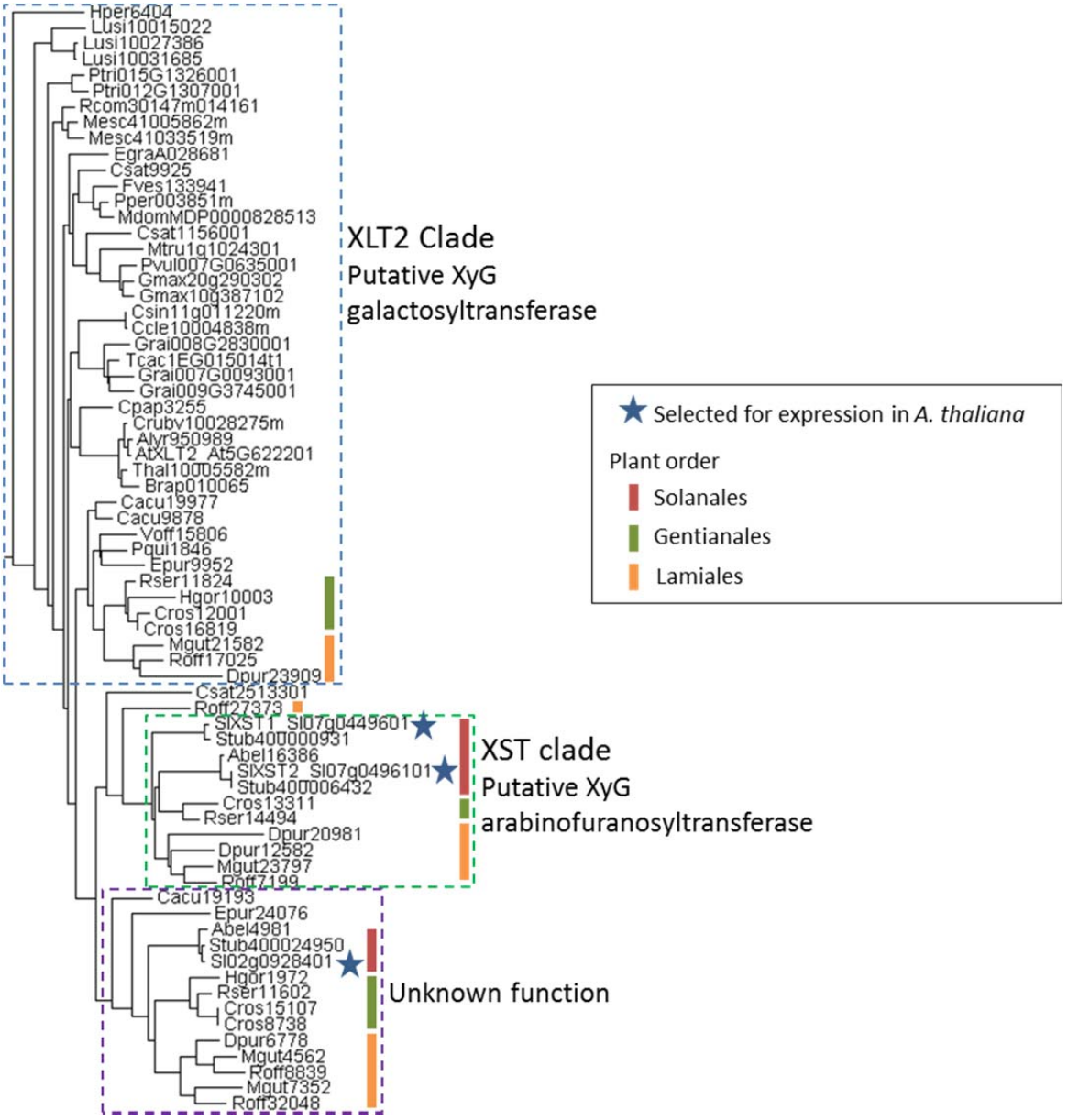


Figure 5-15. Phylogenetic analysis of the XLT2 and XLT2-related clade of GT family 47. The protein sequences are labeled with abbreviations for species (see below) and the gene model (Phytozome) or EST identifier (Medicinal Plants Genomics Resource). Abel - *Atropa belladonna*; Alyr - *Arabidopsis lyrata*; At - *Arabidopsis thaliana*; Brap - *Brassica rapa*; Cacu - *Camptotheca acuminata*; Ccle - Citrus *clementina*; Cpap - *Carica papaya*; Cros - *Catharanthus roseus*; Crub - *Capsella rubella*; Csat - *Cannabis sativa*; Csin - *Citrus sinensis*; Dpur - *Digitalis purpurea*; Dvil - *Dioscorea villosa*; Egra - *Eucalyptus grandis*; Epur - *Echinacea purpurea*; Fver - *Fragaria vesca*; Gbil - *Ginkgo biloba*; Gmax - *Glycine max*; Grai - *Gossypium raimondii*; Hgor - *Hoodia gordonii*; Hper - *Hypericum perforatum*; Lusi - *Linum usitatissimum*; Mdom - *Malus domestica*; Mesc - *Manihot esculenta*; Mtru - *Medicago truncatula*; Pper - *Prunus persica*; Pqui - *Panax quinquefolius*; Ptri - *Populus trichocarpa*; Pvul - *Phaseolus vulgaris*; Rcom - *Ricinus communis*; Roff - *Rosmarinus officinalis*; Rser - *Rauvolfia serpentina*; Sl – Solanum lycopersicum; Tcac - *Theobroma cacao*; Thal - *Thellungiella halophila*; Voff - *Valeriana officinalis*.

5.3.4 Implications for the function of XyG substitution

The *mur3.1 xlt2* double mutant plant, where the XyG is nearly devoid of substitution, has an approximately 30% decrease in plant height and reduced wall extensibility (Figure 5-12, Figure 5-13). These phenotypes are rescued by the expression of *SlMUR3*, indicating that the lack of XyG galactosylation is responsible. Expression of *XST1* and *XST2* also rescue these phenotypes, indicating that arabinofuranosyl substituents are functionally equivalent to the galactosyl residues in regards to these particular phenotypes. The fact that different side-chains can be functionally equivalent suggests a generic effect of XyG substitution that is important with respect to these phenotypes. One hypothesis is that side-chains contribute to polymer solubility and that the solubility of the polymer, not the specific side-chain composition, is important for polymer function. Enzymatic removal of fucosyl and/or galactosyl substituents has previously been shown to reduce XyG solubility (Chambat et al. 2005; Sims et al. 1998). XyG polysaccharides without appropriate solubility could aggregate in the Golgi or apoplast and hence be unable to fulfill their proper function in the extracellular matrix. This model would predict that there would be less enzyme-accessible XyG in the *mur3.1 xlt2* double mutant compared to the wild type or lines expressing *SlMUR3*, *XST1* or *XST2*. Indeed, the double mutant shows more than a 40% decrease in XEG-releasable XyG compared to wild type (Figure 5-6B). The total amount of XEG-releasable XyG is increased in the lines expressing *SlMUR3*, *XST1*, or *XST2* but not in lines expressing *Sl02g09284*. XyG solubility could also have an impact on polysaccharide biosynthesis and trafficking through Golgi bodies. Indeed, an aberrant Golgi structure was observed in mutants defective for MUR3, although the phenotype was believed to be due to an interaction between MUR3 and actin independent from its function as a XyG galactosyltransferase (Tamura et al. 2005). Aberrant Golgi structures could lead to general cellular problems including wall biosynthesis and deposition, making it difficult to draw conclusions about the mechanism by which XyG substitution affects cell wall mechanics and overall plant morphology.

5.3.5 Role of XyG diversity

If galactosyl, arabinofuranosyl, and possibly other XyG side-chains are functionally equivalent, it raises the question as to why so many types of XyG substitution patterns are present throughout the plant kingdom (Figure 1-4). One hypothesis is that these different substitution patterns do not have specific structural implications for the wall but instead create heterogeneity between species, tissues, and developmental time points. Having this heterogeneity may help to resist pathogens that lack the particularly enzymes or metabolic pathways to break down the diverse array of wall components. Examples of *A. thaliana* cell wall mutants having increased resistance to pathogens include the *pmr5* mutant, which is deficient for a putative polysaccharide acetyltransferase and has increased resistance to powdery mildew (Gille et al. 2011b; Vogel et al. 2004), and the *rwa2* mutant, which has decreased acetylation of multiple plant cell wall polysaccharides and increased resistance to *Botrytis cinerea* (Manabe et al. 2011). Notably, both of these cases seem to represent the loss or reduction of a particular cell wall modification rather than the gain of a new type of cell wall structure. To investigate how XyG substitution pattern affects susceptibility of plants to pathogens, *A. thaliana* mutants with various XyG structures were exposed to powdery mildew. All the lines tested show susceptibility to infection, indicating that for this particular pathogen

XyG structures do not convey resistance (Figure 5-14). While powdery mildew did not seem to be affected by plants with altered XyG structures, the growth of other pathogens could be impacted. Future work is required to investigate if wall polysaccharide structure can affect pathogen susceptibility and if plant pathogens could thus act as evolutionary drivers of plant wall polysaccharide structural diversification.

6 Heterologous production of XyG

6.1 Introduction

A major challenge and goal for the study of plant cell walls is to understand the mechanism and regulation of plant cell wall biosynthetic enzymes (Keegstra 2010). Traditional approaches for the identification of such genes have involved the use of protein purification and partial peptide sequencing (Crombie et al. 1998; Edwards et al. 1999; Perrin et al. 1999; Sampedro et al. 2001; Sterling et al. 2006), forward genetics (3) (Arioli et al. 1998; Hauser et al. 1995; Madson et al. 2003; Reiter et al. 1997; Turner and Somerville 1997) or reverse genetics (4) (Brown et al. 2005; Cocuron et al. 2007; Gille et al. 2011a; Persson et al. 2005; Wang et al. 2012). While these approaches have and continue to be used to identify cell wall biosynthetic genes, they have notable limitations with respect to identifying all biosynthetic genes related to a process, conclusively demonstrating biosynthetic activity, and studying the regulation and activity of particular enzymes or enzyme complexes in isolation of endogenous plant proteins that could affect the process in an unknown way. These limitations are discussed in more detail below and provide motivation for the development of a non-plant system for expression of plant cell wall biosynthetic genes for investigation into the respective activities and mechanisms of regulation of the corresponding proteins.

6.1.1 Limitations of direct protein purification, forward genetic, and reverse genetic approaches

One approach that can be used to identify plant cell wall biosynthetic enzymes (or enzymes in general) is direct protein purification from the host tissue. In this approach the enzymatic activity is solubilized and sequentially fractionated with methods such as size exclusion or ion exchange chromatography. The fraction with the enzyme of interest is identified with a suitable enzymatic activity assay, and once an enriched fraction is obtained the peptides are partially sequenced for identification. This approach has been used successfully to identify a mannan galactosyltransferase (Edwards et al. 1999), a XyG glucosidase (Crombie et al. 1998), a XyG fucosyltransferase (Perrin et al. 1999), a XyG xylosidase (Sampedro et al. 2001), and a pectin galacturonosyltransferase (Sterling et al. 2006). Notable drawbacks of this approach include the requirement for a specific and sensitive activity assay and the necessity to fully solubilize the enzyme. In some cases, activity assays can be difficult to set up due to a lack of readily available acceptor or donor substrates, such as radiolabeled nucleotide sugars. Many glycosyltransferases involved in plant cell wall biosynthesis are membrane proteins and some are part of multi-enzyme complexes (Atmodjo et al. 2011; Carpita 2011; Chou et al. 2012; Zeng et al. 2010), potentially making it more difficult to solubilize the enzyme while maintaining functionality.

Forward genetics is another viable approach that has been successfully used to identify genes involved in plant cell wall biosynthesis (Arioli et al. 1998; Gunl et al. 2011; Reiter et al. 1997; Scheible et al. 2001; Turner and Somerville 1997). In a forward genetics approach, one identifies genes that result in particular mutant phenotypes identified by screening a large population. This method is particularly well suited for the identification of genes of unknown function, such as the TBL proteins for involvement in polysaccharide acetylation (Gille et al.

2011b), or implication of previously known genes in a seemingly unrelated process, such as the XyG galactosyltransferase MUR3 being involved in cytoskeleton organization (Tamura et al. 2005). However, forward genetics is usually time consuming, limited to species with a short generation time and capable of being crossed and transformed, constrained by assay method, and could miss genes that are essential for plant growth or redundant for phenotype (3.1.1).

Rather than performing large screens on mutant populations, the availability of insertional mutants in some plant species, such as *A. thaliana* (Alonso 2003), *Brachypodium distachyon* (Thole et al. 2010), and maize (Lunde et al. 2003), allows investigating the effect of the disruption of specific candidate genes through cell wall analytical assays such as monosaccharide composition (York et al. 1986), glycosidic linkage analysis (Ciucanu and Kerek 1984), oligosaccharide mass profiling (Lerouxel et al. 2002), NMR (Cheng et al. 2013) or others. The identification of candidate genes can be aided by homology or coexpression with previously identified genes and has been used successfully to identify genes involved in the biosynthesis of XyG (4) (Cocuron et al. 2007), xylan (Brown et al. 2011; Jensen et al. 2011; Mortimer et al. 2010; Xiong et al. 2013), mannan (Gille et al. 2011a), and lignin (Alejandro et al. 2012). This approach can be successful at identifying novel genes involved in cell wall biosynthesis when knockout lines are available but is less effective in cases of gene redundancy (Jensen et al. 2011) and mutant lethality (Goubet et al. 2003). While there are technologies to utilize reverse genetic approaches in non-model species, such as tilling (Perry et al. 2003) or RNAi (Anders et al. 2012; Voinnet 2002), reverse genetics approaches are most commonly used in model species in which collections of insertional mutants are available, thereby limiting the ability of this approach to investigate processes not present in such model species. While direct protein purification approaches can directly link an enzymatic mechanism, both forward and reverse genetic approaches require additional experimental evidence to make such a connection.

6.1.2 Heterologous pathway reconstruction and *in vivo* activity assays

An alternative or complementary approach to methods described above is the heterologous expression of plant cell wall biosynthetic genes to assay the *in vivo* function of the corresponding protein by detecting the product of the reaction. This differs from *in vitro* activity assays, in which the protein is purified and supplied with exogenous substrate. Expression using a plant host has been used to identify a xylan arabinosyltransferase (Anders et al. 2012) and a XyG arabinosyltransferase (5) and to demonstrate that the CSLH/CslF/(ManS) proteins can act as beta-1,3;1,4 glucan synthases (Doblin et al. 2009). Heterologous expression and *in vivo* activity in a plant host requires that host to have the appropriate acceptor and donor substrates while lacking saturating activity of the candidate gene. Furthermore, creating transgenic plants is usually time consuming and in cases technically challenging. While some species, such as *A. thaliana* (Clough and Bent 1998), *Raphanus sativus* (Curtis and Nam 2001) and *Camelina sativa* (Lu and Kang 2008) can be transformed readily by floral dip in *A. tumefaciens*, the current transformation procedures for many plant species such as *Brachypodium distachyon* (Vogel and Hill 2008), maize, rice, wheat, and sugarcane require a tissue culture stage and take 4 to 5 months to obtain transformed plants (Shrawat and Lorz 2006). Expression in a non-plant system could allow for a faster transformation procedure and

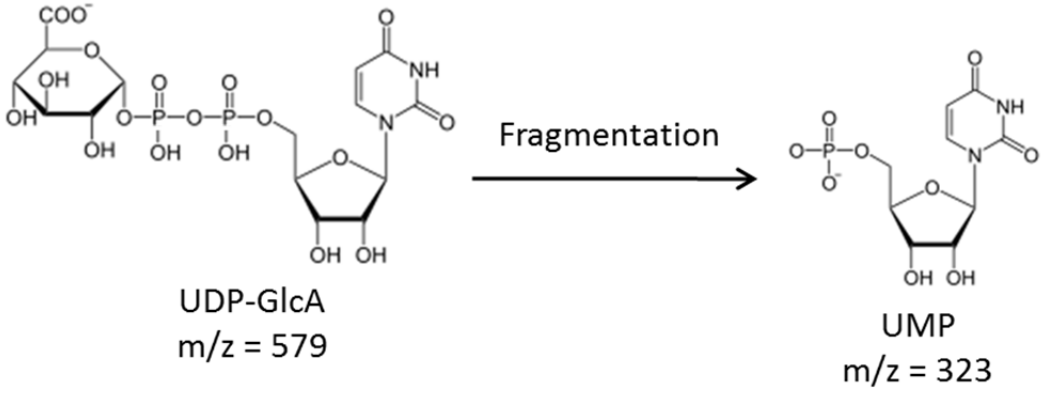
avoid the possibility of endogenous plant cell wall biosynthetic enzymes or wall components from interfering with the evaluation of the activity of a candidate gene.

Heterologous expression and *in vivo* activity has been performed previously for the wall biosynthetic gene CSLC4, a glucan synthase involved in XyG biosynthesis (Cocuron et al. 2007). This work was carried out in the yeast *P. pastoris*, which has been a successful host for heterologous expression of plant proteins for *in vitro* activity assays including the XyG galactosyltransferase MUR3 (Madson et al. 2003), select XyG xylosyltransferases (Cavalier and Keegstra 2006), XyG xylosidase (Gunl and Pauly 2011), and a XyG fucosidase (Gunl et al. 2011). Fast growth, genetic tools, and demonstrated ability to produce several different functional plant cell wall biosynthetic proteins make *P. pastoris* an attractive choice use as a platform for expression and *in vivo* activity assays of plant cell wall biosynthetic genes. Several other systems exist including *Escherichia coli*, filamentous fungi such as *Neurospora crassa*, insect cell culture (S2 cells), and mammalian cell culture. Lacking an endomembrane system and protein glycosylation, *E. coli* would probably be only of limited use. *N. crassa* is genetically tractable but likely contains hydrolytic enzymes that can degrade XyG and other cell wall polysaccharides, potentially leading to product degradation unless the corresponding genes are knocked out (Bauer et al. 2006; Tian et al. 2009). Insect cell culture has been used successfully to express CSLA for *in vitro* demonstration of mannan synthase activity (Liepman et al. 2005), although is not as fast growing or as easy to work with as *P. pastoris* (Drosophila Schneider 2 (S2) Cells, Invitrogen). Expression of a plant GT in mammalian cell culture has not been reported in the literature and is slower growing and more difficult to manipulate genetically than *P. pastoris* (Transfection of DG44 Cells & Development of Stable Cell Lines in Defined Medium, Invitrogen). For these reasons, particularly the previous successes of expressing functional plant GTs, *P. pastoris* was selected as the host for expression of plant cell wall biosynthetic enzymes. The first pathway selected for reconstruction was that of the hemicellulose XyG. The goal of this project was to demonstrate the heterologous production of a complex plant cell wall polysaccharide, in this case XyG, and to use this system as a platform for rapid identification and characterization of genes involved in XyG biosynthesis.

6.1.3 Considerations for the reconstruction of the XyG biosynthetic pathway in *P. pastoris*

P. pastoris has advantages over plant hosts due to its transformation ease, rapid growth, and lack of endogenous plant cell wall components. However it lacks some of the nucleotide sugar donor substrates required by plant cell wall GTs. *P. pastoris* likely has UDP-glucose and GDP-mannose but is likely missing UDP-glucuronic acid, UDP-xylose, UDP-galactose and UDP-arabinose as suggested by the presence or lack of the corresponding glycosyl groups in its glycoproteins and polysaccharides (Cocuron et al. 2007; Hamilton et al. 2006). Because the biosynthetic pathways for most nucleotide sugars are known in plants (Figure 1-3) (Seifert 2004), the appropriate pathway can be engineered into *P. pastoris* as required. The production of UDP-xylose from UDP-glucose requires the activity of UDP-glucose dehydrogenase (UGD) and UDP-xylose synthase (UXS) (Bar-Peled et al. 2001; Dougherty and Vanderijn 1993) (Figure 1-3). *Saccharomyces cerevisiae* has previously been engineered to produce UDP-xylose in the cytosol using these genes (Oka and Jigami 2006), indicating that a similar approach can likely be used to

produce UDP-xylose in the Golgi lumen, the location of the active site of XXT2 (Sogaard et al. 2012). The single enzyme required for conversion of UDP-glucose to UDP-galactose, UDP-galactose epimerase (UGE), is also known (Barber et al. 2006). Less well understood are the nucleotide sugar transporters required for transport of these donor substrates from the cytosol into the Golgi. Only a few of the more than 40 putative NSTs in *A. thaliana* have been associated with specific nucleotide sugars, with none having demonstrated activity for UDP-xylose, UDP-glucuronic acid, or GDP-fucose (Reyes and Orellana 2008). However, nucleotide sugar transporters from animal systems with these specificities have been identified and characterized (Ashikov et al. 2005; Kobayashi et al. 2006; Luhn et al. 2001). If one of these exogenous NSTs is indeed required for XyG production in *P. pastoris*, it would allow for easy screening to find plant NSTs with similar specificity. To do this, one could express the entire functional pathway minus the required NST of known specificity, then express candidate plant NSTs to look for ones allowing for the production of XyG. The nucleotide sugar content of *P. pastoris* can be measured by an LC-ESMS based method (Figure 6-1) (Soo et al. 2004).



Class	Nucleotide sugars	Precursor ion (m/z)	Fragment ion (m/z)
UDP-Hexose	UDP-Glc, UDP-Gal	565	323
UDP-Hexuronic Acid	UDP-GlcA, UDP-GalA	579	323
UDP-Pentose	UDP-Xyl, UDP-Ara	535	323
GDP-Hexose	GDP-Man, GDP-Glc	604	362
GDP-Deoxyhexose	GDP-Fuc	588	362

Figure 6-1. Method for analysis of nucleotide sugars. To assay for the production of nucleotide sugars, metabolites are isolated from *P. pastoris* cells and subjected to LC-ESMS analysis. The ESMS analysis allows for precise identification of ions based on initial mass and subsequent fragmentation pattern. Each class of nucleotide sugar has a characteristic precursor ion and fragment ion resulting from the fragmentation. The retention time can be used to separate isomers, which can be identified by comparison to standards.

Approximately five genes are likely required to engineer *P. pastoris* to produce a simple XyG polysaccharide consisting of a glucan backbone with regular xylosyl substituents (Figure 6-2). For an *A. thaliana*-like XyG, complete with galactosyl, fucosyl and acetyl substituents, an additional ten genes may be required including those involved in nucleotide sugar interconversion/transporter and XyG glycosylation. The approach that was selected to accomplish the expression of multiple genes was sequential genomic insertion utilizing a recyclable marker system (Figure 6-3) (Yang et al. 2009). This system utilizes a gene providing resistance to the antibiotic zeocin (ZeoR) as a positive selection for the identification of transformants and a methanol-inducible mazF counter-selectable marker. Induction of mazF, an RNase, results in the degradation of mRNA ultimately leading to cell death (Zhang et al. 2003). The ZeoR and mazF genes are flanked by direct repeats, allowing for a homologous recombination event to excise the cassette. Growth of *P. pastoris* on methanol thus allows for the selection of cells that have undergone the recombination event to lose the selectable marker. Using this method, genes were sequentially inserted into *P. pastoris* and the nucleotide sugar and polysaccharide products were assayed using LC-ESMS and glycosidic linkage analysis respectively.

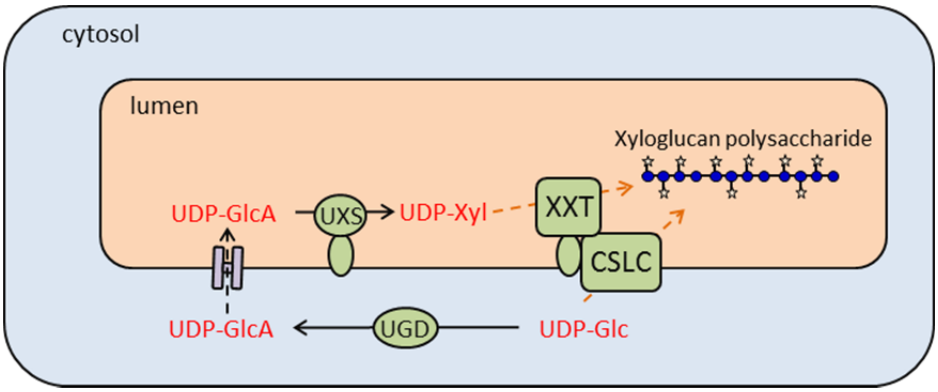


Figure 6-2. Schematic of XyG production in transgenic *P. pastoris*. To engineer *P. pastoris* to produce a simple XyG polysaccharide likely requires the addition of five genes. CSLC and XXT are the two glycosyltransferases likely needed for the synthesis of the xylosated glucan backbone. *P. pastoris* contains UDP-glucose but lacks UDP-xylose, the production of which requires the UGD and UXS enzymes. A UXS enzyme containing an N-terminal transmembrane sequence was selected, which presumably targets the active site of the UXS protein to the Golgi lumen. It is unclear if *P. pastoris* contains an endogenous nucleotide sugar transporter capable of transporting UDP-glucuronic acid.

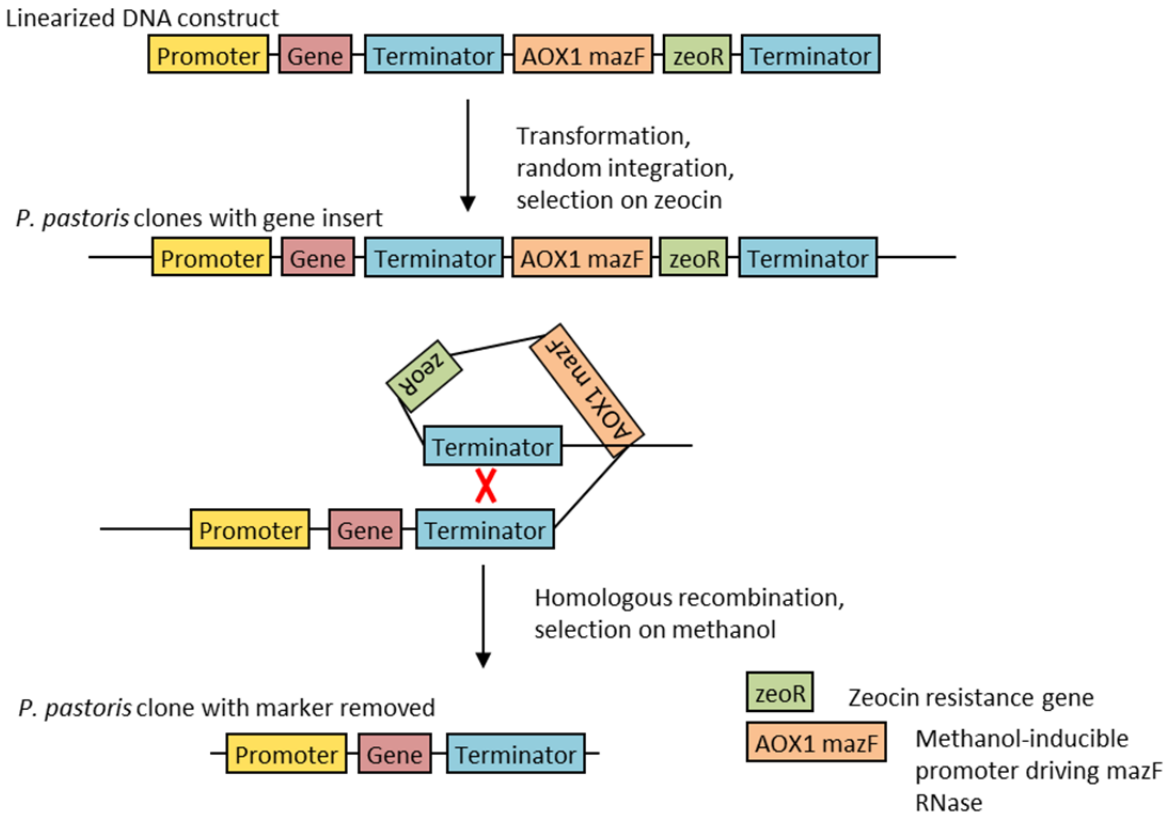


Figure 6-3. Markerless modification of *P. pastoris*. Linearized DNA constructs containing a promoter, the gene of interest, and the markerless cassette are transformed into *P. pastoris*. Transformed cells that have integrated the construct into the genome are selected by growth on media containing zeocin. After selection, expression of the mazF counter-selectable marker can be induced by growth on methanol, thereby selecting for cells which have lost the markerless cassette by homologous recombination. These cells can then be used for addition rounds of gene insertion.

6.2 Results

6.2.1 Sequential gene insertion in *P. pastoris*

Codon-optimized versions of CSLC4 and XXT2 genes (Appendix 1) from nasturtium were synthesized and transformed into *P. pastoris* (Figure 10-1, 2.2.5, 2.2.8). Genotyping PCR was performed to verify both the presence of the transgene and, after counter-selection on methanol, the excision of the selectable cassette (Figure 6-4, 2.2.2, 2.2.9). The results indicated that the system worked as expected and allowed for sequential gene insertion. However, not all transformants were able to undergo the homologous recombination event to excise the marker. Furthermore, in some transformants the new construct resulted in the loss of a previously integrated transgene (Table 6-1). These problems were overcome by screening transformants for colonies with the correct genotype, in which the marker could be

subsequently excised. Ultimately, *P. pastoris* strains incorporating up to five transgenes were obtained (Table 6-1).

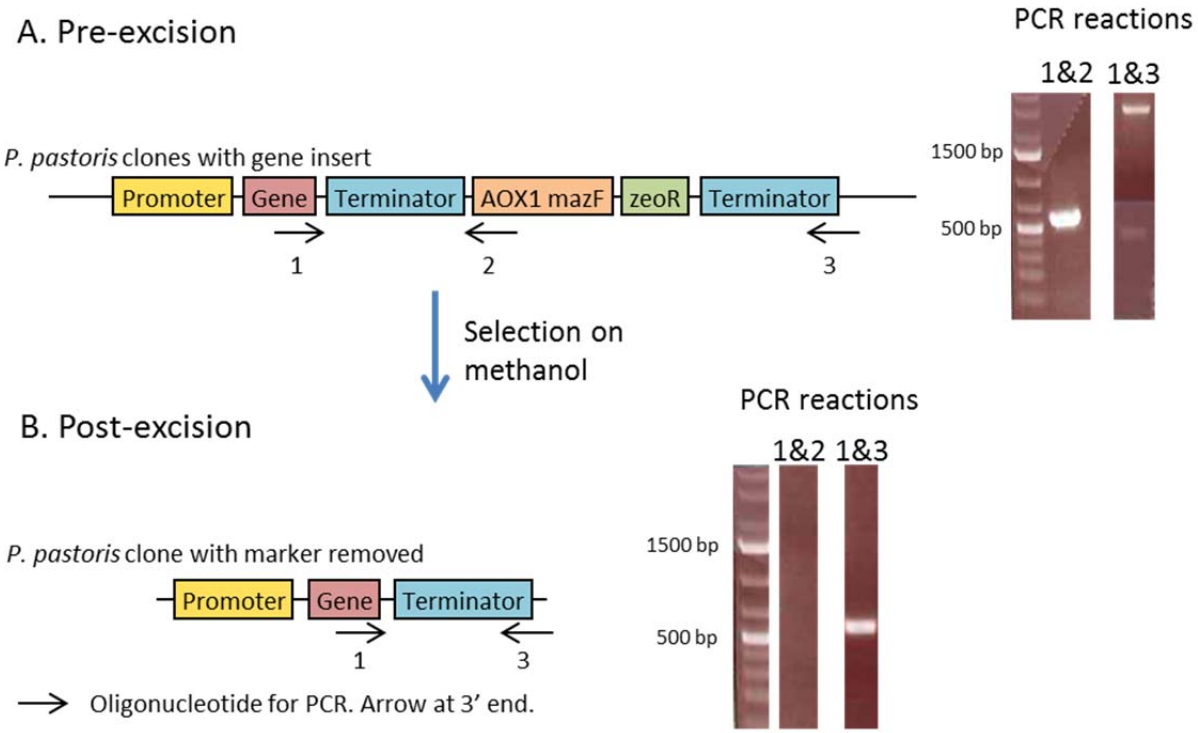


Figure 6-4. Sequential gene insertion results. The CSLC4 gene was cloned into the expression construct with the markerless cassette (Figure 10-1) and transformed into *P. pastoris*. PCR was performed using the indicated primer pairs (1 & 2 and 1 & 3) to assay for transgene presence (A). The same PCR reactions were performed again after selection on methanol (B). After excision of the markerless cassette, a PCR product is not observed for primer pair 1 & 2 and the smaller product for the primer pair 1 & 3 is the only band observed. This method was used to sequentially insert five genes into *P. pastoris*.

Existing transgenes	Colonies genotyped	Correct genotype	Gene lost	Frequency of gene loss
0	17	14	0	0%
1	16	7	7	44%
2	48	14	29	60%
3	15	4	9	60%
4	78	16	59	76%

Table 6-1. Frequency of gene loss following transformation. Colonies were genotyped following transformation for the presence of the new transgene and the existing transgenes. As the number of transgenes previously integrated into the genome increased, a greater frequency of the transformed colonies genotyped had lost one of the previously integrated genes.

6.2.2 Glycosidic linkage analysis of *P. pastoris* expressing CSLC4 and XXT2

Glycosidic linkage analysis (2.5.7) of AIR prepared from *P. pastoris* cells was performed to determine whether the transgenic *P. pastoris* expressing CSLC4 and XXT2 was able to produce XyG or the glucan backbone. A greater than 3-fold increase in the relative abundance of 4-linked glucose was detected in *P. pastoris* expressing both CSLC4 and XXT2 compared to wild type (Figure 6-5). This increase was not observed in *P. pastoris* expressing either of these glycosyltransferases individually.

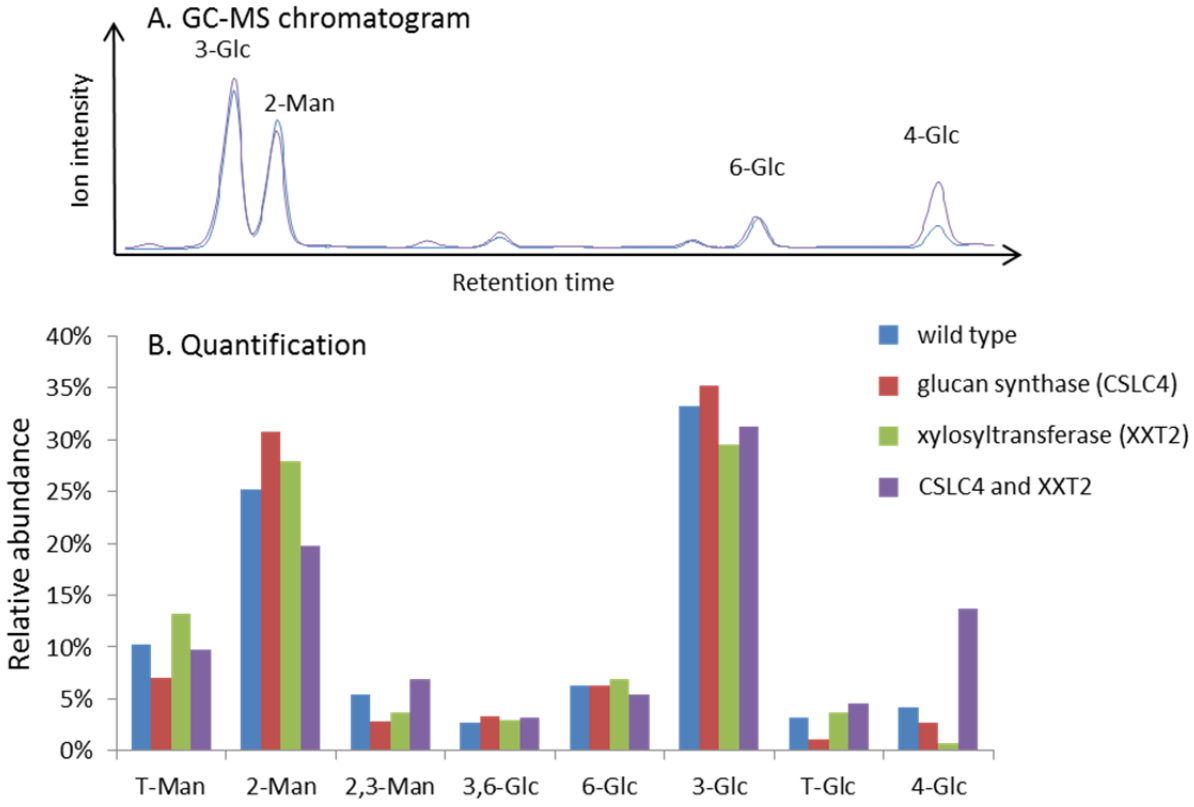


Figure 6-5. Linkage analysis of transgenic *P. pastoris*. Glycosidic linkage analysis was performed on AIR material from wild type *P. pastoris* and transgenic cells expressing CSLC4, XXT2, or both CSLC4 and XXT2. The partial chromatogram from the GC-MS (total ion count) is shown in A whereas the quantified results are in B. An increase in 4-linked glucose was observed only in *P. pastoris* expressing both CSLC4 and XXT2.

6.2.3 Nucleotide sugar analysis of *P. pastoris* expressing UGD and UXS

The XXT2 protein, purified from transgenic *P. pastoris*, has been shown be sufficient to catalyze the transfer of xylose from UDP-xylose to cellohexaose acceptors (Faik et al. 2002). Therefore, the lack of xylose on the glucan produced by *P. pastoris* expressing CSLC4 and XXT2 is presumably caused the inability of *P. pastoris* to produce UDP-xylose. To produce UDP-xylose,

the UDP-glucose Dehydrogenase (UGD) and UDP-xylose Synthase (UXS) genes (Appendix 1) identified in the nasturtium transcriptome (4) were codon optimized and expressed in *P. pastoris* (Figure 10-1, 2.2.5, 2.2.8). Wild type *P. pastoris* was found to contain detectable amounts a UDP-Hexose when a methanol extract is analyzed by LC-ESMS (2.5.9), believed to be UDP-glucose, but no UDP-xylose or UDP-glucuronic acid. *P. pastoris* expressing the UGD gene accumulated large amounts of UDP-glucuronic acid whereas *P. pastoris* expressing both UGD and UXS accumulated both UDP-glucuronic acid and UDP-xylose (Figure 6-6). The identity of the UDP-glucuronic acid and UDP-xylose was confirmed by fragmentation pattern and retention time matching that of standards (Figure 6-1).

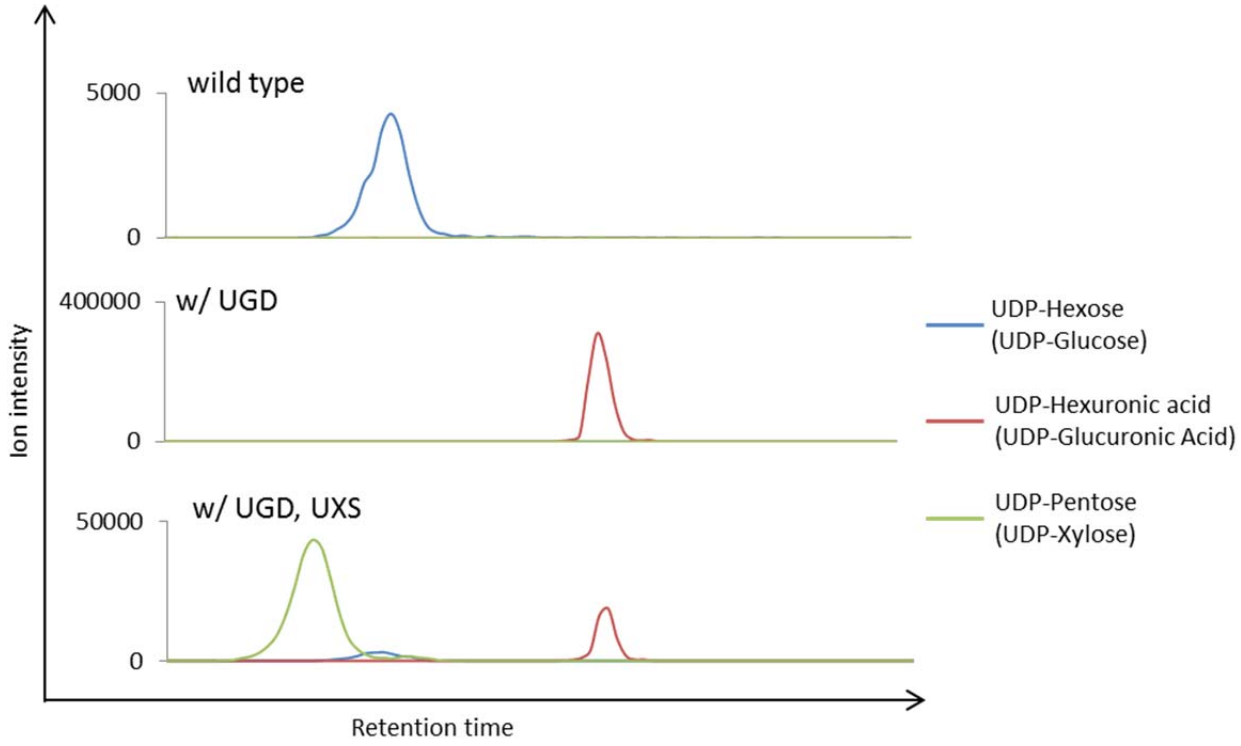


Figure 6-6. Nucleotide sugar analysis of transgenic *P. pastoris*. Wild type *P. pastoris* and *P. pastoris* cells expressing UGD or UGD and UXS were analyzed for nucleotide sugar content using an LC-ESMS based method. UDP-glucose was detected in the wild type cells. UDP-glucuronic acid and UDP-xylose were detected in *P. pastoris* expressing UGD or UGD and UXS respectively.

6.2.4 Expression of CSLC4, XXT2, UGD and UXS in *P. pastoris*

As the previous results demonstrated that CSLC4 and XXT2 expression are sufficient for backbone formation (Figure 6-5) and the expression of UGD and UXS for UDP-xylose production (Figure 6-6), all four genes were coexpressed into *P. pastoris* in an attempt to produce XyG (2.2.11, 2.2.9). This transgenic *P. pastoris* produced a 1,4-glucan and UDP-xylose, but no terminal xylose or 4,6- glucose was detected as would be expected if a xylosylated XyG were produced (Figure 6-7, Figure 6-8, 2.5.7, 2.5.9).

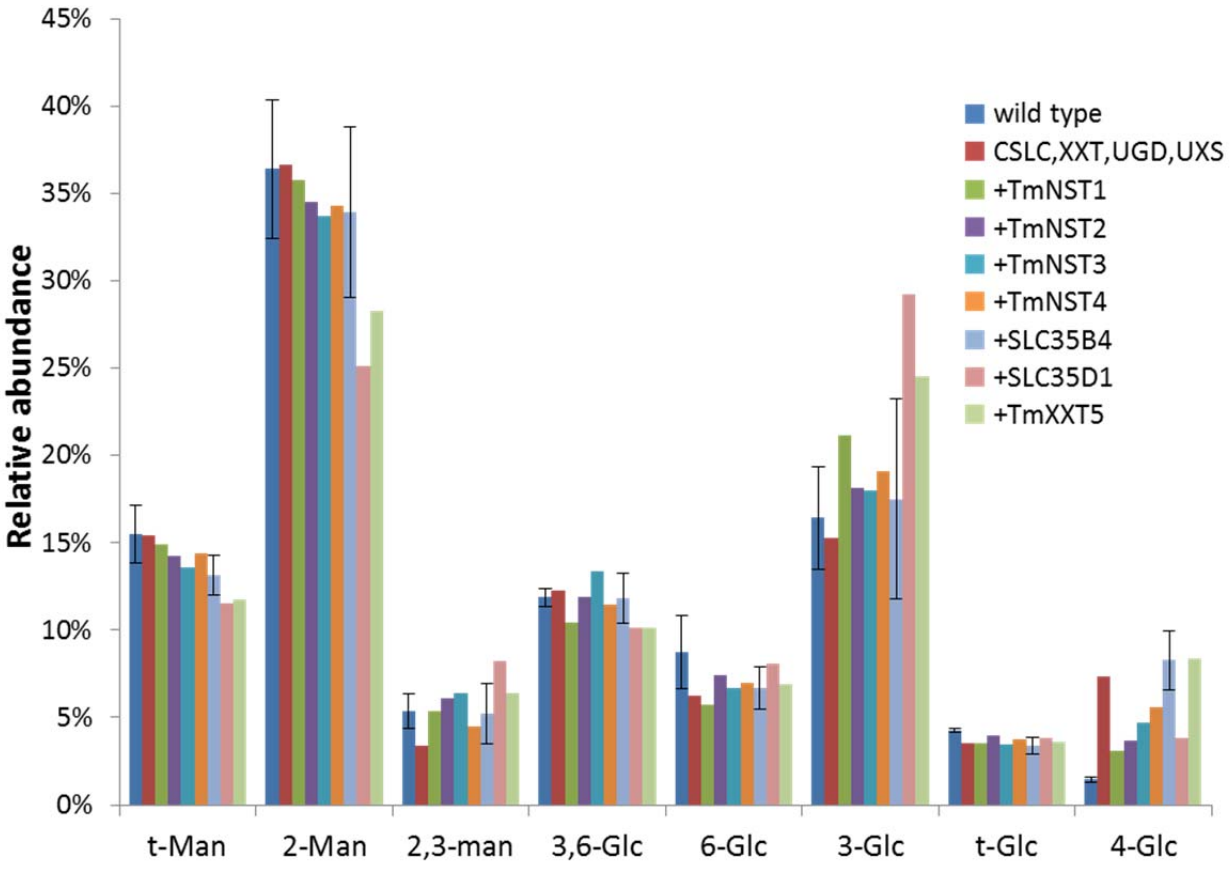


Figure 6-7. Glycosidic linkage analysis of transgenic *P. pastoris*. The + indicates that the line is expressing CSLC4, XXT, UGD and UXS in addition to the transgene listed. SLC35B4 and SLC35D1 are mammalian nucleotide sugar transporters that can act on UDP-xylose and UDP-glucuronic acid respectively. Terminal-xylose or 4,6-glucose were not detected in any of the lines. Error bars indicate standard deviation of biological replicates.

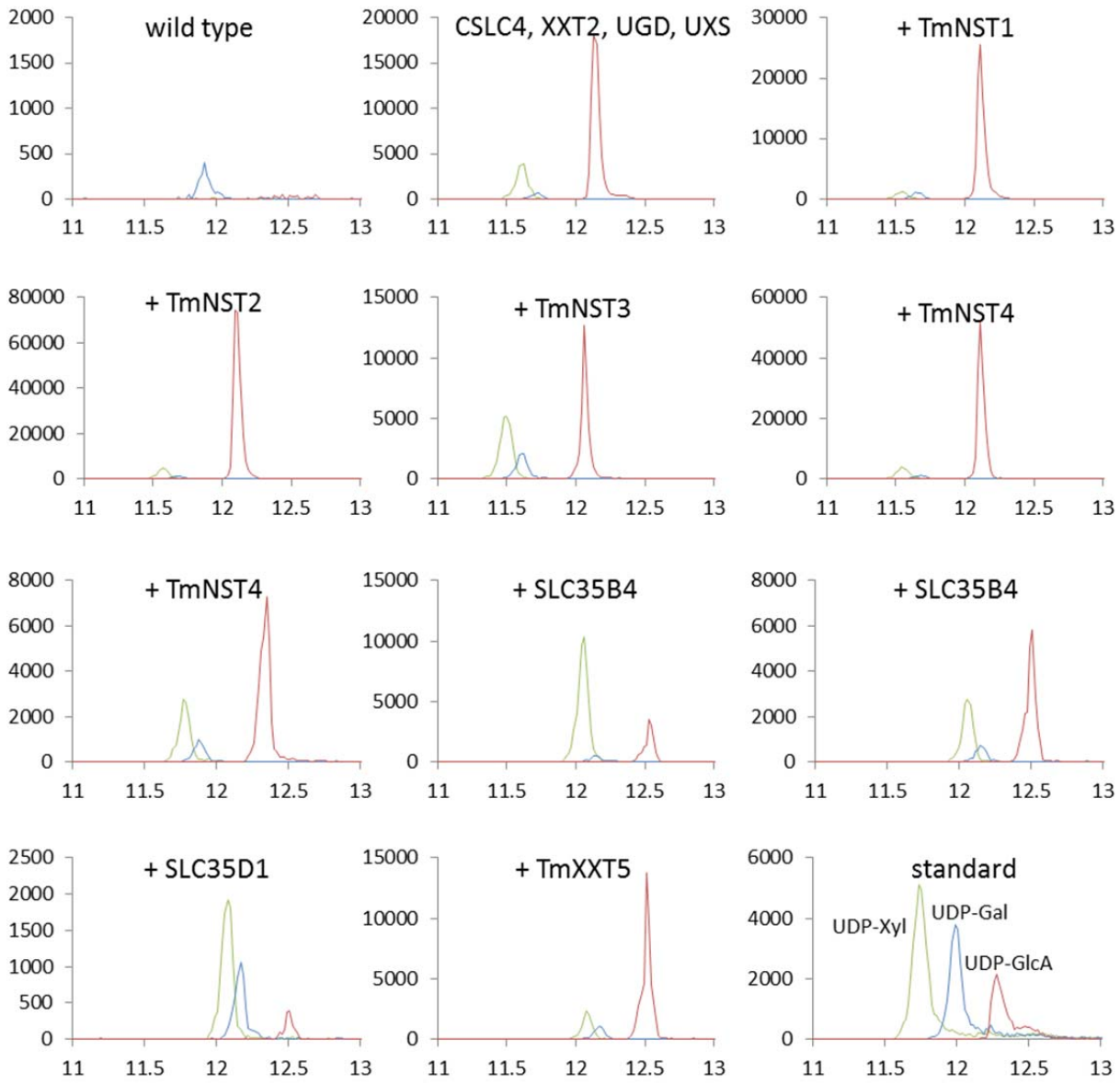


Figure 6-8. Nucleotide sugar analysis of transgenic *P. pastoris*. LC-ESMS was used to assay for the presence of nucleotide sugars. + indicates the cells contain CSLC4, XXT2, UGD and UXS in addition to the indicated gene. The standard shown is 5 µg / mL each of UDP-galactose (UDP-Gal), UDP-glucuronic acid (UDP-GlcA), and UDP-xylose (UDP-Xyl). Note the difference in relative retention time for the UDP-Gal standard and the UDP-hexose detected in *P. pastoris*, believed to be UDP-glucose. Three ion channels are shown: blue – 565->323 (UDP-hexose), red – 579->323 (UDP-hexuronic acid), green – 535->323 (UDP-pentose). X-axis is retention time in minutes, Y-axis is relative ion intensity. UDP-xylose was detected in all transgenic lines with the relative amount of UDP-GlcA notably decreased in the line expressing the mammalian UDP-GlcA transporter SLC35D1.

6.2.5 Addition of nucleotide sugar transporters

The UXS selected for expression in this work has a predicted transmembrane (Krogh et al. 2001) and therefore was hypothesized to result in the production of UDP-xylose in the Golgi lumen. However, the location of the produced UDP-xylose was not confirmed here. The lack of XyG production in *P. pastoris* despite the presence of a glucan acceptor, the XXT, and UDP-xylose suggests that the UDP-xylose might be produced outside the Golgi, possibly in the cytosol, and is not transferred into the Golgi, the site of the xylosyltranferase. In an attempt to mitigate this problem, several nucleotide sugar transporter genes were cloned and expressed in a *P. pastor*is strain already expressing the CSLC4, XXT2, UGD and UXS (2.2.5, 2.2.8). The nucleotide sugar transporters utilized included a UDP-xylose transporter (SLC35B4) (Ashikov et al. 2005), a UDP-Glucuronic acid transporter (SLC35D1) (Kobayashi et al. 2006), and four putative nucleotide transporters with unknown specificity identified from the nasturtium transcriptional analysis (4, Appendix 1, Figure 10-2). The simultaneous expression of these genes with CSLC4, XXT2, UGD and UXS in *P. pastoris* did not result in the production of XyG that was detectable by glycosidic linkage analysis as 4,6-linked glucose or terminal xylose were absent (Figure 6-7). The relative amount of 4-linked glucose detected for this line appears lower than that for the strain expressing only CSLC4 and XXT2 (Figure 6-5, Figure 6-7). However the relative abundance of other linkages, notably 3,6-linked glucose and 2-linked mannose, also changes indicating that such alterations could be due to differences in the specific times of induction and growth, which were similar between experiments but not precisely controlled.

6.2.6 Addition of TmXXT5

An alternative hypothesis is that an additional XXT is required for XyG xylosylation. Evidence for this hypothesis stems from the observation that two XXT genes, XXT2 and XXT5, are expressed in developing nasturtium seeds (4). XXT5 was cloned from nasturtium and expressed in a *P. pastoris* already containing CSLC4, XXT2, UGD and UXS. The addition of XXT5 did not result in the production of XyG detectable by glycosidic linkage analysis (Figure 6-7).

6.2.7 Subcellular localization of XXT2 and CSLC4 in *P. pastoris*

The mechanism by which XyG biosynthetic proteins are retained in the Golgi is not well understood, but is presumably due to a yet-unidentified retention signal within the peptide sequence (Gao et al. 2012; Tu and Banfield 2010). It was unclear if the mechanism for retaining these proteins in the Golgi in plants would also function in *P. pastoris*. A mislocalization of these proteins to the plasma or vacuolar membrane would result in the CSLC4 protein still being able to utilize a cytoplasmic pool of UDP-glucose to produce the glucan backbone (as the orientation of the protein would not change) but no UDP-xylose would be available for the XXT. To investigate this possibility, the CSLC4 and XXT2 genes were tagged at the C-terminus with GFP and expressed in *P. pastoris* (2.2.5, 2.2.8). Both CSLC4-GFP and XXT2-GFP localized to intracellular compartments consistent with previously described Golgi structures of *P. pastoris* (Figure 6-9)(Rossanese et al. 1999). To ensure this localization was reflective of the non-tagged proteins and not an artifact of the GFP fusion, the functionality of the constructs was assayed using glycosidic linkage analysis to look for the production of 4-linked glucan (2.5.7). Both the GFP-tagged CSLC4 and XXT2 remained functional as *P. pastoris* expressing either CSLC4-GFP and XXT or XXT-GFP and CSLC4 resulted in the production of a 4-linked glucan (Figure 6-10).

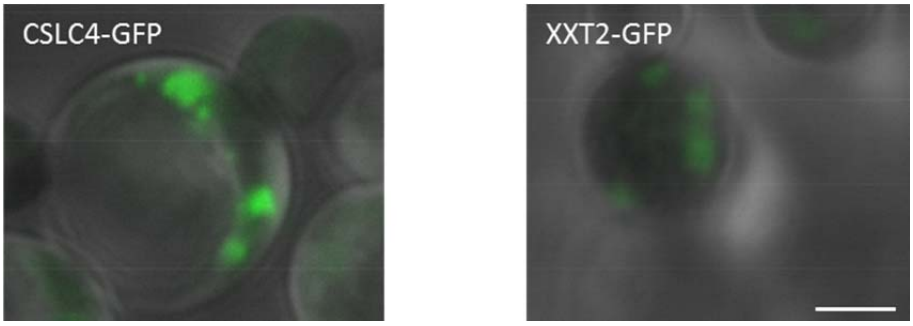


Figure 6-9. Localization of CSLC4 and XXT2 within *P. pastoris*. N-terminally GFP-tagged CSLC4 and XXT2 were expressed in *P. pastoris* to examine localization of the proteins. In both cases the fluorescent signal was observed in small intracellular vesicles, which is consistent with localization to the Golgi. Fluorescence was not observed in the plasma or vacuolar membrane. Scale bar = 2 µm.

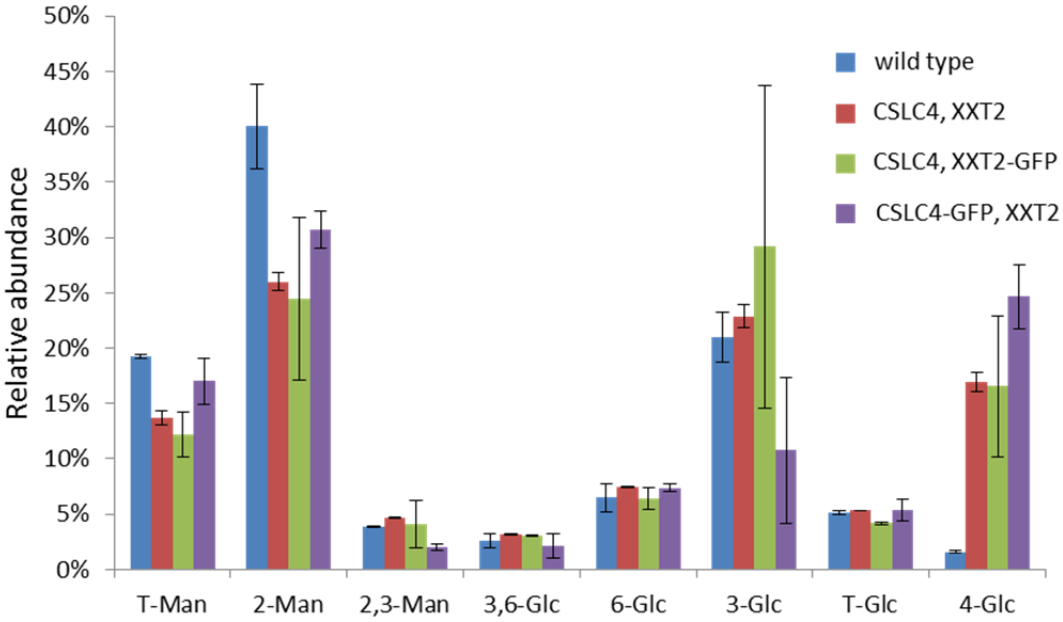


Figure 6-10. Functionality of GFP tagged CSLC4 and XXT2. Glycosidic linkage analysis was performed on *P. pastoris* expressing native CSLC4 and XXT2 or a GFP-tagged version of the peptide along with the native partner. The abundance of 4-Glc detected indicates the GFP tagged proteins are functional.

6.3 Discussion

6.3.1 Efficacy of the mazF markerless cassette for sequential gene insertion

The genotyping results indicated that the markerless gene insertion system was successful in sequentially integrating five genes into the *P. pastoris* genome (Figure 6-4, Table 6-1). However, the current system has several drawbacks. The use of methanol to induce expression of both the heterologous gene and the counterselectable marker means that the marker must be excised before checking for the functionality of the transgene (Figure 6-3). One way around this problem is to use a non-excisable zeocin resistance cassette rather than the entire markerless cassette. This would allow for the transformed cells to be induced and phenotyped immediately after selection and genotype confirmation, which facilitates screening candidate genes. This method was used to assay the effect of the nucleotide sugar transporters and the XXT5 (Figure 6-7, Figure 6-8). The disadvantage of this approach is that the marker cannot be excised, so if an additional gene is required one must repeat the transformation and genotyping to generate a strain with the markerless cassette. Having a counterselectable marker that is only active (transcriptionally or mechanistically) under certain conditions would eliminate this problem. Examples of mechanistically inducible counter-selectable markers include orotidine 5-phosphate decarboxylase (URA3), ribosomal protein S12 (rpsL), and thymidine kinase (Kit 1962, Anthony 1995, Converse 2004)(Bayliss and Ingraham 1974; Boeke et al. 1987; Lederberg 1951; Reyrat et al. 1998). Alternatively, an inducible promoter could be used to drive the expression of constitutively-active counter-selectable markers, such as MazF, ccdB, or growth-inhibitory (GIN) sequences, as long as the promoter is not the same as that used to drive transgene expression (Akada et al. 1997; Bernard and Couturier 1992; Zhang et al. 2003). In addition to the methanol-inducible promoter used in this work, a copper-inducible and several glucose-inducible promoters are known that can be utilized in *P. pastoris* (Koller et al. 2000; Prielhofer et al. 2013). Alternatively, exogenous transcriptional induction systems such as the tetracycline-inducible system could be engineered into *P. pastoris* (Gari et al. 1997; Gossen and Bujard 1992). By utilizing one or several of these or similar available tools / approaches, it might be possible to develop a system suited for gene expression without the need to first undergo marker excision.

A second issue with the sequential insertion system is that *P. pastoris*, rather than always randomly integrating exogenous linear pieces of DNA into the genome, sometimes recombines them with homologous genomic DNA. This is a problem when the same promoter and terminator are used for each gene, as the addition of a gene has the potential to knock out one previously inserted (Table 6-1). With additional transgenes in the genome, the chances of knocking out a previous gene increase in addition to the number of PCR reactions required to screen each colony for presence of all transgenes. This was found to be a manageable but significant issue when doing the sequential insertion of five genes into *P. pastoris*, with the observed frequency of homologous recombination (HR) and excision of a previously integrated gene being 76% (Table 6-1). As random integration is believed to occur through the non-homologous end joining (NHEJ) DNA repair pathway, it may be possible to increase the frequency of random integration by increasing the rate of NHEJ or by decreasing / preventing homologous recombination genetically (Naatsaari et al. 2012; Ninomiya et al. 2004; Poggeler

and Kuck 2006) or chemically (Noguchi et al. 2006). It might be possible to increase NHEJ by the introduction of double strand breaks using UV light. HR could be reduced by knocking out or conditionally repressing some of the known components of the HR pathway, such as Rad51 or Rad54 (Heyer et al. 2010). Knocking out homologous recombination machinery would prevent the direct repeats flanking the markerless cassette from recombining, but an alternate system utilizing a recombinase (FLP or Cre/Lox, both of which have been used successfully in *P. pastoris*) could be used instead (Cregg and Madden 1989; Marx et al. 2008).

An alternative to a sequential gene insertion system would be to place all or a subset of the desired genes into a single construct for either genomic integration or expression from a plasmid. A plasmid-based system would allow for faster expression of a core set of genes (such as CSLC4, XXT, UGD and UXS). With an appropriate origin of DNA replication, plasmids can be propagated in *P. pastoris* (Cregg et al. 1985). The disadvantage of using a plasmid system is that the plasmid could be lost in the absence of selective pressure, preventing repeated use of the same selectable marker. Additionally, if the same promoter and terminator sequences are used for each transgene, the stability of these sequences may be reduced if there are multiple copies of the plasmid present due to homologous recombination. Despite these concerns, a fully or partially plasmid-based system could be faster and provide more consistent results than a system reliant on genomic integration and therefore remains a viable option for future efforts to engineer *P. pastoris*.

6.3.2 Production of a 4-linked glucan in *P. pastoris*

The expression of TmCSLC4 and TmXXT2 in *P. pastoris* resulted in the production of a 4-linked glucan as revealed by glycosidic linkage analysis (Figure 6-5). This confirms the results of a previous study which showed production of a glucan chain using the CSLC4 and XXT2 genes from *A. thaliana* (Cocuron et al. 2007). That the expression of both CSLC4 and XXT2 is required for the production of a glucan chain suggests that these proteins interact, with XXT2 having a non-catalytic role in the production of the glucan chain. An interaction between the *A. thaliana* CSLC and XXT was investigated in a separate study using bifluorescence complementation and pull down assays which indicated CSLC4 likely interacts with XXT5 and may have a weaker interaction with XXT2 (Chou et al. 2012). The requirement to have XXT present for CSLC4 to be active may be a way to prevent the CSLC4 from producing a non-xylosyated glucan backbone, which could aggregate within the endomembrane system *in planta* (Whistler 1973). Like CSLC and XXT for XyG biosynthesis, multiple GTs involved in xylan biosynthesis form a complex (Zeng et al. 2010). CSLA alone is sufficient for mannan backbone formation and does not require the galactosyl transferase (also GT34 like the XXT enzymes) for activity (Liepman et al. 2005). The topology of CSLA suggests the active site of this enzyme is in the Golgi lumen (Davis et al. 2010). In contrast, protein topology experiments suggest that the active site of CSLC resides in the cytosol, thus taking cytosolic UDP-glucose and pushing the glucan chain through the membrane into the Golgi (Davis et al. 2010). These differences suggest that despite a common evolutionary origin, the biosynthetic processes for mannan and XyG biosynthesis are functionally quite different.

6.3.3 Production of UDP-xylose in *P. pastoris*

Expression of UGD and UXS in *P. pastoris* was sufficient for the formation of UDP-GlcA and UDP-xylose *in vivo* (Figure 6-6). Unlike a previous reconstruction of the UDP-xylose pathway, which used a cytosolic form of UXS (Oka and Jigami 2006), the UXS selected for expression in this work has an N-terminal transmembrane domain that is predicted to localize the active site of the enzyme to the lumen of the endomembrane system. This is in contrast to the UGD enzyme, which lacks a putative transmembrane domain and is predicted to localize to the cytoplasm. That the expression of these two enzymes was sufficient for the production of UDP-xylose suggests that either there is an endogenous UDP-GlcA transporter able to transport UDP-GlcA into the ER/Golgi or that some UXS activity is present in the cytosol. Nucleotide sugar transporters can have a broad range of specificity (Berninsone et al. 2001; Hong et al. 2000), so it is possible that an endogenous transporter could transport UDP-GlcA. The subcellular distribution of the produced UDP-xylose in transgenic *P. pastoris* is therefore not clear but is important as the XXT proteins likely require a luminal pool of UDP-Xylose for activity.

6.3.4 Lack of detectable XyG after coexpression of CSLC4, XXT2, UGD and UXS

Simultaneous expression of CSLC4, XXT, UGD and UXS in *P. pastoris* failed to result in production of xylosylated glucan polysaccharide that was detectable by glycosidic linkage analysis. There are multiple possible explanations, several of which were explored with additional experiments. One hypothesis was that the CSLC4 and XXT2 proteins could be mislocalized, which could result in a lack of UDP-xylose being available for XXT2. GFP-fusion proteins of both CSLC4 and XXT2 localized to structures consistent with Golgi (Figure 6-9) (Rossanese et al. 1999). A second hypothesis was that an additional protein is required for activity. The *A. thaliana* CSLC4 protein had a stronger affinity for XXT5 in interaction experiments that XXT1 or XXT2 (Chou 2012). A nasturtium ortholog of XXT5 was found to be expressed, though not induced, during the deposition of storage XyG during seed development suggesting that both XXT5 and XXT2 may be required for XyG xylosylation. Expression of a non-codon optimized version of XXT5 from nasturtium along with CSLC4, XXT2, UGD and UXS in *P. pastoris* failed to result in the production of detectable XyG (Figure 6-7). Therefore, there must be another limiting factor preventing production of XyG in *P. pastoris*.

A third scenario is that the CSLC4/XXT2 complex is complete and correctly localized but UDP-xylose is not present in the Golgi. Significant amounts of UDP-xylose were produced in the cells (Figure 6-6, Figure 6-8), but a lack of UDP-xylose in the Golgi lumen would account for XXT2 being unable to xylosylate the glucan backbone. That UXS was able to produce UDP-xylose without an exogenous UDP- glucuronic acid transporter suggesting that perhaps UXS is making UDP-xylose in the cytosol. To address this, a total of six nucleotide sugar transporters were cloned and expressed in *P. pastoris* along with the CSLC4, XXT2, UGD and UXS. Four of these NSTs were highly expressed in the nasturtium dataset and the other two were obtained from mammalian systems, one with specificity for UDP-xylose and UDP-N-acetylglucosamine and a second with the ability to transport both UDP-glucuronic acid and UDP-acetylgalactosamine (Ashikov et al. 2005; Bakker et al. 2008; Muraoka et al. 2001). The addition of these transporters failed to result in the production of XyG detectable by glycosidic linkage analysis (Figure 6-7), although each line was able to produce the 4-linked glucan backbone and UDP-

xylose (Figure 6-8). Interestingly, the amount of 4-linked glucan decreased in several of the lines. The effect was particularly strong for the mammalian transporter SLC35D1 and intermediate for the nasturtium transporters with TmNST1 being the strongest. No effect was observed for the UDP-xylose transporter SLC35B4. A decrease in the relative amount of UDP-glucuronic acid and UDP-xylose was also observed in the strain expressing SLC35D1 (Figure 6-8). It is unclear if this may be due to possible reduced expression of CSLC4/XXT2 (which could be affected by the addition or expression of the transgene), a decrease in the availability of UDP-glucose due to the action of the transporter, or the production of a small amount of soluble, xylosylated glucan that was secreted to the media. The complex media used to grow *P. pastoris* is not well suited for examining the soluble fraction as it contains large amounts of yeast extract and peptone (30 mg / mL total) which may make it difficult to detect small amounts xylosated glucan due to the complex nature of these materials and the fact that peptone contains animal proteins, which can be xylosated (Moremen et al. 2012). Moving to a minimal media system would facilitate analysis of the soluble fraction, which may harbor small amounts XyG.

6.3.5 Conclusion

If ultimately successful, having a heterologously reconstructed XyG biosynthetic pathway would allow for rapid functionality testing of candidate XyG biosynthetic genes. In contrast to a plant system (5), working with a yeast system, which does not normally produce XyG, would allow for study of XyG biosynthetic genes in isolation from endogenous plant proteins, which may modulate their activity. At this stage, determining the limiting factors for *in vivo* XyG production in *P. pastoris* remains a challenge but if solved could provide insight into how the process occurs *in planta*. This system could also be used to reconstruct the pathways for other cell wall polysaccharides. While altering XyG structure *in planta* has been facilitated by the relatively mild morphological phenotypes (Cavalier et al. 2008; Madson et al. 2003; Vanzin et al. 2002), the modification of other wall components can be quite severe or lethal (Geshi et al. 2013; Goubet et al. 2003; Sitaraman et al. 2008). Hence, the identification and study of these genes, made difficult by such phenotypes, could be facilitated by a heterologous production system such as that attempted in this work.

7 Concluding remarks

Several approaches were utilized in this work to identify previously uncharacterized genes involved in the biosynthesis of the plant hemicellulose XyG. The identification of these genes provided insight not only into the how this polysaccharide is produced but also into general pathways of polysaccharide biosynthesis (e.g. polysaccharide acetylation) and the genetic basis for and evolution of polysaccharide structural diversity.

Investigation into an *A. thaliana* ems mutant having altered XyG resulted in the identification of the AXY9 gene, a defect which results in decreased XyG acetylation (Figure 3-18). As AXY9 also affects the acetylation of other cell wall polysaccharides (Figure 3-23), it appears that AXY9 is part of a pathway for delivering activated acetate to the Golgi lumen for use by polysaccharide-specific acetyltransferases (3.3.4.3). As part of this pathway, AXY9 may use acetyl-CoA as a donor substrate and transfer the acetyl group to an intermediate carrier molecule, which is subsequently transported into the Golgi for use by the TBL proteins. The severe growth phenotype of the *axy9.2* mutant (Figure 3-20), attributed to decreased xylan acetylation (3.3.3), demonstrates that this particular wall polymer substituent is critical for the function of this polysaccharide (3.3.3.2). Future work is required to elucidate the specific biochemical activity of AXY9 and to better understand the mechanism by which decreased xylan acetylation apparently leads to a collapsed xylem and severe growth phenotype of the *axy9.2* mutant.

Transcriptional profiling (4) and comparative genomics (5) approaches resulted in the identification of a XyG galactosyltransferase (XLT2) and two XyG arabinofuranosyltransferases (XST1 and XST2), respectively, by homology to a previously identified XyG galactosyltransferase (Madson et al. 2003). The discovery of these genes allowed for genetic experiments to investigate the *in vivo* function of specific glycosyl side-chains. It was found that arabinofuranosyl side-chains can act in a functionally equivalent manner to galactosyl residues in some respects (Figure 5-11, Figure 5-12), suggesting that XyG substitution may have a somewhat generic function, such as regulating the solubility of the polysaccharide, as opposed to exclusively mediating specific interactions with receptors or molecules (5.3.4). The identification of these genes validates the hypothesis that this clade of GT47 has GTs responsible for XyG structural diversity and provides insight into the evolution of GTs with novel donor and acceptor substrate specificity (5.3.3). The comparative genomics and heterologous expression strategy employed for the identification of candidate XyG arabinofuranosyltransferases proved effective and can likely be applied in efforts to identify the genetic basis for structural variation in a wide variety of polysaccharides and glycoproteins. Such an approach is increasingly powerful and applicable with progressively more genomic and transcriptomic data available (Paterson et al. 2010). This work also demonstrates the potential for significantly reengineering plant cell wall polysaccharides. While genetic reduction of XyG galactosylation resulted in decreased plant growth (Figure 4-8), the addition of arabinofuranose resulted in a rescue of plant growth (Figure 5-11). This suggests that efforts to alter cellulosic biomass for particular purposes may be able to mitigate detrimental effects to the plant cell wall by engineering it to resemble the walls of other plant species with desirable properties. As a great diversity of cell wall structures exist throughout the plant kingdom, the space for designing functional cell walls is quite large (Pauly and Keegstra 2010).

The heterologous expression of XyG GTs was effective for the identification of the XyG arabinofuranosyltransferases, however in a plant system one cannot exclude the involvement of other proteins which may modulate enzymatic activity and can be limited by acceptor substrate availability. To avoid these problems and develop a truly heterologous system, an effort was made to reconstruct the XyG biosynthetic pathway in the yeast *P. pastoris* (6). Such a system would enable for rapid screening of XyG biosynthetic genes and would provide a platform to explore the regulation of XyG biosynthesis in isolation of endogenous plant factors. While ultimately not successful in producing a XyG polysaccharide (6.3.4), the progress made here will hopefully serve as stepping stone for future efforts to reconstruct plant cell wall biosynthetic pathways in heterologous systems.

Overall, this work increases our understanding of XyG and general plant cell wall polysaccharide biosynthesis and function. Future work is needed to fully understand the role of AXY9 in the polysaccharide acetylation pathway and why this type of modification is apparently so critical for xylan function. Structural studies of the GT47s presented here and others can likely provide insights into the specific amino acids of these proteins responsible for acceptor and donor specificity, likely facilitating GT specificity prediction and potentially allowing for the production of GTs with specificities unique from those observed in nature. The approaches presented in this work will likely be successful for the discovery of the genetic basis for additional plant cell wall polysaccharide diversity, thereby allowing for better understanding not only of these processes but also of the functional significance of this diversity.

8 References

Aist JR (1976) Papillae and Related Wound Plugs of Plant-Cells. Annu Rev Phytopathol 14: 145-163

Akada R, Yamamoto J, Yamashita I (1997) Screening and identification of yeast sequences that cause growth inhibition when overexpressed. Molecular & General Genetics 254: 267-274

Akoh CC, Lee GC, Liaw YC, Huang TH, Shaw JF (2004) GDSL family of serine esterases/lipases. Prog Lipid Res 43: 534-552

Alejandro S, Lee Y, Tohge T, Sudre D, Osorio S, Park J, Bovet L, Lee Y, Geldner N, Fernie AR (2012) AtABCG29 is a monolignol transporter involved in lignin biosynthesis. Curr Biol 22: 1207-1212

Alonso JM (2003) Genome-wide insertional mutagenesis of Arabidopsis thaliana (vol 301, pg 653, 2003). Science 301: 1849-1849

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen HM, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide Insertional mutagenesis of Arabidopsis thaliana. Science 301: 653-657

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402

Amor Y, Haigler CH, Johnson S, Wainscott M, Delmer DP (1995) A Membrane-Associated Form of Sucrose Synthase and Its Potential Role in Synthesis of Cellulose and Callose in Plants. Proc Natl Acad Sci U S A 92: 9353-9357

Amrhein N, Frank G, Lemm G, Luhmann HB (1983) Inhibition of lignin formation by L-alpha-aminooxy-beta-phenylpropionic acid, an inhibitor of phenylalanine ammonia-lyase. Eur J Cell Biol 29: 139-144

Anders N, Wilkinson MD, Lovegrove A, Freeman J, Tryfona T, Pellny TK, Weimar T, Mortimer JC, Stott K, Baker JM, Defoin-Platel M, Shewry PR, Dupree P, Mitchell RAC (2012) Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. Proc Natl Acad Sci U S A 109: 989-993

Anderson CT, Carroll A, Akhmetova L, Somerville C (2010) Real-time imaging of cellulose reorientation during cell wall expansion in Arabidopsis roots. Plant Physiol 152: 787-796

Annis SL, Goodwin PH (1997) Recent advances in the molecular genetics of plant cell wall-degrading enzymes produced by plant pathogenic fungi. Eur J Plant Pathol 103: 1-14

Anthon GE, Barrett DM (2010) Changes in pectin methylesterification and accumulation of methanol during production of diced tomatoes. J Food Eng 97: 367-372

Arioli T, Peng LC, Betzner AS, Burn J, Wittke W, Herth W, Camilleri C, Hofte H, Plazinski J, Birch R, Cork A, Glover J, Redmond J, Williamson RE (1998) Molecular analysis of cellulose biosynthesis in Arabidopsis. Science 279: 717-720

Ashikov A, Routier F, Fuhlrott J, Helmus Y, Wild M, Gerardy-Schahn R, Bakker H (2005) The human solute carrier gene SLC35B4 encodes a bifunctional nucleotide sugar transporter with specificity for UDP-xylose and UDP-N-acetylglucosamine. J Biol Chem 280: 27230-27235

Atmodjo MA, Sakuragi Y, Zhu X, Burrell AJ, Mohanty SS, Atwood JA, 3rd, Orlando R, Scheller HV, Mohnen D (2011) Galacturonosyltransferase (GAUT)1 and GAUT7 are the core of a plant cell wall pectin biosynthetic homogalacturonan:galacturonosyltransferase complex. Proc Natl Acad Sci U S A 108: 20225-20230

Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang JF, Fung P, Gong YC, Wang PW, McCourt P, Guttman DS (2011) Next-generation mapping of Arabidopsis genes. Plant Journal 67: 715-725

Bakker H, Routier F, Ashikov A, Neumann D, Bosch D, Gerardy-Schahn R (2008) A CMP-sialic acid transporter cloned from< i> Arabidopsis thaliana</i>. Carbohydr Res 343: 2148-2152

Bar-Peled M, Griffith CL, Doering TL (2001) Functional cloning and characterization of a UDP-glucuronic acid decarboxylase: The pathogenic fungus Cryptococcus neoformans elucidates UDP-xylose synthesis. Proc Natl Acad Sci U S A 98: 12003-12008

Bar-Peled M, O'Neill MA (2011) Plant Nucleotide Sugar Formation, Interconversion, and Salvage by Sugar Recycling. Annual Review of Plant Biology, Vol 62 62: 127-155

Barber C, Rosti J, Rawat A, Findlay K, Roberts K, Seifert GJ (2006) Distinct properties of the five UDP-D-glucose/UDP-D-galactose 4-epimerase isoforms of Arabidopsis thaliana. Journal of Biological Chemistry 281: 17276-17285

Basu D, Liang Y, Liu X, Himmeldirk K, Faik A, Kieliszewski M, Held M, Showalter AM (2013) Functional Identification of a Hydroxyproline-O-galactosyltransferase Specific for Arabinogalactan Protein Biosynthesis in Arabidopsis. Journal of Biological Chemistry 288: 10132-10143

Bauer S, Vasu P, Mort AJ, Somerville CR (2005) Cloning, expression, and characterization of an oligoxyloglucan reducing end-specific xyloglucanobiohydrolase from Aspergillus nidulans. Carbohydr Res 340: 2590-2597

Bauer S, Vasu P, Persson S, Mort AJ, Somerville CR (2006) Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. Proc Natl Acad Sci U S A 103: 11417-11422

Bauer WD, Talmadge KW, Keegstra K, Albersheim P (1973) The Structure of Plant Cell Walls: II. The Hemicellulose of the Walls of Suspension-cultured Sycamore Cells. Plant Physiol 51: 174-187

Bayliss FT, Ingraham JL (1974) Mutation in Saccharomyces-Cerevisiae Conferring Streptomycin and Cold Sensitivity by Affecting Ribosome Formation and Function. J Bacteriol 118: 319-328

Belfield EJ, Gan X, Mithani A, Brown C, Jiang C, Franklin K, Alvey E, Wibowo A, Jung M, Bailey K, Kalwani S, Ragoussis J, Mott R, Harberd NP (2012) Genome-wide analysis of mutations in mutant lineages selected following fast-neutron irradiation mutagenesis of Arabidopsis thaliana. Genome Res 22: 1306-1315

Bernard P, Couturier M (1992) Cell Killing by the F-Plasmid Ccdb Protein Involves Poisoning of DNA-Topoisomerase-Ii Complexes. J Mol Biol 226: 735-745

Berninsone P, Hwang HY, Zemtseva I, Horvitz HR, Hirschberg CB (2001) SQV-7, a protein involved in Caenorhabditis elegans epithelial invagination and early embryogenesis, transports UDP-glucuronic acid, UDP-N- acetylgalactosamine, and UDP-galactose. Proc Natl Acad Sci U S A 98: 3738-3743

Biely P, MacKenzie C, Puls J, Schneider H (1986) Cooperativity of esterases and xylanases in the enzymatic degradation of acetyl xylan. Nat Biotechnol 4: 731-733

Bischoff V, Selbig J, Scheible WR (2010) Involvement of TBL/DUF231 proteins into cell wall biology. Plant Signal Behav 5: 1057-1059
Bishop PD, Makus DJ, Pearce G, Ryan CA (1981) Proteinase Inhibitor-Inducing Factor Activity in Tomato Leaves Resides in Oligosaccharides Enzymically Released from Cell-Walls. P Natl Acad Sci-Biol 78: 3536-3540
Bleecker AB, Estelle MA, Somerville C, Kende H (1988) Insensitivity to Ethylene Conferred by a Dominant Mutation in Arabidopsis-Thaliana. Science 241: 1086-1089
Boeke JD, Trueheart J, Natsoulis G, Fink GR (1987) 5-Fluoroorotic Acid as a Selective Agent in Yeast Molecular-Genetics. Methods Enzymol 154: 164-175
Bombarely A, Menda N, Tecle IY, Buels RM, Strickler S, Fischer-York T, Pujar A, Leto J, Gosselin J, Mueller LA (2011) The Sol Genomics Network (solgenomics.net): growing tomatoes using Perl. Nucleic Acids Res 39: D1149-1155
Bosca S, Barton CJ, Taylor NG, Ryden P, Neumetzler L, Pauly M, Roberts K, Seifert GJ (2006) Interactions between MUR10/CesA7-dependent secondary cellulose biosynthesis and primary cell wall structure. Plant Physiol 142: 1353-1363
Brown D, Wightman R, Zhang Z, Gomez LD, Atanassov I, Bukowski JP, Tryfona T, McQueen-Mason SJ, Dupree P, Turner S (2011) Arabidopsis genes IRREGULAR XYLEM (IRX15) and IRX15L encode DUF579-containing proteins that are essential for normal xylan deposition in the secondary cell wall. The Plant Journal 66: 401-413
Brown DM, Goubet F, Wong VW, Goodacre R, Stephens E, Dupree P, Turner SR (2007) Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. Plant J 52: 1154-1168
Brown DM, Zeef LAH, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell 17: 2281-2295
Brown DM, Zhang ZN, Stephens E, Dupree P, Turner SR (2009) Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. Plant Journal 57: 732-746
Brumer H, 3rd, Zhou Q, Baumann MJ, Carlsson K, Teeri TT (2004) Activation of crystalline cellulose surfaces through the chemoenzymatic modification of xyloglucan. J Am Chem Soc 126: 5715-5721
Buckeridge MS (2010) Seed Cell Wall Storage Polysaccharides: Models to Understand Cell Wall Biosynthesis and Degradation. Plant Physiol 154: 1017-1023
Buckeridge MS, dos Santos HP, Tine MAS (2000) Mobilisation of storage cell wall polysaccharides in seeds. Plant Physiol Bioch 38: 141-156
Budde RJA, Fang TK, Randall DD, Miernyk JA (1991) Acetyl-Coenzyme-a Can Regulate Activity of the Mitochondrial Pyruvate-Dehydrogenase Complex Insitu. Plant Physiol 95: 131-136
Burget EG, Reiter WD (1999) The mur4 mutant of arabidopsis is partially defective in the de novo synthesis of uridine diphospho L-arabinose. Plant Physiol 121: 383-389
Burton RA, Gidley MJ, Fincher GB (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. Nat Chem Biol 6: 724-732
Carpita NC (2011) Update on mechanisms of plant cell wall biosynthesis: how plants make cellulose and other (1->4)-beta-D-glycans. Plant Physiol 155: 171-184

Carpita NC, Gibeaut DM (1993) Structural Models of Primary-Cell Walls in Flowering Plants - Consistency of Molecular-Structure with the Physical-Properties of the Walls during Growth. Plant Journal 3: 1-30
Cassan-Wang H, Goue N, Saidi MN, Legay S, Sivadon P, Goffner D, Grima-Pettenati J (2013) Identification of novel transcription factors regulating secondary cell wall formation in Arabidopsis. Front Plant Sci 4: 189
Cavalier DM, Keegstra K (2006) Two xyloglucan xylosyltransferases catalyze the addition of multiple xylosyl residues to cellohexaose. Journal of Biological Chemistry 281: 34197-34207
Cavalier DM, Lerouxel O, Neumetzler L, Yamauchi K, Reinecke A, Freshour G, Zabotina OA, Hahn MG, Burgert I, Pauly M, Raikhel NV, Keegstra K (2008) Disrupting two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. Plant Cell 20: 1519-1537
Chambat G, Karmous M, Costes M, Picard M, Joseleau J-P (2005) Variation of xyloglucan substitution pattern affects the sorption on celluloses with different degrees of crystallinity. Cellulose 12: 117-125
Chen Z, Hong X, Zhang H, Wang Y, Li X, Zhu JK, Gong Z (2005) Disruption of the cellulose synthase gene, AtCesA8/IRX1, enhances drought and osmotic stress tolerance in Arabidopsis. Plant J 43: 273-283
Cheng K, Sorek H, Zimmermann H, Wemmer DE, Pauly M (2013) Solution-state 2D NMR spectroscopy of plant cell walls enabled by a dimethylsulfoxide-d6/1-ethyl-3-methylimidazolium acetate solvent. Anal Chem 85: 3213-3221
Cho H-T, Cosgrove DJ (2000) Altered expression of expansin modulates leaf growth and pedicel abscission in Arabidopsis thaliana. Proceedings of the National Academy of Sciences 97: 9783-9788
Chou YH, Pogorelko G, Zabotina OA (2012) Xyloglucan xylosyltransferases XXT1, XXT2, and XXT5 and the glucan synthase CSLC4 form Golgi-localized multiprotein complexes. Plant Physiol 159: 1355-1366
Ciucanu I, Kerek F (1984) A Simple and Rapid Method for the Permethylation of Carbohydrates. Carbohydr Res 131: 209-217
Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant Journal 16: 735-743
Cocuron JC, Lerouxel O, Drakakaki G, Alonso AP, Liepman AH, Keegstra K, Raikhel N, Wilkerson CG (2007) A gene from the cellulose synthase-like C family encodes a beta-1,4 glucan synthase. Proc Natl Acad Sci U S A 104: 8550-8555
Cosgrove DJ (1989) Characterization of Long-Term Extension of Isolated Cell-Walls from Growing Cucumber Hypocotyls. Planta 177: 121-130
Cosgrove DJ (1998) Cell wall loosening by expansins. Plant Physiol 118: 333-339
Cosgrove DJ (2000) Loosening of plant cell walls by expansins. Nature 407: 321-326
Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol 6: 850-861
Coutu C, Brandle J, Brown D, Brown K, Miki B, Simmonds J, Hegedus DD (2007) pORE: a modular binary vector series suited for both monocot and dicot plant transformation. Transgenic Res 16: 771-781
Cregg JM, Barringer KJ, Hessler AY, Madden KR (1985) Pichia-Pastoris as a Host System for Transformations. Mol Cell Biol 5: 3376-3385

Cregg JM, Madden KR (1989) Use of Site-Specific Recombination to Regenerate Selectable Markers. Molecular & General Genetics 219: 320-323
Crombie HJ, Chengappa S, Hellyer A, Reid JS (1998) A xyloglucan oligosaccharide-active, transglycosylating beta-D-glucosidase from the cotyledons of nasturtium (Tropaeolum majus L) seedlings--purification, properties and characterization of a cDNA clone. Plant J 15: 27-38
Crouch ML, Sussex IM (1981) Development and storage-protein synthesis in Brassica napus L. embryos in vivo and in vitro. Planta 153: 64-74
Curtis IS, Nam HG (2001) Transgenic radish (Raphanus sativus L. longipinnatus Bailey) by floral-dip method--plant development and surfactant are important in optimizing transformation efficiency. Transgenic Res 10: 363-371
Curtis MD, Grossniklaus U (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. Plant Physiol 133: 462-469
Davis J, Brandizzi F, Liepman AH, Keegstra K (2010) Arabidopsis mannan synthase CSLA9 and glucan synthase CSLC4 have opposite orientations in the Golgi membrane. Plant J 64: 1028-1037
Desilva J, Jarman CD, Arrowsmith DA, Stronach MS, Chengappa S, Sidebottom C, Reid JSG (1993) Molecular Characterization of a Xyloglucan-Specific Endo-(1-]4)-Beta-D-Glucanase (Xyloglucan Endotransglycosylase) from Nasturtium Seeds. Plant Journal 3: 701-711
Desnos T, Orbovic V, Bellini C, Kronenberger J, Caboche M, Traas J, Hofte H (1996) Procuste1 mutants identify two distinct genetic pathways controlling hypocotyl cell elongation, respectively in dark- and light-grown Arabidopsis seedlings. Development 122: 683-693
Desprez T, Juraniec M, Crowell EF, Jouy H, Pochylova Z, Parcy F, Hofte H, Gonneau M, Vernhettes S (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. Proc Natl Acad Sci U S A 104: 15572-15577
Desveaux D, Faik A, Maclachlan G (1998) Fucosyltransferase and the biosynthesis of storage and structural xyloglucan in developing nasturtium fruits. Plant Physiol 118: 885-894
Dhugga KS, Barreiro R, Whitten B, Stecca K, Hazebroek J, Randhawa GS, Dolan M, Kinney AJ, Tomes D, Nichols S, Anderson P (2004) Guar seed beta-mannan synthase is a member of the cellulose synthase super gene family. Science 303: 363-366
Doblin MS, De Melis L, Newbigin E, Bacic A, Read SM (2001) Pollen tubes of Nicotiana alata express two genes from different beta-glucan synthase families. Plant Physiol 125: 2040-2052
Doblin MS, Pettolino FA, Wilson SM, Campbell R, Burton RA, Fincher GB, Newbigin E, Bacic A (2009) A barley cellulose synthase-like CSLH gene mediates (1,3;1,4)-beta-D-glucan synthesis in transgenic Arabidopsis. Proc Natl Acad Sci U S A 106: 5996-6001
Dougherty BA, Vanderijn I (1993) Molecular Characterization of Hasb from an Operon Required for Hyaluronic-Acid Synthesis in Group-a Streptococci - Demonstration of Udp-Glucose Dehydrogenase-Activity. Journal of Biological Chemistry 268: 7118-7124
Drake JW, Baltz RH (1976) The biochemistry of mutagenesis. Annu Rev Biochem 45: 11-37
Dumville JC, Fry SC (2000) Uronic acid-containing oligosaccharins: Their biosynthesis, degradation and signalling roles in non-diseased plant tissues. Plant Physiol Bioch 38: 125-140
Dunwell JM, Purvis A, Khuri S (2004) Cupins: the most functionally diverse protein superfamily? Phytochemistry 65: 7-17
Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5: 113

Edwards M, Dea IC, Bulpin PV, Reid JG (1985) Xyloglucan (amyloid) mobilisation in the cotyledons of Tropaeolum majus L. seeds following germination. Planta 163: 133-140
Edwards M, Dea ICM, Bulpin PV, Reid JSG (1986) Purification and Properties of a Novel Xyloglucan-Specific Endo-(1-]4)-Beta-D-Glucanase from Germinated Nasturtium Seeds (Tropaeolum-Majus L). Journal of Biological Chemistry 261: 9489-9494
Edwards ME, Dickson CA, Chengappa S, Sidebottom C, Gidley MJ, Reid JSG (1999) Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis. Plant Journal 19: 691-697
Eklof JM, Brumer H (2010) The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling. Plant Physiol 153: 456-466
El-Battari A, Prorok M, Angata K, Mathieu S, Zerfaoui M, Ong E, Suzuki M, Lombardo D, Fukuda M (2003) Different glycosyltransferases are differentially processed for secretion, dimerization, and autoglycosylation. Glycobiology 13: 941-953
Faik A (2010) Xylan biosynthesis: news from the grass. Plant Physiol 153: 396-402
Faik A, Price NJ, Raikhel NV, Keegstra K (2002) An Arabidopsis gene encoding an alpha-xylosyltransferase involved in xyloglucan biosynthesis. Proc Natl Acad Sci U S A 99: 7797-7802
Fanutti C, Gidley MJ, Reid JS (1996) Substrate subsite recognition of the xyloglucan endo-transglycosylase or xyloglucan-specific endo-(1-->4)-beta-D-glucanase from the cotyledons of germinated nasturtium (Tropaeolum majus L.) seeds. Planta 200: 221-228
Fanutti C, Gidley MJ, Reid JSG (1993) Action of a Pure Xyloglucan Endo-Transglycosylase (Formerly Called Xyloglucan-Specific Endo-(1-]4)-Beta-D-Glucanase) from the Cotyledons of Germinated Nasturtium Seeds. Plant Journal 3: 691-700
Farkas V, Sulova Z, Stratilova E, Hanna R, Maclachlan G (1992) Cleavage of xyloglucan by nasturtium seed xyloglucanase and transglycosylation to xyloglucan subunit oligosaccharides. Arch Biochem Biophys 298: 365-370
Fatland BL, Ke JS, Anderson MD, Mentzen WI, Cui LW, Allred CC, Johnston JL, Nikolau BJ, Wurtele ES (2002) Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in Arabidopsis. Plant Physiol 130: 740-756
Fatland BL, Nikolau BJ, Wurtele ES (2005) Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP-citrate lyase in Arabidopsis. Plant Cell 17: 182-203
Fink H, Ahrenstedt L, Bodin A, Brumer H, Gatenholm P, Krettek A, Risberg B (2011) Bacterial cellulose modified with xyloglucan bearing the adhesion peptide RGD promotes endothelial cell adhesion and metabolism--a promising modification for vascular grafts. J Tissue Eng Regen Med 5: 454-463
Frandsen GI, Mundy J, Tzen JT (2001) Oil bodies and their associated proteins, oleosin and caleosin. Physiol Plantarum 112: 301-307
Fry SC (1989) The Structure and Functions of Xyloglucan. J Exp Bot 40: 1-11
Fry SC, Aldington S, Hetherington PR, Aitken J (1993a) Oligosaccharides as Signals and Substrates in the Plant-Cell Wall. Plant Physiol 103: 1-5
Fry SC, Smith RC, Renwick KF, Martin DJ, Hodge SK, Matthews KJ (1992) Xyloglucan Endotransglycosylase, a New Wall-Loosening Enzyme-Activity from Plants. Biochemical Journal 282: 821-828
Fry SC, York WS, Albersheim P, Darvill A, Hayashi T, Joseleau JP, Kato Y, Lorences EP, Maclachlan GA, Mcneil M, Mort AJ, Reid JSG, Seitz HU, Selvendran RR, Voragen AGJ, White AR (1993b) An

Unambiguous Nomenclature for Xyloglucan-Derived Oligosaccharides. Physiol Plantarum 89: 1-3
Gao CJ, Yu CKY, Qu S, San MWY, Li KY, Lo SW, Jiang LW (2012) The Golgi-Localized Arabidopsis Endomembrane Protein12 Contains Both Endoplasmic Reticulum Export and Golgi Retention Signals at Its C Terminus C. Plant Cell 24: 2086-2104
Gari E, Piedrafita L, Aldea M, Herrero E (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. Yeast 13: 837-848
Geshi N, Johansen JN, Dilokpimol A, Rolland A, Belcram K, Verger S, Kotake T, Tsumuraya Y, Kaneko S, Tryfona T, Dupree P, Scheller HV, Hofte H, Mouille G (2013) A galactosyltransferase acting on arabinogalactan protein glycans is essential for embryo development in Arabidopsis. Plant J
Gidley MJ, Lillford PJ, Rowlands DW, Lang P, Dentini M, Crescenzi V, Edwards M, Fanutti C, Reid JSG (1991) Structure and Solution Properties of Tamarind-Seed Polysaccharide. Carbohydr Res 214: 299-314
Gille S, Cheng K, Skinner ME, Liepman AH, Wilkerson CG, Pauly M (2011a) Deep sequencing of voodoo lily (Amorphophallus konjac): an approach to identify relevant genes involved in the synthesis of the hemicellulose glucomannan. Planta 234: 515-526
Gille S, de Souza A, Xiong G, Benz M, Cheng K, Schultink A, Reca IB, Pauly M (2011b) O-acetylation of Arabidopsis hemicellulose xyloglucan requires AXY4 or AXY4L, proteins with a TBL and DUF231 domain. Plant Cell 23: 4041-4053
Gille S, Hansel U, Ziemann M, Pauly M (2009) Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. Proc Natl Acad Sci U S A 106: 14699-14704
Giovane A, Servillo L, Balestrieri C, Raiola A, D'Avino R, Tamburrini M, Ciardiello MA, Camardella L (2004) Pectin methylesterase inhibitor. Bba-Proteins Proteom 1696: 245-252
Gossen M, Bujard H (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc Natl Acad Sci U S A 89: 5547-5551
Goubet F, Misrahi A, Park SK, Zhang Z, Twell D, Dupree P (2003) AtCSLA7, a cellulose synthase-like putative glycosyltransferase, is important for pollen tube growth and embryogenesis in Arabidopsis. Plant Physiol 131: 547-557
Goujon T, Minic Z, El Amrani A, Lerouxel O, Aletti E, Lapierre C, Joseleau JP, Jouanin L (2003) AtBXL1, a novel higher plant (Arabidopsis thaliana) putative beta-xylosidase gene, is involved in secondary cell wall metabolism and plant development. Plant J 33: 677-690
Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27: 221-224
Gu Y, Kaplinsky N, Bringmann M, Cobb A, Carroll A, Sampathkumar A, Baskin TI, Persson S, Somerville CR (2010) Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. Proc Natl Acad Sci U S A 107: 12866-12871
Guerriero G, Fugelstad J, Bulone V (2010) What Do We Really Know about Cellulose Biosynthesis in Higher Plants? J Integr Plant Biol 52: 161-175
Guillon F, Bouchet B, Jamme F, Robert P, Quemener B, Barron C, Larre C, Dumas P, Saulnier L (2011) Brachypodium distachyon grain: characterization of endosperm cell walls. J Exp Bot 62: 1001-1015

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307-321
Gunl M, Neumetzler L, Kraemer F, de Souza A, Schultink A, Pena M, York WS, Pauly M (2011) AXY8 encodes an alpha-fucosidase, underscoring the importance of apoplastic metabolism on the fine structure of Arabidopsis cell wall polysaccharides. Plant Cell 23: 4025-4040
Gunl M, Pauly M (2011) AXY3 encodes a alpha-xylosidase that impacts the structure and accessibility of the hemicellulose xyloglucan in Arabidopsis plant cell walls. Planta 233: 707-719
Habib H, Fazili KM (2007) Plant protease inhibitors: a defense strategy in plants. Biotechnol Mol Biol Rev 2: 68-85
Hahn MG, Darvill AG, Albersheim P (1981) Host-Pathogen Interactions .19. The Endogenous Elicitor, a Fragment of a Plant-Cell Wall Polysaccharide That Elicits Phytoalexin Accumulation in Soybeans. Plant Physiol 68: 1161-1169
Hajdukiewicz P, Svab Z, Maliga P (1994) The Small, Versatile Ppzp Family of Agrobacterium Binary Vectors for Plant Transformation. Plant Mol Biol 25: 989-994
Hall Q, Cannon MC (2002) The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in Arabidopsis. Plant Cell 14: 1161-1172
Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang YW, Rios S, Bobrowicz P, Stadheim TA, Li HJ, Choi BK, Hopkins D, Wischnewski H, Roser J, Mitchell T, Strawbridge RR, Hoopes J, Wildt S, Gerngross TU (2006) Humanization of yeast to produce complex terminally sialylated glycoproteins. Science 313: 1441-1443
Harholt J, Jensen JK, Sorensen SO, Orfila C, Pauly M, Scheller HV (2006) ARABINAN DEFICIENT 1 is a putative arabinosyltransferase involved in biosynthesis of Pectic Arabinan in Arabidopsis. Plant Physiol 140: 49-58
Harholt J, Jensen JK, Verhertbruggen Y, Sogaard C, Bernard S, Nafisi M, Poulsen CP, Geshi N, Sakuragi Y, Driouich A, Knox JP, Scheller HV (2012) ARAD proteins associated with pectic Arabinan biosynthesis form complexes when transiently overexpressed in planta. Planta 236: 115-128
Harholt J, Suttangkakul A, Vibe Scheller H (2010) Biosynthesis of pectin. Plant Physiol 153: 384-395
Hauser MT, Morikami A, Benfey PN (1995) Conditional Root Expansion Mutants of Arabidopsis. Development 121: 1237-1252
Hayashi T (1989) Xyloglucans in the Primary-Cell Wall. Annu Rev Plant Phys 40: 139-168
Hayashi T, Kaida R (2011) Functions of Xyloglucan in Plant Cells. Mol Plant 4: 17-24
Hematy K, Cherk C, Somerville S (2009) Host-pathogen warfare at the plant cell wall. Curr Opin Plant Biol 12: 406-413
Herbers K, Lorences EP, Barrachina C, Sonnewald U (2001) Functional characterisation of Nicotiana tabacum xyloglucan endotransglycosylase (NtXET-1): generation of transgenic tobacco plants and changes in cell wall xyloglucan. Planta 212: 279-287
Heyer WD, Ehmsen KT, Liu J (2010) Regulation of Homologous Recombination in Eukaryotes. Annu Rev Genet 44: 113-139

Hoffman M, Jia Z, Pena MJ, Cash M, Harper A, Blackburn AR, 2nd, Darvill A, York WS (2005) Structural analysis of xyloglucans in the primary cell walls of plants in the subclass Asteridae. Carbohydr Res 340: 1826-1840
Hong K, Ma D, Beverley SM, Turco SJ (2000) The Leishmania GDP-mannose transporter is an autonomous, multi-specific, hexameric complex of LPG2 subunits. Biochemistry-Us 39: 2013-2022
Hoth A, Blaschek W, Franz G (1986) Xyloglucan (amyloid) formation in the cotyledons of Tropaeolum majus L. seeds. Plant Cell Rep 5: 9-12
Huckelhoven R (2007) Cell wall - Associated mechanisms of disease resistance and susceptibility. Annu Rev Phytopathol 45: 101-127
Huynh QK, Hironaka CM, Levine EB, Smith C, Borgmeyer J, Shah D (1992) Antifungal proteins from plants. Purification, molecular cloning, and antifungal properties of chitinases from maize seed. Journal of Biological Chemistry 267: 6635-6640
Indiveri C, Iacobazzi V, Giangregorio N, Palmieri F (1997) The mitochondrial carnitine carrier protein: cDNA cloning, primary structure and comparison with other mitochondrial transport proteins. Biochemical Journal 321: 713-719
Iwai H, Masaoka N, Ishii T, Satoh S (2002) A pectin glucuronyltransferase gene is essential for intercellular attachment in the plant meristem. Proc Natl Acad Sci U S A 99: 16319-16324
Jarvis MC (1984) Structure and Properties of Pectin Gels in Plant-Cell Walls. Plant Cell Environ 7: 153-164
Jensen JK, Kim H, Cocuron JC, Orler R, Ralph J, Wilkerson CG (2011) The DUF579 domain containing proteins IRX15 and IRX15-L affect xylan synthesis in Arabidopsis. Plant J 66: 387-400
Jensen JK, Schultink A, Keegstra K, Wilkerson CG, Pauly M (2012) RNA-Seq analysis of developing nasturtium seeds (Tropaeolum majus): identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis. Mol Plant 5: 984-992
Jensen JK, Sorensen SO, Harholt J, Geshi N, Sakuragi Y, Moller I, Zandleven J, Bernal AJ, Jensen NB, Sorensen C, Pauly M, Beldman G, Willats WGT, Scheller HV (2008) Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in Arabidopsis. Plant Cell 20: 1289-1302
Jia Z, Qin Q, Darvill AG, York WS (2003) Structure of the xyloglucan produced by suspension-cultured tomato cells. Carbohydr Res 338: 1197-1208
Jonas MC, Pehar M, Puglielli L (2010) AT-1 is the ER membrane acetyl-CoA transporter and is essential for cell viability. J Cell Sci 123: 3378-3388
Juge N, Payan F, Williamson G (2004) XIP-I, a xylanase inhibitor protein from wheat: a novel protein function. Biochim Biophys Acta 1696: 203-211
Kato Y, Ito S, Mitsuishi Y (2004) Study on the structures of xyloglucans using xyloglucan specific enzymes. Trends Glycosci Glyc 16: 393-406
Keegstra K (2010) Plant Cell Walls. Plant Physiol 154: 483-486
Keppler BD, Showalter AM (2010) IRX14 and IRX14-LIKE, two glycosyl transferases involved in glucuronoxylan biosynthesis and drought tolerance in Arabidopsis. Mol Plant 3: 834-841
Kerr EM, Fry SC (2003) Pre-formed xyloglucans and xylans increase in molecular weight in three distinct compartments of a maize cell-suspension culture. Planta 217: 327-339

Kiefer LL, York WS, Darvill AG, Albersheim P (1989) Structure of Plant-Cell Walls .27. Xyloglucan Isolated from Suspension-Cultured Sycamore Cell-Walls Is O-Acetylated. Phytochemistry 28: 2105-2107
Kircher M, Kelso J (2010) High-throughput DNA sequencing--concepts and limitations. Bioessays 32: 524-536
Kispal G, Sumegi B, Dietmeier K, Bock I, Gajdos G, Tomcsanyi T, Sandor A (1993) Cloning and sequencing of a cDNA encoding Saccharomyces cerevisiae carnitine acetyltransferase. Use of the cDNA in gene disruption studies. J Biol Chem 268: 1824-1829
Knox JP (2008) Revealing the structural and functional diversity of plant cell walls. Curr Opin Plant Biol 11: 308-313
Kobayashi T, Sleeman JE, Coughtrie MW, Burchell B (2006) Molecular and functional characterization of microsomal UDP-glucuronic acid uptake by members of the nucleotide sugar transporter (NST) family. Biochem J 400: 281-289
Koller A, Valesco J, Subramani S (2000) The CUPI promoter of Saccharomyces cerevisiae is inducible by copper in Pichia pastoris. Yeast 16: 651-656
Konieczny A, Ausubel FM (1993) A Procedure for Mapping Arabidopsis Mutations Using Codominant Ecotype-Specific Pcr-Based Markers. Plant Journal 4: 403-410
Kooiman P (1971) Structures of the galactomannans from seeds of Annona muricata, Arenga saccharifera, Cocos nucifera, Convolvulus tricolor, and Sophora japonica. Carbohydr Res 20: 329-337
Koornneef M, Bosma T, Hanhart C, Van der Veen J, Zeevaart J (1990) The isolation and characterization of gibberellin-deficient mutants in tomato. Theoretical and Applied Genetics 80: 852-857
Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567-580
Krysan PJ, Young JC, Sussman MR (1999) T-DNA as an insertional mutagen in Arabidopsis. Plant Cell 11: 2283-2290
Laitinen RAE, Schneeberger K, Jelly NS, Ossowski S, Weigel D (2010) Identification of a Spontaneous Frame Shift Mutation in a Nonreference Arabidopsis Accession Using Whole Genome Sequencing. Plant Physiol 153: 652-654
Lawand S, Dorne AJ, Long D, Coupland G, Mache R, Carol P (2002) Arabidopsis a bout de souffle, which is homologous with mammalian carnitine acyl carrier, is required for postembryonic growth in the light. Plant Cell 14: 2161-2173
Le Dizet P (1972) [Analysis of the structure of amyloid from nasturtiums]. Carbohydr Res 24: 505-509
Lederberg J (1951) Streptomycin Resistance - a Genetically Recessive Mutation. J Bacteriol 61: 549-550
Lee C, Teng Q, Zhong R, Ye ZH (2011) The four Arabidopsis reduced wall acetylation genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. Plant Cell Physiol 52: 1289-1301
Leonard R, Pabst M, Bondili JS, Chambat G, Veit C, Strasser R, Altmann F (2008) Identification of an Arabidopsis gene encoding a GH95 alpha1,2-fucosidase active on xyloglucan oligo- and polysaccharides. Phytochemistry 69: 1983-1988

Lerouxel O, Cavalier DM, Liepman AH, Keegstra K (2006) Biosynthesis of plant cell wall polysaccharides - a complex process. Curr Opin Plant Biol 9: 621-630
Lerouxel O, Choo TS, Seveno M, Usadel B, Faye L, Lerouge P, Pauly M (2002) Rapid structural phenotyping of plant cell wall mutants by enzymatic oligosaccharide fingerprinting. Plant Physiol 130: 1754-1763
Levy S, Maclachlan G, Staehelin LA (1997) Xyloglucan sidechains modulate binding to cellulose during in vitro binding assays as predicted by conformational dynamics simulations. Plant Journal 11: 373-386
Li X, Cordero I, Caplan J, Molhoj M, Reiter WD (2004) Molecular analysis of 10 coding regions from Arabidopsis that are homologous to the MUR3 xyloglucan galactosyltransferase. Plant Physiol 134: 940-950
Liepman AH, Wilkerson CG, Keegstra K (2005) Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. Proc Natl Acad Sci U S A 102: 2221-2226
Liu J-JJ, Krenz DC, Galvez AF, de Lumen BO (1998) Galactinol synthase (GS): increased enzyme activity and levels of mRNA due to cold and desiccation. Plant Sci 134: 11-20
Liu YB, Lu SM, Zhang JF, Liu S, Lu YT (2007) A xyloglucan endotransglucosylase/hydrolase involves in growth of primary root and alters the deposition of cellulose in Arabidopsis. Planta 226: 1547-1560
Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ (2012) Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol 30: 434-+
Long D, Martin M, Sundberg E, Swinburne J, Puangsomlee P, Coupland G (1993) The maize transposable element system Ac/Ds as a mutagen in Arabidopsis: identification of an albino mutation induced by Ds insertion. Proc Natl Acad Sci U S A 90: 10370-10374
Lu C, Kang J (2008) Generation of transgenic plants of a potential oilseed crop Camelina sativa by Agrobacterium-mediated transformation. Plant Cell Rep 27: 273-278
Luhn K, Wild MK, Eckhardt M, Gerardy-Schahn R, Vestweber D (2001) The gene defective in leukocyte adhesion deficiency II encodes a putative GDP-fucose transporter. Nat Genet 28: 69-72
Lunde CF, Morrow DJ, Roy LM, Walbot V (2003) Progress in maize gene discovery: a project update. Functional & integrative genomics 3: 25-32
Mabusela WT, Stephen AM, Botha MC (1990) Carbohydrate Polymers from Aloe-Ferox Leaves. Phytochemistry 29: 3555-3558
Madson M, Dunand C, Li X, Verma R, Vanzin GF, Caplan J, Shoue DA, Carpita NC, Reiter WD (2003) The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. Plant Cell 15: 1662-1670
Malik K, Wu K, Li XQ, Martin-Heller T, Hu M, Foster E, Tian L, Wang C, Ward K, Jordan M, Brown D, Gleddie S, Simmonds D, Zheng S, Simmonds J, Miki B (2002) A constitutive gene expression system derived from the tCUP cryptic promoter elements. Theor Appl Genet 105: 505-514
Manabe Y, Nafisi M, Verhertbruggen Y, Orfila C, Gille S, Rautengarten C, Cherk C, Marcus SE, Somerville S, Pauly M, Knox JP, Sakuragi Y, Scheller HV (2011) Loss-of-function mutation of REDUCED WALL ACETYLATION2 in Arabidopsis leads to reduced cell wall acetylation and increased resistance to Botrytis cinerea. Plant Physiol 155: 1068-1078

Marcus SE, Verhertbruggen Y, Herve C, Ordaz-Ortiz JJ, Farkas V, Pedersen HL, Willats WGT, Knox JP (2008) Pectic homogalacturonan masks abundant sets of xyloglucan epitopes in plant cell walls. Bmc Plant Biol 8
Mardis ER (2008) The impact of next-generation sequencing technology on genetics. Trends Genet 24: 133-141
Maris A, Kaewthai N, Eklof JM, Miller JG, Brumer H, Fry SC, Verbelen JP, Vissenberg K (2011) Differences in enzymic properties of five recombinant xyloglucan endotransglucosylase/hydrolase (XTH) proteins of Arabidopsis thaliana. J Exp Bot 62: 261-271
Marlett JA, McBurney MI, Slavin JL (2002) Position of the American Dietetic Association: health implications of dietary fiber. J Am Diet Assoc 102: 993-1000
Maruyama K, Goto C, Numata M, Suzuki T, Nakagawa Y, Hoshino T, Uchiyama T (1996) O-acetylated xyloglucan in extracellular polysaccharides from cell-suspension cultures of Mentha. Phytochemistry 41: 1309-1314
Marx H, Mattanovich D, Sauer M (2008) Overexpression of the riboflavin biosynthetic pathway in Pichia pastoris. Microb Cell Fact 7
Masterson C, Wood C (2000) Pea chloroplast carnitine acetyltransferase. P Roy Soc B-Biol Sci 267: 1-6
McCormac AC, Elliott MC, Chen DF (1998) A simple method for the production of highly competent cells of Agrobacterium for transformation via electroporation. Mol Biotechnol 9: 155-159
Mcdougall GJ, Fry SC (1989) Structure-Activity-Relationships for Xyloglucan Oligosaccharides with Antiauxin Activity. Plant Physiol 89: 883-887
Mcdougall GJ, Fry SC (1990) Xyloglucan Oligosaccharides Promote Growth and Activate Cellulase - Evidence for a Role of Cellulase in Cell Expansion. Plant Physiol 93: 1042-1048
McKinney EC, Ali N, Traut A, Feldmann KA, Belostotsky DA, McDowell JM, Meagher RB (1995) Sequence-based identification of T-DNA insertion mutations in Arabidopsis: actin mutants act2-1 and act4-1. Plant J 8: 613-622
McNeil M, Darvill AG, Fry SC, Albersheim P (1984) Structure and function of the primary cell walls of plants. Annu Rev Biochem 53: 625-663
Min Kuo T, Lowell CA, Smith PT (1997) Changes in soluble carbohydrates and enzymic activities in maturing soybean seed tissues. Plant Sci 125: 1-11
Minic Z, Rihouey C, Do CT, Lerouge P, Jouanin L (2004) Purification and characterization of enzymes exhibiting β-D-xylosidase activities in stem tissues of Arabidopsis. Plant Physiol 135: 867-878
Mishra A, Malhotra AV (2009) Tamarind xyloglucan: a polysaccharide with versatile application potential. J Mater Chem 19: 8528-8536
Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol 11: 266-277
Moore PJ, Staehelin LA (1988) Immunogold localization of the cell-wall-matrix polysaccharides rhamnogalacturonan I and xyloglucan during cell expansion and cytokinesis inTrifolium pratense L.; implication for secretory pathways. Planta 174: 433-445
Moreira LR, Filho EX (2008) An overview of mannan structure and mannan-degrading enzyme systems. Appl Microbiol Biotechnol 79: 165-178
Moremen KW, Tiemeyer M, Nairn AV (2012) Vertebrate protein glycosylation: diversity, synthesis and function. Nat Rev Mol Cell Biol 13: 448-462

Morozova O, Hirst M, Marra MA (2009) Applications of new sequencing technologies for transcriptome analysis. Annu Rev Genomics Hum Genet 10: 135-151
Morse AM, Whetten RW, Dubos C, Campbell MM (2009) Post-translational modification of an R2R3-MYB transcription factor by a MAP Kinase during xylem development. New Phytologist 183: 1001-1013
Mortimer JC, Miles GP, Brown DM, Zhang Z, Segura MP, Weimar T, Yu X, Seffen KA, Stephens E, Turner SR, Dupree P (2010) Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. Proc Natl Acad Sci U S A 107: 17409-17414
Motose H, Sugiyama M, Fukuda H (2004) A proteoglycan mediates inductive interaction during plant vascular development. Nature 429: 873-878
Mouille G, Robin S, Lecomte M, Pagant S, Hofte H (2003) Classification and identification of Arabidopsis cell wall mutants using Fourier-Transform InfraRed (FT-IR) microspectroscopy. Plant Journal 35: 393-404
Mueller SC, Brown RM (1980) Evidence for an Intramembrane Component Associated with a Cellulose Microfibril-Synthesizing Complex in Higher-Plants. J Cell Biol 84: 315-326
Muller K, Levesque-Tremblay G, Bartels S, Weitbrecht K, Wormit A, Usadel B, Haughn G, Kermode AR (2013) Demethylesterification of cell wall pectins in Arabidopsis plays a role in seed germination. Plant Physiol 161: 305-316
Muraoka M, Kawakita M, Ishida N (2001) Molecular characterization of human UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter, a novel nucleotide sugar transporter with dual substrate specificity. Febs Lett 495: 87-93
Naatsaari L, Mistlberger B, Ruth C, Hajek T, Hartner FS, Glieder A (2012) Deletion of the Pichia pastoris KU70 Homologue Facilitates Platform Strain Generation for Gene Expression and Synthetic Biology. PLoS One 7
Nakagawa T, Kurose T, Hino T, Tanaka K, Kawamukai M, Niwa Y, Toyooka K, Matsuoka K, Jinbo T, Kimura T (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. J Biosci Bioeng 104: 34-41
Nelson BK, Cai X, Nebenfuhr A (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. Plant Journal 51: 1126-1136
Nguema-Ona E, Andeme-Onzighi C, Aboughe-Angone S, Bardor M, Ishii T, Lerouge P, Driouich A (2006) The reb1-1 mutation of Arabidopsis. Effect on the structure and localization of galactose-containing cell wall polysaccharides. Plant Physiol 140: 1406-1417
Niemann C, Carpita NC, Whistler RL (1997) Arabinose-Containing Oligosaccharides from Tamarind Xyloglucan. Starch-Stärke 49: 154-159
Ninomiya Y, Suzuki K, Ishii C, Inoue H (2004) Highly efficient gene replacements in Neurospora strains deficient for nonhomologous end-joining. Proc Natl Acad Sci U S A 101: 12248-12253
Nishikubo N, Takahashi J, Roos AA, Derba-Maceluch M, Piens K, Brumer H, Teeri TT, Stalbrand H, Mellerowicz EJ (2011) Xyloglucan endo-Transglycosylase-Mediated Xyloglucan Rearrangements in Developing Wood of Hybrid Aspen. Plant Physiol 155: 399-413
Nishitani K, Tominaga R (1992) Endoxyloglucan Transferase, a Novel Class of Glycosyltransferase That Catalyzes Transfer of a Segment of Xyloglucan Molecule to Another Xyloglucan Molecule. Journal of Biological Chemistry 267: 21058-21064

Noguchi M, Yu D, Hirayama R, Ninomiya Y, Sekine E, Kubota N, Ando K, Okayasu R (2006) Inhibition of homologous recombination repair in irradiated tumor cells pretreated with Hsp90 inhibitor 17-allylamino-17-demethoxygeldanamycin. Biochem Biophys Res Commun 351: 658-663

Oka T, Jigami Y (2006) Reconstruction of de novo pathway for synthesis of UDP-glucuronic acid and UDP-xylose from intrinsic UDP-glucose in Saccharomyces cerevisiae. Febs J 273: 2645-2657

Oliver DJ, Nikolau BJ, Wurtele ES (2009) Acetyl-CoA-Life at the metabolic nexus. Plant Sci 176: 597-601

Ossowski S, Schneeberger K, Lucas-Lledo JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M (2010) The Rate and Molecular Spectrum of Spontaneous Mutations in Arabidopsis thaliana. Science 327: 92-94

Page DR, Grossniklaus L (2002) The art and design of genetic screens: Arabidopsis thaliana. Nat Rev Genet 3: 124-136

Pande SV, Parvin R (1980) Carnitine-acylcarnitine translocase catalyzes an equilibrating unidirectional transport as well. J Biol Chem 255: 2994-3001

Park YB, Cosgrove DJ (2012a) Changes in Cell Wall Biomechanical Properties in the Xyloglucan-Deficient xxt1/xxt2 Mutant of Arabidopsis. Plant Physiol 158: 465-475

Park YB, Cosgrove DJ (2012b) A Revised Architecture of Primary Cell Walls Based on Biomechanical Changes Induced by Substrate-Specific Endoglucanases. Plant Physiol 158: 1933-1943

Paterson AH, Freeling M, Tang H, Wang X (2010) Insights from the comparison of plant genome sequences. Annual review of plant biology 61: 349-372

Pattathil S, Avci U, Baldwin D, Swennes AG, McGill JA, Popper Z, Bootten T, Albert A, Davis RH, Chennareddy C, Dong RH, O'Shea B, Rossi R, Leoff C, Freshour G, Narra R, O'Neil M, York WS, Hahn MG (2010) A Comprehensive Toolkit of Plant Cell Wall Glycan-Directed Monoclonal Antibodies. Plant Physiol 153: 514-525

Pauly M, Albersheim P, Darvill A, York WS (1999a) Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. Plant Journal 20: 629-639

Pauly M, Andersen LN, Kauppinen S, Kofod LV, York WS, Albersheim P, Darvill A (1999b) A xyloglucan-specific endo-beta-1,4-glucanase from Aspergillus aculeatus: expression cloning in yeast, purification and characterization of the recombinant enzyme. Glycobiology 9: 93-100

Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink A, Xiong G (2013) Hemicellulose biosynthesis. Planta

Pauly M, Keegstra K (2010) Plant cell wall polymers as precursors for biofuels. Curr Opin Plant Biol 13: 305-312

Pauly M, Scheller HV (2000) O-Acetylation of plant cell wall polysaccharides: identification and partial characterization of a rhamnogalacturonan O-acetyl-transferase from potato suspension-cultured cells. Planta 210: 659-667

Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM (1996) Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. Proc Natl Acad Sci U S A 93: 12637-12642

Pena MJ, Darvill AG, Eberhard S, York WS, O'Neill MA (2008) Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. Glycobiology 18: 891-904
Pena MJ, Kong Y, York WS, O'Neill MA (2012) A galacturonic Acid-containing xyloglucan is involved in Arabidopsis root hair tip growth. Plant Cell 24: 4511-4524
Perez S, Rodriguez-Carvajal MA, Doco T (2003) A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function. Biochimie 85: 109-121
Perrin RM, DeRocher AE, Bar-Peled M, Zeng WQ, Norambuena L, Orellana A, Raikhel NV, Keegstra K (1999) Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis. Science 284: 1976-1979
Perrot-Rechenmann C (2010) Cellular responses to auxin: division versus expansion. Cold Spring Harb Perspect Biol 2: a001446
Perry JA, Wang TL, Welham TJ, Gardner S, Pike JM, Yoshida S, Parniske M (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume Lotus japonicus. Plant Physiol 131: 866-871
Persson S, Wei HR, Milne J, Page GP, Somerville CR (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. Proc Natl Acad Sci U S A 102: 8633-8638
Poggeler S, Kuck U (2006) Highly efficient generation of signal transduction knockout mutants using a fungal strain deficient in the mammalian ku70 ortholog. Gene 378: 1-10
Pogorelko G, Lionetti V, Fursova O, Sundaram RM, Qi M, Whitham SA, Bogdanove AJ, Bellincampi D, Zabotina OA (2013) Arabidopsis and Brachypodium distachyon transgenic plants expressing Aspergillus nidulans acetylesterases have decreased degree of polysaccharide acetylation and increased resistance to pathogens. Plant Physiol 162: 9-23
Porchia AC, Sørensen SO, Scheller HV (2002) Arabinoxylan biosynthesis in wheat. Characterization of arabinosyltransferase activity in Golgi membranes. Plant Physiol 130: 432-441
Prielhofer R, Maurer M, Klein J, Wenger J, Kiziak C, Gasser B, Mattanovich D (2013) Induction without methanol: novel regulated promoters enable high-level expression in Pichia pastoris. Microb Cell Fact 12
Puhlmann J, Bucheli E, Swain MJ, Dunning N, Albersheim P, Darvill AG, Hahn MG (1994) Generation of monoclonal antibodies against plant cell-wall polysaccharides. I. Characterization of a monoclonal antibody to a terminal alpha-(1-->2)-linked fucosyl-containing epitope. Plant Physiol 104: 699-710
Quettier A-L, Eastmond PJ (2009) Storage oil hydrolysis during early seedling growth. Plant Physiol Bioch 47: 485-490
Razem FA, Baron K, Hill RD (2006) Turning on gibberellin and abscisic acid signaling. Curr Opin Plant Biol 9: 454-459
Reiter WD, Chapple C, Somerville CR (1997) Mutants of Arabidopsis thaliana with altered cell wall polysaccharide composition. Plant Journal 12: 335-345
Reyes F, Orellana A (2008) Golgi transporters: opening the gate to cell wall polysaccharide biosynthesis. Curr Opin Plant Biol 11: 244-251
Reyrat JM, Pelicic V, Gicquel B, Rappuoli R (1998) Counterselectable markers: Untapped tools for bacterial genetics and pathogenesis. Infect Immun 66: 4011-4017

Ridley BL, O'Neill MA, Mohnen DA (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry 57: 929-967
Rose JKC, Braam J, Fry SC, Nishitani K (2002) The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: Current perspectives and a new unifying nomenclature. Plant and Cell Physiology 43: 1421-1435
Rossanese OW, Soderholm J, Bevis BJ, Sears IB, O'Connor J, Williamson EK, Glick BS (1999) Golgi structure correlates with transitional endoplasmic reticulum organization in Pichia pastoris and Saccharomyces cerevisiae. The Journal of cell biology 145: 69-81
Rosso MG, Li Y, Strizhov N, Reiss B, Dekker K, Weisshaar B (2003) An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol Biol 53: 247-259
Roubroeks JP, Kondo T (2012) Nano- and microstructures in stretched and non-stretched blend gels of cellulose and hemicelluloses. Holzforschung 66: 993-1000
Rozen S, Skaletsky H (1999) Primer3 on the WWW for general users and for biologist programmers. Bioinformatics methods and protocols. Springer, pp 365-386
Ruprecht C, Mutwil M, Saxe F, Eder M, Nikoloski Z, Persson S (2011) Large-scale co-expression approach to dissect secondary cell wall formation across plant species. Front Plant Sci 2
Saez-Aguayo S, Ralet MC, Berger A, Botran L, Ropartz D, Marion-Poll A, North HM (2013) PECTIN METHYLESTERASE INHIBITOR6 Promotes Arabidopsis Mucilage Release by Limiting Methylesterification of Homogalacturonan in Seed Coat Epidermal Cells. Plant Cell 25: 308-323
Sampedro J, Gianzo C, Iglesias N, Guitian E, Revilla G, Zarra I (2012) AtBGAL10 is the main xyloglucan beta-galactosidase in Arabidopsis, and its absence results in unusual xyloglucan subunits and growth defects. Plant Physiol 158: 1146-1157
Sampedro J, Pardo B, Gianzo C, Guitian E, Revilla G, Zarra I (2010) Lack of alpha-xylosidase activity in Arabidopsis alters xyloglucan composition and results in growth defects. Plant Physiol 154: 1105-1115
Sampedro J, Sieiro C, Revilla G, Gonzalez-Villa T, Zarra I (2001) Cloning and expression pattern of a gene encoding an alpha-xylosidase active against xyloglucan oligosaccharides from Arabidopsis. Plant Physiol 126: 910-920
Samuels AL, Rensing KH, Douglas CJ, Mansfield SD, Dharmawardhana DP, Ellis BE (2002) Cellular machinery of wood production: differentiation of secondary xylem in Pinus contorta var. latifolia. Planta 216: 72-82
Sandhu APS, Randhawa GS, Dhugga KS (2009) Plant Cell Wall Matrix Polysaccharide Biosynthesis. Mol Plant 2: 840-850
Scheible WR, Eshed R, Richmond T, Delmer D, Somerville C (2001) Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in Arabidopsis Ixr1 mutants. Proc Natl Acad Sci U S A 98: 10079-10084
Scheller HV, Ulvskov P (2010) Hemicelluloses. Annu Rev Plant Biol 61: 263-289
Schmalix W, Bandlow W (1993) The Ethanol-Inducible Yat1 Gene from Yeast Encodes a Presumptive Mitochondrial Outer Carnitine Acetyltransferase. Journal of Biological Chemistry 268: 27428-27439

Schultink A, Cheng K, Park YB, Cosgrove DJ, Pauly M (2013) The identification of two arabinosyltransferases from tomato reveals functional equivalency of xyloglucan side chain substituents. Plant Physiol 163: 86-94
Seifert GJ (2004) Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. Curr Opin Plant Biol 7: 277-284
Shewry PR, Napier JA, Tatham AS (1995) Seed storage proteins: structures and biosynthesis. Plant Cell 7: 945
Shrawat AK, Lorz H (2006) Agrobacterium-mediated transformation of cereals: a promising approach crossing barriers. Plant Biotechnol J 4: 575-603
Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7: 539
Simi CK, Abraham TE (2010) Biodegradable biocompatible xyloglucan films for various applications. Colloid Polym Sci 288: 297-306
Simkovic I (2013) Unexplored possibilities of all-polysaccharide composites. Carbohyd Polym 95: 697-715
Simpson DJ, Fincher GB, Huang AH, Cameron-Mills V (2003) Structure and function of cereal and related higher plant (1→ 4)-β-xylan endohydrolases. Journal of Cereal Science 37: 111-127
Sims IM, Gane AM, Dunstan D, Allan GC, Boger DV, Melton LD, Bacic A (1998) Rheological properties of xyloglucans from different plant species. Carbohyd Polym 37: 61-69
Sims IM, Munro SL, Currie G, Craik D, Bacic A (1996) Structural characterisation of xyloglucan secreted by suspension-cultured cells of Nicotiana plumbaginifolia. Carbohydr Res 293: 147-172
Sitaraman J, Bui M, Liu Z (2008) LEUNIG_HOMOLOG and LEUNIG perform partially redundant functions during Arabidopsis embryo and floral development. Plant Physiol 147: 672-681
Smith RC, Fry SC (1991) Endotransglycosylation of xyloglucans in plant cell suspension cultures. Biochem J 279 (Pt 2): 529-535
Sogaard C, Stenbaek A, Bernard S, Hadi M, Driouich A, Scheller HV, Sakuragi Y (2012) GO-PROMTO Illuminates Protein Membrane Topologies of Glycan Biosynthetic Enzymes in the Golgi Apparatus of Living Tissues. PLoS One 7
Somerville C (2006) Cellulose synthesis in higher plants. Annu Rev Cell Dev Biol 22: 53-78
Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredez A, Persson S, Raab T, Vorwerk S, Youngs H (2004) Toward a systems approach to understanding plant-cell walls. Science 306: 2206-2211
Soo EC, Aubry AJ, Logan SM, Guerry P, Kelly JF, Young NM, Thibault P (2004) Selective detection and identification of sugar nucleotides by CE-electrospray-MS and its application to bacterial metabolomics. Anal Chem 76: 619-626
Sorensen I, Pettolino FA, Wilson SM, Doblin MS, Johansen B, Bacic A, Willats WG (2008) Mixed-linkage (1-->3),(1-->4)-beta-D-glucan is not unique to the Poales and is an abundant component of Equisetum arvense cell walls. Plant J 54: 510-521
Sparkes IA, Runions J, Kearns A, Hawes C (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nat Protoc 1: 2019-2025

Sterling JD, Atmodjo MA, Inwood SE, Kumar Kolli VS, Quigley HF, Hahn MG, Mohnen D (2006) Functional identification of an Arabidopsis pectin biosynthetic homogalacturonan galacturonosyltransferase. Proc Natl Acad Sci U S A 103: 5236-5241

Sterling JD, Quigley HF, Orellana A, Mohnen D (2001) The catalytic site of the pectin biosynthetic enzyme α-1, 4-galacturonosyltransferase is located in the lumen of the Golgi. Plant Physiol 127: 360-371

Suzuki M, Kato A, Nagata N, Komeda Y (2002) A xylanase, AtXyn1, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the Arabidopsis thaliana genome. Plant Cell Physiol 43: 759-767

Takeda T, Furuta Y, Awano T, Mizuno K, Mitsuishi Y, Hayashi T (2002) Suppression and acceleration of cell elongation by integration of xyloglucans in pea stem segments. Proc Natl Acad Sci U S A 99: 9055-9060

Tamura K, Shimada T, Kondo M, Nishimura M, Hara-Nishimura I (2005) KATAMARI1/MURUS3 Is a novel golgi membrane protein that is required for endomembrane organization in Arabidopsis. Plant Cell 17: 1764-1776

Tedman-Jones JD, Lei R, Jay F, Fabro G, Li X, Reiter WD, Brearley C, Jones JD (2008) Characterization of Arabidopsis mur3 mutations that result in constitutive activation of defence in petioles, but not leaves. Plant J 56: 691-703

Thole V, Worland B, Wright J, Bevan MW, Vain P (2010) Distribution and characterization of more than 1000 T-DNA tags in the genome of Brachypodium distachyon community standard line Bd21. Plant Biotechnol J 8: 734-747

Thomas DR, Mcneil PH (1976) Effect of Carnitine on Oxidation of Saturated Fatty-Acids by Pea Cotyledon Mitochondria. Planta 132: 61-63

Thomas DR, Wood C (1982) Oxidation of Acetate, Acetyl Coa and Acetylcarnitine by Pea Mitochondria. Planta 154: 145-149

Tian CG, Beeson WT, Iavarone AT, Sun JP, Marletta MA, Cate JHD, Glass NL (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus Neurospora crassa. Proc Natl Acad Sci U S A 106: 22157-22162

Tu L, Banfield DK (2010) Localization of Golgi-resident glycosyltransferases. Cell Mol Life Sci 67: 29-41

Turner SR, Somerville CR (1997) Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. Plant Cell 9: 689-701

Uchida N, Sakamoto T, Kurata T, Tasaka M (2011) Identification of EMS-induced causal mutations in a non-reference Arabidopsis thaliana accession by whole genome sequencing. Plant Cell Physiol 52: 716-722

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W (2010) Lignin Biosynthesis and Structure. Plant Physiol 153: 895-905

Vanzin GF, Madson M, Carpita NC, Raikhel NV, Keegstra K, Reiter WD (2002) The mur2 mutant of Arabidopsis thaliana lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. Proc Natl Acad Sci U S A 99: 3340-3345

Velasquez M, Salter JS, Dorosz JG, Petersen BL, Estevez JM (2012) Recent advances on the posttranslational modifications of EXTs and their roles in plant cell walls. Front Plant Sci 3

Verhertbruggen Y, Marcus SE, Chen J, Knox JP (2013) Cell Wall Pectic Arabinans Influence the Mechanical Properties of Arabidopsis thaliana Inflorescence Stems and Their Response to Mechanical Stress. Plant Cell Physiol 54: 1278-1288
Vierhuis E, York WS, Kolli VS, Vincken J, Schols HA, Van Alebeek GW, Voragen AG (2001) Structural analyses of two arabinose containing oligosaccharides derived from olive fruit xyloglucan: XXSG and XLSG. Carbohydr Res 332: 285-297
Vogel J, Hill T (2008) High-efficiency Agrobacterium-mediated transformation of Brachypodium distachyon inbred line Bd21-3. Plant Cell Rep 27: 471-478
Vogel JP, Raab TK, Somerville CR, Somerville SC (2004) Mutations in PMR5 result in powdery mildew resistance and altered cell wall composition. The Plant Journal 40: 968-978
Voinnet O (2002) RNA silencing: small RNAs as ubiquitous regulators of gene expression. Curr Opin Plant Biol 5: 444-451
Vuttipongchaikij S, Brocklehurst D, Steele-King C, Ashford DA, Gomez LD, McQueen-Mason SJ (2012) Arabidopsis GT34 family contains five xyloglucan alpha-1,6-xylosyltransferases. New Phytol 195: 585-595
Wang Y, Alonso AP, Wilkerson CG, Keegstra K (2012) Deep EST profiling of developing fenugreek endosperm to investigate galactomannan biosynthesis and its regulation. Plant Mol Biol 79: 243-258
Wang Z, Fang B, Chen J, Zhang X, Luo Z, Huang L, Chen X, Li Y (2010) De novo assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweet potato (Ipomoea batatas). BMC Genomics 11: 726
Weigel D, Ahn JH, Blazquez MA, Borevitz JO, Christensen SK, Fankhauser C, Ferrandiz C, Kardailsky I, Malancharuvil EJ, Neff MM, Nguyen JT, Sato S, Wang ZY, Xia Y, Dixon RA, Harrison MJ, Lamb CJ, Yanofsky MF, Chory J (2000) Activation tagging in Arabidopsis. Plant Physiol 122: 1003-1013
Whistler RL (1973) Solubility of Polysaccharides and Their Behavior in Solution. Adv Chem Ser: 242-255
Wiegand G, Remington SJ (1986) Citrate Synthase - Structure, Control, and Mechanism. Annu Rev Biophys Bio 15: 97-117
Winters SJ, Ghooray D, Fujii Y, Moore JP, Jr., Nevitt JR, Kakar SS (2007) Transcriptional regulation of follistatin expression by GnRH in mouse gonadotroph cell lines: evidence for a role for cAMP signaling. Mol Cell Endocrinol 271: 45-54
Wood C, Jalil MNH, Ariffin A, Yong BCS, Thomas DR (1983) Carnitine Short-Chain Acyltransferase in Pea Mitochondria. Planta 158: 175-178
Wu AM, Rihouey C, Seveno M, Hornblad E, Singh SK, Matsunaga T, Ishii T, Lerouge P, Marchant A (2009) The Arabidopsis IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. Plant Journal 57: 718-731
Xiong G, Cheng K, Pauly M (2013) Xylan O-Acetylation Impacts Xylem Development and Enzymatic Recalcitrance as Indicated by the Arabidopsis Mutant tbl29. Mol Plant 6: 1373-1375
Yang J, Jiang W, Yang S (2009) mazF as a counter-selectable marker for unmarked genetic modification of Pichia pastoris. FEMS Yeast Res 9: 600-609

Yaoi K, Mitsuishi Y (2002) Purification, characterization, cloning, and expression of a novel xyloglucan-specific glycosidase, oligoxyloglucan reducing end-specific cellobiohydrolase. J Biol Chem 277: 48276-48281
York WS, Darvill AG, Albersheim P (1984) Inhibition of 2,4-Dichlorophenoxyacetic Acid-Stimulated Elongation of Pea Stem Segments by a Xyloglucan Oligosaccharide. Plant Physiol 75: 295-297
York WS, Darvill AG, Mcneil M, Stevenson TT, Albersheim P (1986) Isolation and Characterization of Plant-Cell Walls and Cell-Wall Components. Methods Enzymol 118: 3-40
York WS, Impallomeni G, Hisamatsu M, Albersheim P, Darvill AG (1995) Eleven newly characterized xyloglucan oligoglycosyl alditols: the specific effects of sidechain structure and location on 1H NMR chemical shifts. Carbohydr Res 267: 79-104
York WS, Kolli VSK, Orlando R, Albersheim P, Darvill AG (1996) The structures of arabinoxyloglucans produced by solanaceous plants. Carbohydr Res 285: 99-128
York WS, Oates JE, van Halbeek H, Darvill AG, Albersheim P, Tiller PR, Dell A (1988) Location of the O-acetyl substituents on a nonasaccharide repeating unit of sycamore extracellular xyloglucan. Carbohydr Res 173: 113-132
Yuan S, Wu YJ, Cosgrove DJ (2001) A fungal endoglucanase with plant cell wall extension activity. Plant Physiol 127: 324-333
Yuan Y, Teng Q, Zhong R, Ye ZH (2013) The Arabidopsis DUF231 domain-containing protein ESK1 mediates 2-O- and 3-O-acetylation of xylosyl residues in xylan. Plant Cell Physiol 54: 1186-1199
Zablackis E, York WS, Pauly M, Hantus S, Reiter WD, Chapple CC, Albersheim P, Darvill A (1996) Substitution of L-fucose by L-galactose in cell walls of Arabidopsis mur1. Science 272: 1808-1810
Zabotina OA (2012) Xyloglucan and its biosynthesis. Front Plant Sci 3: 134
Zabotina OA, van de Ven WTG, Freshour G, Drakakaki G, Cavalier D, Mouille G, Hahn MG, Keegstra K, Raikhel NV (2008) Arabidopsis XXT5 gene encodes a putative alpha-1,6-xylosyltransferase that is involved in xyloglucan biosynthesis. Plant Journal 56: 101-115
Zeng W, Jiang N, Nadella R, Killen TL, Nadella V, Faik A (2010) A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. Plant Physiol 154: 78-97
Zhang YL, Zhang JJ, Hoeflich KP, Ikura M, Qing GL, Inouye M (2003) MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in Escherichia coli. Mol Cell 12: 913-923
Zhao Q, Dixon RA (2011) Transcriptional networks for lignin biosynthesis: more complex than we thought? Trends Plant Sci 16: 227-233
Zhao YJ, Song DL, Sun JY, Li LG (2013) Populus endo-beta-mannanase PtrMAN6 plays a role in coordinating cell wall remodeling with suppression of secondary wall thickening through generation of oligosaccharide signals. Plant Journal 74: 473-485

9 Appendix 1: Gene sequences

CSLC4

Native CDS:

ATGGCTCCCAACTCAGTTGTTGTAACAATTGAGAATCCCACCAGCATCTCAGTAAAGGAGATGATGAAGAAGTCAGATTCAGAATCTCTTTTC
CCGGATAAACAGAAGAAATCTGTTAGTTCTAAGCAATTCACATGGTTTCTTCTCCTAAAAGCTCATAAAGCACTAACCTGTCTTTCATGGTTAG
CAATGGCTTCGAAATCCCTGTGTTCATCAGTCAAGAAAAGGGTTTCATCATCTGATATAAGTGAAGAAGAGGAAGAACATCCAAAAAGCAGA
GGGAAATTGTATAGATTCATCAAAGTATTTCTCTCTATATCTGTTTTAGCTCTTATAGTGGAAATCATTGCGTATTACATGAAATGGAACTTGG
GGAGTATGATTCAGCCATGGGGGATCCAAGGTTTCCTCCATTGGTCTTACATGGCTTGGCTTTCGTTTCGGCTGGATTATATTGCACCTTTGGT
TGGACTTCTATCTCAATTCTGCATTGTGTTGTTCATGATCCAATCTCTAGATCGGTTGATTCTGTGTATTGGATGTTTTTGGATTAAGTACAAGA
ATTTGAAACCAAAGATTGATGAAAATGCTTATGATATTGAAGACAACTCAAGCTTCCCAATGGTACTAGTTCAGATTCCTATGTGCAATGAAA
GAGAGGTGTATGAGCAATCGATTGGTGCTGTTTGTCAACTCGATTGGCCGAAAGATCGAATCCTGATTCAAGTTCTGGATGACTCAGATGATT
CAAGCCTACAACTTTTGATCAAAGAAGAAGTATCCTCATGGCGCCAGAAAGGAGTGAACATAATCTACAGGCATCGTTTAATCAGAACCGGTT
ACAAAGCAGGCAATATGAAATCAGCAATGGGTTGTGATTACGTTAAAGACTACGAATTCGTCGCCATATTCGACGCCGATTTCCAGCCGAATC
CAGATTTCCTTAAGCTTACAATGCCTCATTTCAAGGGTGATCCTGAAGTCGGATTAGTTCAAGCTCGTTGGTCGTTTGTTAACAAAGATGAGAA
CTTGCTGACAAGACTTCAAAACATTAACTTGTGCTTCCATTTCGAAGTCGAGCAACAAGTGAACGGTGTGTTCCTCAATTTCTTCGGGTTCAAC
GGAACCGCAGGGGTGTGGAGGATCAAGGCGCTGGAAGAGTCGGGAGGATGGCTCGAAAGAACCACGGTCGAGGATATGGATATCGCGGT
TCGCGCGCACTTAAACGGGTGGAAGTTTGTTTTCCTCAACGATGTGAAAGTACTTTGTGAATTGCCTGAATCTTATGAAGCTTATAAGAAACA
ACAACATCGTTGGCATTCGGGTCCAATGCAGCTCTTTCGGTTATGTCTTCCATCCGTCTTATCTTCAAAGATCTCAGCATGGAAGAAGGCTAAT
TTGATCTTTCTTTTCTTCCTTTTAAGGAAACTTATACTCCCATTTTATTCATTCACATTGTTTTGTATTATACTCCCTTTAACCATGTTTATACCCGA
AACCGAGTTACCCCTTTGGGTCATCTGTTACATACCCATTTTCATGTCCTTCTTGAACATTCTCCCCTCACCCAAATCTTTCCCTTTTCTAGTCCCT
TACCTACTCTTCGAAAACACAATGTCCGTCACAAAGTTCAACGCCATGATCTCCGGCTTGTTCCAACTCGGAAGCGCCTACGAATGGGTCGTC
ACCAAAAAAAACCGGTCGATCATCCGAGTCCGACCTACTTGCCCTCGCCGAGAAAGAGGAGAAGATACTACGACGAAACTCCGAGTCCGGTTT
GGAAATGTTGAGTAAACTTAAGGAACAAGAAGAAATCGCGCCGCTGAAGAAGGTCACCAAGAAAAAAAAGAACAAGATATTCCGGAAAGA
GCTCGCCCTCGCGTTTCTTCTTCTTACCGCTGCCGGAAGAAGCTTGTTGTCTGCTCAAGGAGTTCATTTTTACTTCTTGTTGTTTCAAGGCTTGT
CTTTCTTAGTTGTAGGTTTGGATTTGATTGGTGAGCAAATGAGTTAG

Codon optimized

ATGGCCCCTAACTCCGTTGTTGTTACTATAGAAAATCCCACCAGTATTTCCGTGAAAGAGATGATGAAAAAGAGTGACTCAGAAAGTCTGTTT
CCAGACAAACAGAAAAAGTCAGTATCTTCTAAACAGTTTACTTGGTTCCTGCTATTGAAAGCCCACAAGGCTTTAACGTGCTTGTCATGGCTG
GCCATGGCTTCAAAGTCTTTATGTTCATCTGTCAAAAAGAGAGTGTCCTCCAGTGACATCTCCGAAGAGGAAGAGGAGCATCCAAAGAGTCG
TGGGAAATTGTATCGATTCATCAAAGTGTTTTTGTCTATTTCTGTGTTAGCACTTATTGTCGAGATTATCGCCTATTACATGAAATGGAATTTGG
GTTCAATGATACAACCATGGGGAATACAAGGCTTTCTGCACTGGTCCTATATGGCTTGGTTGTCTTTCAGACTTGATTACATTGCACCATTGGT
TGGCCTGCTATCTCAGTTCTGTATTGTGTTATTCATGATTCAATCATTGGATCGTTTGATTTTGTGTATTGGATGCTTTTGGATTAAGTACAAAA
ACTTGAAGCCAAAGATAGATGAGAATGCATACGATATCGAAGATAACAGTTCCTTCCCTATGGTTTTAGTCCAAATTCCTATGTGCAATGAGC
GAGAAGTCTACGAACAGTCTATCGGAGCTGTCTGTCAATTAGATTGGCCAAAAGATAGAATCCTTATCCAAGTTTTGGATGACTCAGATGATT
CCTCTTTACAGTTATTGATCAAGGAGGAAGTTTCCTCATGGCGTCAAAAAGGCGTAAACATCATCTATCGTCATAGATTGATTAGGACAGGGT
ATAAGGCAGGGAATATGAAATCTGCAATGGGATGTGATTACGTTAAAGACTATGAGTTCGTGGCTATCTTTGATGCAGACTTCCAACCTAATC
CCGACTTTTTGAAACTGACTATGCCACATTTCAAAGGTGACCCCGAAGTTGGTCTAGTTCAAGCTAGATGGTCCTTTGTCAACAAAGATGAGA
ACTTGCTTACAAGGTTGCAAAACATCAATCTTTGTTTTCACTTTGAAGTCGAGCAACAGGTTAACGGTGTTTTCCTTAACTTTTTCGGTTTCAAT
GGAACCGCCGGTGTTTGGAGAATCAAGGCTCTTGAGGAATCCGGAGGCTGGCTAGAGAGAACAACTGTCGAGGACATGGATATTGCAGTTA
GAGCCCACTTGAATGGGTGGAAGTTTGTCTTTTTGAATGACGTAAAGGTTCTATGCGAACTGCCAGAATCTTATGAAGCATACAAAAAGCAAC
AACATAGATGGCATAGTGGTCCAATGCAGTTGTTTCGTCTTTGTCTGCCATCTGTGCTATCTTCCAAGATATCCGCTTGGAAAAAGGCTAACCT
GATCTTTTTGTTTTTCCTTTTGAGGAAATTGATTTTGCCCTTTTACTCTTTCACTCTTTTCTGCATCATCCTACCTTTGACAATGTTCATACCAGAA
ACTGAATTACCTCTTTGGGTGATATGTTACATTCCCATATTCATGTCATTCCTTAACATTCTGCCTAGTCCAAAATCCTTTCCTTTTCTGGTGCCTT
ATCTATTGTTCGAAAACACCATGAGTGTCACTAAGTTTAACGCCATGATTTCCGGTTTGTTTCAGTTGGGTTCAGCTTATGAGTGGGTTGTAAC
GAAAAAGACCGGTAGATCATCTGAATCTGACTTACTTGCCCTGGCTGAAAAGGAAGAGAAGATATTGAGAAGGAATTCAGAATCAGGACTG
GAAATGCTGTCAAAGCTTAAGGAACAGGAGGAGATTGCACCTCTGAAAAAGGTAACCAAGAAGAAGAAAAACAAGATTTTCCGAAAGGAAT
TAGCTCTTGCCTTTTTGCTACTTACGGCTGCTGGTAGATCTCTACTTTCTGCTCAGGGTGTCCACTTCTACTTTTTGCTGTTCCAAGGATTAAGTT
TTCTAGTCGTAGGATTAGACCTAATTGGAGAACAAATGTCTTAG

XXT2

Native CDS:

ATGATAGAGAGGTGTTTAGGAGCTCAAAGATCCCGGAAGATCCAGAGAGCTTTGCGTCACTGTAAGGTCACCGCTCTTTGTCTTGTTCTCACC
GTCGTCGTTCTACGTGGTACTATCGGTGCCGGTAAATTCGGTACTCCAGAGCAGGACTTCGTCGACATTCGGGACCATTTTTATTCGAGGAAG

CGATCCGAGCCGCATCGTGTTCTTGAGGAGGTTCAAACAACCACGTCTTCATCGTCTTCCTCTGATTCCTCCAAGTCTTCAGGTAACAACAATA
ATTACGAGCAATTCGATGTTAATTCGATTTTTGTTCACGAAGGTGAGGATGATAAGCCTGATCCGAGTAAACCTTACTCACTAGGACCTAAGA
TCTCAGATTGGGATGAACAGAGATCTCAATGGCTTAAGCAAAACCCTGATTTCCCTAATTTCGTTCAACCCTCTAAGCCACGGGTTTTATTGGT
TACTGGGTCTTCCCCGAAACCGTGTGAGAACCCGGTTGGGGATCATTACTTGTTGAAATCTATCAAGAATAAGATTGATTACTGTAGGGTACA
CGGTATCGAGATTTTCTATAACATGGCGCTTCTCGATGCGGAAATGGCTGGGTTTTGGGCTAAGCTCCCTTTGATTCGGAAGCTTTTATTGTCT
CACCCTGAAGTGGAGTTCCTTTGGTGGATGGACAGTGACGCTATGTTCACCGATATGGCGTTTGAGCTTCCATGGGAGAGGTATAAGGATGT
TAACTTAGTGATGCACGGTTGGAATGAGATGGTGTATGATGAAAAGAATTGGATTGGGTTGAACACTGGAAGTTTCTTGTTGAGGAACACAC
AGTGGGCTTTGGATCTTCTTGATGTATGGGCACCAATGGGTCCTAAAGGAAAGGTTAGGGATGATGCAGGGAAAATACTTACTCGTGAGCTT
AAAAACAGGCCGGTTTTCGAAGCTGACGATCAGTCAGCAATGGTTTACTTACTAGCAACACAGAGAGAGCAATGGGGGAACAAAGTTTATCT
CGAAAACGCGTATTACCTGCACGGATACTGGGGGATTTTGGTTGATAGATATGAAGAGATGATCGAGAATTATCATCCTGGTTTCGGTGATC
ATAGATGGCCATTGGTGACCCATTTCGTGGGTTGTAAACCTTGTGGAAAATTTGGTGATTACCCAGTTGAGAGATGTTTAAGGCAAATGGAC
AGAGCATTCAACTTTGGGGATAACCAGATACTTCAAATATACGGTTTCACTCATAAAACATTGGCTAGTAAACGAGTCAAAAGAGTTCGAAAC
GAAACAAGCATGCCTCTTGAGGTTAAAGATGAACTTGGGTTGATTCATCCACCCTTTAAGGCTGTTAAGGTAGCCTCTGGTTAA

Codon optimized:

ATGATTGAACGATGCCTTGGCGCTCAGCGATCTAGAAAGATCCAGCGTGCTTTGAGACATTGCAAAGTTACGGCACTGTGTCTTGTCTTAACT
GTGGTTGTTCTGCGTGGTACTATAGGTGCAGGGAAGTTTGGTACTCCAGAGCAAGATTTTGTAGATATTCGAGATCATTTCTACAGTAGGAAG
AGAAGTGAACCACATAGAGTATTGGAAGAGGTGCAAACTACAACCTCCAGTTCCTCATCTTCTGACTCTTCAAAATCATCAGGGAACAATAAC
AACTACGAGCAATTTGATGTCAATTCCATATTCGTACATGAAGGCGAAGATGATAAGCCTGACCCTTCAAAGCCTTACAGTTTGGGACCAAAG
ATTTCCGATTGGGACGAGCAGAGGTCACAATGGCTTAAACAAAATCCAGATTTTCCAAACTTTGTCCAGCCTTCTAAACCCAGAGTTCTGTTAG
TCACCGGTTCATCCCCTAAGCCTTGCGAGAATCCAGTGGGAGATCATTACCTTTTGAAATCCATCAAAAACAAGATTGACTATTGTCGTGTGCA
TGGGATCGAGATTTTCTACAATATGGCACTACTGGATGCTGAAATGGCCGGATTCTGGGCAAAGTTGCCACTTATTCGAAAACTATTGTTATC
TCACCCTGAAGTTGAGTTTTTGTGGTGGATGGATTCTGACGCTATGTTTACAGACATGGCTTTTGAACTACCTTGGGAAAGGTACAAGGATGT
TAATCTGGTCATGCACGGTTGGAACGAAATGGTGTATGACGAGAAAAACTGGATTGGATTGAATACTGGAAGTTTCCTTTTGAGAAACACTC
AATGGGCATTAGACTTGTTAGATGTTTGGGCTCCAATGGGACCCAAAGGAAAGGTTAGGGACGATGCTGGCAAGATTCTAACCAGAGAATT
GAAAAACAGGCCAGTCTTTGAGGCCGATGATCAATCTGCTATGGTCTATTTGTTGGCTACCCAACGTGAGCAATGGGGTAACAAAGTTTACTT
GGAGAATGCCTACTATCTACACGGTTATTGGGGTATCTTAGTAGACAGATATGAAGAGATGATTGAAAACTACCACCCTGGCTTTGGAGATC
ACAGATGGCCCTTGGTTACTCACTTCGTGGGATGTAAACCCTGTGGAAAGTTCGGTGACTATCCCGTCGAAAGATGTCTTAGACAGATGGAC
AGAGCCTTCAATTTCGGAGACAATCAGATTTTGCAAATCTATGGTTTTACGCATAAGACATTAGCTTCCAAAAGAGTTAAGCGTGTTAGAAAC
GAAACATCTATGCCATTGGAGGTTAAAGACGAACTGGGTCTGATACACCCACCTTTCAAGGCCGTAAAGGTTGCTAGTGGTTGA

UGD

Native CDS:

ATGGTGAAGATCTGCTGTATTGGGGCTGGATATGTTGGTGGCCCCACCATGGCTGTGATTGCACTGAAGTGCCCAAATATTGAAGTTGCTGTC
GTCGATATTTCTGTACCAAGAATCAACGCCTGGAACAGCGAACAACTCCCAATTTATGAACCGGGTCTGGATGATGTGGTGAAGGAGTGCAG
AGGAAGAAACCTTTTCTTCAGCACTGATGTGGAGAAACATGTTTCTGAGGCTGATATCGTCTTTGTTTCCGTCAACACACCAACCAAAACCCAA
GGTCTCGGAGCTGGAAAAGCAGCTGATCTGACCTACTGGGAGAGCGCAGCCCGTATGATTGCTGATGTATCCAAATCATCAAAAATTGTGGT
TGAGAAATCAACAGTGCCTGTGAAAACAGCTGAGGCAATTGAAAAGATCTTGACACACAACAGCAAAGGCATTAGCTTTCAGATTCTTTCAA
ACCCGGAGTTTCTTGCTGAGGGAACCGCCATTGAAGACCTTTTTGGTCCGGACCGTGTTCTCATCGGAGGTAGGGAAACCCCAGAAGGTGCT
AAGGCTATTAAAGCCTTGAAAGATGTTTATGCCCATTGGGTACCCGAGGATCGCATTATCTGCACCAATCTCTGGTCTGCCGAGCTATCGAAG
CTCGCGGCCAACGCTTTTTTGGCGCAGAGGATCTCTTCTGTGAATGCCATGTCAGCTTTATGTGAGGCAACTGGTGCAAATGTGTCCGAAGTA
TCTCATGCGATCGGTAAGGACACCAGAATCGGTGCCAAGTTCTTGAACGCCAGCGTGGGTTTCGGTGGGTCCTGTTTCCAGAAAGATATCTTG
AACTTGGTCTACATCTGCGAGTGTAATGGCCTCCCAGAGGTTGCAAACTACTGGAAGCAGGTCATTAAGGTTAATGACTACCAGAAGTTTCGA
TTCGTAAACCGTGTCGTTTCTTCCATGTTTAACACAGTGTCCGCTAAGAAAGTGGCCATTCTAGGTTTCGCTTTCAAGAAGGACACGGGTGAC
ACTAGAGAGACCCCAGCCATTGATGTCTGCAATGGACTGTTAGCAGATAAGGCCCAGCTGAGCATTTACGATCCCCAAGTCACTGAAGATCA
GATCCGAAGGGATCTCGCTATGAAGAAATTCGACTGGGATCATCCCGCCCATCTTCAACCGATGAGCCCTACCGACACCAAGCAGCAAGTCA
CCGTAGTTTCTGGCGCTTACGAGGCGACAAAGGATGCACACGGTATATGTTTCCTGACTGAGTGGGACGAATTCAAAACACTCGACTTCCAG
AAGATTTACGATAACATGCAGAAACCGGCATTCGTGTTCGATGGTCGCAACATCGTGGATCTCGAGAAGCTGAGGCAAATCGGGTTTATCGT
TTACTCGATCGGTAAGCCGTTGGATTCATGGCTCAAGGACATGCCTGCTGTGGCTTAA

Codon optimized:

ATGGTAAAGATTTGTTGTATAGGTGCCGGTTACGTTGGAGGGCCTACTATGGCCGTTATTGCATTGAAATGTCCTAACATTGAAGTCGCAGTG
GTGGACATCTCTGTCCCAAGGATTAACGCCTGGAATTCAGAACAACTGCCAATCTACGAGCCAGGTTTAGATGATGTGGTCAAGGAATGCAG
AGGTAGAAACCTGTTTTTCTCCACCGACGTTGAGAAACATGTGTCAGAGGCTGATATTGTCTTTGTTTCTGTAAACACACCTACAAAGACACAA
GGACTTGGTGCTGGAAAAGCTGCAGATCTGACCTATTGGGAATCCGCAGCTCGAATGATTGCAGATGTCAGTAAGTCCTCCAAAATCGTGGT

GGAAAAGTCTACTGTCCCTGTGAAAACTGCCGAAGCTATCGAGAAAATCTTAACACACAATAGTAAAGGTATCAGTTTCCAGATCCTATCTAA
CCCAGAGTTTTTGGCCGAGGGGACAGCTATAGAGGATTTGTTTGGTCCAGATAGGGTTCTAATTGGGGGTAGAGAGACTCCAGAAGGAGCT
AAGGCAATCAAGGCTCTAAAGGATGTGTACGCACATTGGGTTCCAGAGGACAGAATCATTTGCACGAACTTATGGTCCGCCGAGCTTTCTAA
GCTGGCCGCAAATGCCTTTCTGGCCCAGAGAATCTCTTCTGTGAACGCTATGTCTGCATTATGCGAAGCTACTGGTGCAAATGTGTCCGAGGT
TTCTCATGCTATTGGTAAGGACACTCGTATCGGAGCTAAGTTTTTGAACGCCTCAGTTGGATTCGGCGGTTCATGTTTTCAGAAAGACATCCTG
AATCTGGTCTACATTTGTGAATGCAATGGCTTGCCTGAGGTTGCTAACTATTGGAAGCAAGTAATCAAAGTTAATGACTACCAAAAGTTTAGA
TTCGTGAATAGGGTAGTTAGTTCAATGTTTAACACGGTTTCCGCCAAAAAGGTTGCCATATTGGGATTTGCATTCAAAAAGGACACTGGAGAC
ACAAGAGAAACCCCTGCAATTGACGTTTGTAACGGATTACTAGCTGACAAAGCTCAGCTTTCTATCTACGACCCTCAGGTCACGGAAGATCAG
ATTAGGAGAGATCTAGCTATGAAAAAGTTTGATTGGGATCACCCCGCTCACCTTCAACCTATGTCTCCCACCGACACTAAGCAGCAAGTTACT
GTTGTTTCAGGTGCTTACGAAGCTACGAAAGATGCACATGGCATTTGCTTTCTGACTGAATGGGACGAGTTCAAAACATTGGACTTTCAAAAG
ATATACGACAATATGCAAAAGCCAGCTTTCGTATTTGACGGCAGAAACATTGTTGATCTTGAAAAGCTGAGACAAATTGGTTTCATAGTTTAC
TCTATTGGAAAACCATTGGACAGTTGGTTGAAAGATATGCCAGCCGTCGCTTGA

UXS

Native CDS:

ATGGCGTCGGAGCTTATATACAGAGGGCATGAAACCGACCAACCTGTATCAGATTGGTACTCACCGAAACGTATCAAGCCATGTTTTGCGGT
GACTCAATCGATTCGTTACTTGCTCCGGGAACAGAGGCTAATCTTCGTGATAGTGGGGATGGTGTTAGCGACTTTGGTCTTCGCGTTTTTCCCT
CATTCGTCTAGGTCAACAAACCTTCCCGATTACGTCGACCCGATATCCGGGTCGGACTCTTACTACCCCATCCAAACACAGAAGAAGACCAACT
TGGAGTACTTGAACCGAGTCGGATTTGGATCGGCTGGTGGTAAGGTTCCGCTAGGTTTGAAACGGAAAGGATTAAGAATCGTGGTGACCGG
TGGGGCCGGATTCGTCGGGTCCCATCTCGTGGATCGTTTGATTGCGAGAGGTGATAGCGTGATCGTCGTGGATAATTTCTTCACTGGGAATA
AAGAGAATGTGATGCATCATTTCGGGAATCCGAATTTCGAGCTCATCAGACATGACGTTGTTGAACCTTTGCTCCTTGAAGTCGATCAGATCT
ATCATCTCGCTTGCCCTGCATCGCCTGTTCATTACAAATTCAACCCCGTCAAGACTATCAAAACGAATGTGGTTGGAACACTCAACATGCTTGG
GTTAGCAAAGCGAGTTGGAGCGAGATTTTTGCTGACGAGCACCAGTGAAGTATACGGAGATCCTCTGCAACATCCTCAGGTCGAAACATATT
GGGGAAACGTGAATCCCATCGGTGTAAGAAGCTGTTATGATGAGGGAAAACGTACTGCTGAAACTTTGACCATGGACTACCATAGAGGAGC
TGGTGTAGAGGTACGAATTGCTAGGATTTTCAATACTTATGGACCTCGAATGTGCATTGATGACGGTCGTGTTGTTAGCAATTTCGTTGCTCA
GGCATTAAGGAAAGAGCCTATGACTGTTTATGGTGACGGGAAGCAGACAAGGAGTTTCCAATTTGTTTCCGACCTGGTTGAGGGATTGATGC
GCCTAATGGAAGGAGAGCATGTAGGACCTTTCAATCTTGGAAATCCCGGGGAATTCACCATGCTTGAGCTTGCTAAGGTGGTACAAGAAACT
ATTGACTCAAATGCAAAGATAGAGTTCAGGCCGAACACAGAAGATGATCCTCACAAGAGGAAGCCGGATATCACAAGGGCTAAAGACCTTCT
TGGCTGGGAACCAAAGGTATCCCTTCAAAAGGGTCTACCTCTTATGGTGGCGGACTTCCGACACCGTATTTTCGGTGATCACAAAGAGGGTA
GCACCACCACCACCACCACCACCACCACAACAACTAAGAACTCGTCTTAA

Codon optimized:

ATGGCTTCCGAATTGATATACAGAGGCCATGAGACTGATCAACCTGTTTCCGATTGGTATTCTCCCAAGAGGATCAAGCCTTGTTTTGCTGTTA
CACAATCAATTAGATATCTGTTAAGAGAGCAAAGATTGATATTCGTTATAGTCGGTATGGTGCTGGCTACTTTGGTATTTGCCTTTTTCCCACA
CTCCTCTAGATCCACGAATCTTCCAGACTATGTTGATCCTATTAGTGGTTCTGATTCCTATTACCCTATACAAACTCAGAAAAAGACTAATCTGG
AATATCTGAACAGGGTGGGATTCGGCTCTGCAGGGGGCAAAGTGCCATTAGGATTGAAACGAAAGGGGTTGCGTATTGTCGTCACGGGTGG
GGCTGGTTTTGTTGGATCTCATCTTGTTGACCGTTTGATTGCCAGAGGTGATTCTGTAATCGTAGTAGACAACTTTTTCACAGGTAACAAGGAA
AACGTAATGCATCACTTTGGTAATCCTAACTTTGAATTGATAAGACACGATGTAGTGGAACCACTATTGCTTGAAGTAGATCAAATCTACCACC
TTGCATGTCCAGCTTCTCCTGTTCACTACAAATTCAATCCTGTTAAGACCATTAAGACCAATGTCGTGGGAACCTTAAACATGTTAGGATTGGC
CAAACGTGTTGGTGCCAGATTCCTGCTAACCAGTACTTCAGAAGTTTATGGTGACCCATTGCAACACCCACAGGTTGAGACTTATTGGGGAAA
TGTTAATCCCATTGGTGTTCGTAGTTGCTACGATGAAGGCAAAAGGACAGCTGAAACCTTAACAATGGATTATCACCGAGGGGCTGGAGTGG
AGGTTAGAATTGCCAGGATCTTCAATACTTATGGCCCCAGAATGTGTATTGATGATGGCCGAGTTGTTTCTAACTTTGTAGCTCAGGCATTACG
TAAGGAGCCAATGACCGTCTATGGTGATGGTAAGCAAACTAGAAGTTTCCAATTCGTATCAGATTTGGTTGAAGGATTGATGCGACTTATGG
AAGGTGAACATGTCGGGCCTTTCAACTTGGGTAATCCCGGAGAATTCACCATGCTTGAACTAGCTAAAGTCGTGCAGGAAACTATAGACTCA
AATGCTAAGATCGAATTTCGTCCTAACACAGAGGATGATCCACATAAGCGTAAGCCAGATATCACCAGAGCCAAAGATCTACTTGGCTGGGA
GCCCAAAGTCTCACTGCAAAAAGGTCTGCCTCTTATGGTTGCCGACTTCAGACATAGGATTTTCGGTGACCATAAGGAGGGATCAACCACTAC
GACCACTACAACTACTACAACCACGAAAAACTCATCATGA

TmNST1

Native CDS:

ATGACCGAGGGCCAAACTTTTCAGCTGGGGACTGTAGGAGCTTTGAGTTTATCTGTGGTATCATCTGTGTCAATTGTGATATGTAACAAGGCC
CTTATTAGCACGCTTGGATTTACCTTTGCCACAACTTTGACTAGCTGGCATCTTTTGGTAACATTTTGCTCCCTTCATGCGGCACTATGGATGAA
TTTGTTCGAACACAAACCTTTTGATGCAAGAGCGGTGATGGGATTTGGTGTATTAAATGGCATTTCTATTGGGCTTTTAAATCTTAGTTTGGGT
TTCAATTCTGTTGGTTTTTATCAGATGACAAAACTGGCTATAATTCCATGCACTGTCCTTTTGGAGACCCTTTTCTTCAGGAAAAAGTTCAGTCA

AAGCATCCAGTTTTCTCTAGTCATCCTACTTGTGGGTGTGGGCATTGCAACAGTGACCGATCTTCAACTCAATGTCTTGGGTTCTGTCTTGTCTC
TGCTGGCAGTTATTACTACATGCGTCGCTCAAATTATGACGAACAACATTCAGAAGAAATTCAAAGTTTCTTCAACCCAACTTCTATATCAGTC
TTGTCCTTATCAAGCATTAACTTTGTTCGTTGTGGGGCCATTTCTAGATGGTCTTTTGACTAACCTAAATGTGTTTGCTTTCAAGTACACCCCTC
AAGTCTTGTTCTTCATTGTTCTGTCCTGTCTGATATCCGTGTCTGTAAACTTCAGTACGTTTCTCGTCATTGGAAAGACGTCTCCAGTTACATATC
AGGTTCTAGGACATCTTAAAACATGTCTGGTTTTGGCCTTTGGCTATACTCTACTTCATGACCCATTTAGCTGGCGCAACATTCTGGGTATACTT
GTCGCTGTAGTTGGAATGATACTTTATTCTTATTGCTGCACGATTGAGAGCCAAAAGGCTAGTGAAGCATCAACACAAACGCCCCAGGAAAA
GGAAATCGAATCGGAGCCCTTAGTCAGTGTGGAAACTGGAATCGGAAACCTAGCAGATGGTGTTGCTCAAACTTCTCCTGTGTGGAATTCAA
ACAAGGATCTGCATGCATGA

TmNST2

Native CDS:

ATGGATTCGGAGAAGAAACATTCGACGGTATCTAATTTGGGAGCATGGGGCATGAATGTTGTTAGCTCCGTCGGTATAATTATGGCTAATAA
GCAGCTCATGTCTGCCTCCGGCTACGGTTTTAGCTTTGCTACGACGCTCACGGGATTCCACTTTGGCCTGACGGCTCTGGTTGGTTTGATATCT
AAAGCCACAGGCCTATCGGCATCCAAGCAAGTTCCGCTATGGGAGCTTGTTTGGTTCTCTATTGTTGCTAATATCTCCATCGCTGCCATGAACT
TCAGTCTCATGCTTAATTCCGTCGGCTTTTACCAGATTTCGAAATTGAGCATGATTCCAGTTGTATGCGTGATGGAATGGATCATTCATCACAA
ACACTACTCCAAAGAAGTGAAAATGTCTGTATTGGTTGTGGTGATAGGTGTCGGTATATGCACCGTGACAGATGTTCATGTCAACGCCAAAG
GTTTCATTTGTGCCTGCACAGCTATCATGTCTACATCTTTGCAGCAAATCTCAATAGGATCTTTGCAGAAAAAATACTCAATTGGATCTTTTGAA
TTATTGAGCAAGACAGCCCCAATCCAAGCCCTCTCTCTTCTTGTTCTTGGTCCCTTTGTGGATTACTCTCTCAGTGGAAAACTCATTACAGAATA
CAAATCGTCCTATGGTGCCATTCTGTGCATTCTTGTTTCATGCACCTTAGCAGTATTCTGCAACGTGAGCCAGTACCTCTGCATCGGGCGATTTT
CCGCAACTTCTTTCCAGGTTTTAGGTCACATGAAAACCGTATGTGTTCTCACATTGGGGTGGCTTCTGTTCGACTCGGAGCTTACTTTAAAGAA
CATCGCCGGTATGATTCTTGCTGTGGTTGGTATGGTGATCTATAGTTGGGCTGTTGAAATCGAAAAACAATCAAAGGCTATACCTATGAAAAA
CAGCCTTACTGAAGAAGAGATTAGGCTTTTGAAGGAAGGAGTAGAGAATAACTCATTGAAGAACATTCAGCTCGGAGAAAATAAAGCTTAG

TmNST3

Native CDS:

ATGGGTAAGGGTGGATCAATGAGCGAGAATGTAATCAAGAAGATCATCCTCTCCTACACCTATGTCGCCATCTGGATCTTCCTCAGTTTCACC
GTCATCGTCTACAACAAATACATCCTCGACAAGAAGATGTACAATTGGCCATTCCCCGTTTCACTAACCATGATCCACATGTCCTTTTGCGCCA
CTCTCGCTTTCCTCCTCATCAAGGTCTTCAGATTCGTAGAGCCTGTTACCATGTCTCGCGAGATGTACTTCAAATCCGTTGTTCCGATTGGCGCC
TTGTACTCCCTCTCCCTATGGCTTTCCAATTCCGCTTACATTTACCTCTCCGTGTCTTTTATTCAGATGCTTAAAGCTCTCATGCCTGTCGCCGTTT
ACTCGATCGGGGTTTTGTTTAAGAAAGAAGGGTTCAAATCTGACACCATGGCCAACATGCTCTCGATCTCGTTCGGTGTTGCCATCGCGGCGT
ACGGGGAGGCTAGATTCGATACGTGGGGTGTTATGTTGCAACTCGGTGCTGTTGCTTTCGAGGCTACACGTTTGGTAATGATCCAAATCTTGC
TTACATCTAAGGGTATAACTTTGAACCCTATCACCTCATTGTATTATGTTGCTCCATGTTGCTTTGTTTTCTTGTTTGTTCCATGGATTATCGTGG
AGTATCCGATTTTAAGGGATACATCGAGTTTCCATTTCGATTTTGTCATCTTCGGCACCAATTCATTGTGTGCTTTCGCTTTGAACCTTGCCGTG
TTCTTGCTTGTGGGTAAGACTTCTGCTTTGACCATGAATGTGGCCGGTGTGGTCAAGGATTGGTTGTTGATTGCTTTCTCCTGGTCGGTGATAA
AGGACACCGTTACACCTATTAACCTGGTTGGGTATGGTGTTGCGTTTATGGGTGTTGCTTATTACAATCACTCCAAATTGCAAGCCTTAAAGGC
TAAGGAGGCACAAAAGAAAACTCAACAAGCTGATGAGGAAGCCGGAAGGCTTCTGGAGGATAGGGAAGGAGGAGAAGGCACTACTAAGA
GAAATGAATCTCAGAATTAA

TmNST4

Native CDS:

ATGGGAAAGGGCGGCTCAATGAGCGAGAGTGTGATCAAAAAGATCCTCCTCTCCTACACCTATGTCGCCATCTGGATCTTCCTCAGTTTCACG
GTCATCGTTTACAACAAGTACATCCTCGACAAAAAGATGTATAACTGGCCTTTCCCCGTTTCCCTAACCATGATCCACATGTCCTTCTGTGCAAC
TCTTGCTTTCCTCCTCATCAAGGTTTTCAAGCTGGTTGAGCCTGTTACTATGTCTCGCGAGATGTACTTGAAATCCGTTGTCCCCATCGGCGCTT
TGTACTCTCTCTCCCTCTGGCTTTCTAATTCCGCTTACATTTACCTCTCGGTTTCCTTTATTCAGATGCTCAAGGCCTTGATGCCTGTAGCTGTTTA
CTCCATCGGCGTTCTTTTCAAGAAAGAAGGGTTCAAGAGTGATACGATGATGAACATGCTTTCGATCTCGTTCGGTGTAGCTATTGCGGCTTA
TGGAGAGGCCAGATTCGATACATGGGGAGTGATGTTGCAATTAGGCGCCGTAGCTTTTGAGGCTACTCGTTTGGTAATGATTCAAATCTTACT
TACATCTAAGGGTATTACATTGAATCCAATTACTTCCTTATATTACGTGGCTCCCTGTTGCTTCGTTTTCTTGTTTTTTCCATGGATTTTCGTTGA
ATTCCCGATATTGAGAGACACATCGAGTTTCCATTTCGATTTCGTCATCTTTGGTACCAATTCCTTGTGTGCTTTCGCATTGAATCTAGCTGTAT
TCTTGCTTGTCGGCAAGACTTCTGCATTGACCATGAATGTGGCCGGTGTGGTCAAGGATTGGTTGCTGATTGCCTTCTCGTGGTCGGTCATTA
AGGATACCGTCACACCGATTAACCTGGTGGGGTACGGGCTCGCGTTCTTGGGCGTAGGGTATTACAATCATTCCAAACTCCAGGCATTAAAG
GCTAAGGAGGCACAGAAGAAAACTCAACAGACTGACGAGGAAGCGGGAAGGCTTCTGGAGGAGAGGGAAGGAGAAAGCAGTGCAAAGA
GAAATGATTCTCAAAACTAA

TmXXT5

Native CDS:

ATGGGTCAGGAGATTTTCACGCCTCAAAAGCGAGCTTCCGGTGGCGGAGGCGGTCTACCGACGACCACCGCAAATGTTAATCGAGGTCGAG
CGGCGCCACGTGGGAGGCAGATCCATAAGACATTTAACAACATTAAGATTACGCTTCTCTGCGGTTTCGTTACCATCCTCGTCCTCCGTGGGA
CAATCGGCGTCGGTAATCTCGGTGGATCTGATGACACCGTCAACCAGAATATAATCGAGGAGACTAATCGGATCCTTGCGGAGATTAGATCC
GATTCCGAACCCGGTGATCCCGAAGGAAACGCTGAAGTTTTTATGAACCCTAACGAGACCTACGCACTTGGTCCGAAGATAACGAATTGGGA
TGACGAAAGGAAGGTATGGCTTACTCAAAACCCTGAATTTCCTAGTCTTGTTCATGGTAAAGCTAGGATTCTCCTTGTAACTGGATCGCCTCCA
AAACCCTGTGATAACGCGATTGGGGATCATTACTTGTTGAAAGCGATTAAGAACAAGATTGATTACTGTAGACTTCATGGGATAGAGATAGT
GTATAATTTGGCTCATTTGGATAAAGAACTTGCTGGATACTGGGCAAAACTTCCATTGATTAGGAAGTTAATGTTATCTCATCCTGAAGTAGA
GTGGATTTGGTGGATGGATAGTGATGCTTTATTTACAGACATGGTTTTCGAAATCCCCTTATCTAAATACGAAAAGTCGAACCTAGTCATTCAT
GGGTATCCAGATTTGTTGTATGATCAGAAATCATGGATTGCTTTGAACACAGGAAGTTTTCTGTTCCGGAATTGTCAGTGGTCGTTGGATTTG
TTGGACGTTTGGGCACCTATGGGGCCTAAAGGTCCTATCCGAGAGGAGGCTGGAAAAGTTCTAACtGGGTTTTTGAAAGGCCGACCAGCTTT
CGAAGCCGACGATCAATCCGCTTTGATTTACTTGCTTCTTTCTCAGAAAGATAAATGGATGGACAAGGTGTACGTAGAGAATCAGTACTATTT
GCACGGtTTCTGGGAAGGATTGGTCGATAGGTACGAGGAAATGATGGAGAAGTATCATCCTGGATTAGGTGACGAAAGGTGGCCTTTCGTT
ACCCATTTCGTCGGCTGCAAACCTTGCGGGAGCTACGCGGATTACGCGGTCGAGCGATGCTTGAAGAGCATGGAAAGAGCTTTCAATTTCGC
AGATAATCAAGTTCTTAAACTGTATGGATTCGGGCATAGAGGGTTGTTGAGTCCCAAGATTAAGAGGATCAGGAATGAGACGAGTACACCTT
TGGAGTTTGTAGAAAGGTTTGATATTCGCAGATCAGCTGGACACGTTGGAAAAGTTGGATCATAG

SLC35D1

Native CDS:

ATGGCGGAAGTTCATAGACGTCAGCATGCTCGGGTTAAAGGAGAAGCCCCCGCGAAATCCTCCACACTCCGAGATGAGGAGGAGCTGGGGA
TGGCGTCGGCCGAAACGCTGACCGTGTTTCTGAAGCTGCTGGCCGCCGGCTTTTACGGCGTGAGCTCCTTCCTGATCGTGGTGGTGAATAAG
AGCGTGCTCACCAATTACAGATTTCCCTCCTCACTATGTGTTGGACTTGGCCAGATGGTGGCCACAGTGGCAGTTCTCTGGGTGGGAAAGGC
GCTCAGAGTAGTCAAGTTTCCTGACCTTGACAGAAATGTACCTCGAAAGACGTTTCCACTACCTCTACTATATTTTGGGAACCAAATCACGGG
ACTGTTCAGCACAAAGAAACTGAACTTGCCAATGTTTACAGTTCTGAGAAGGTTCTCCATCCTGTTTACAATGTTTGCTGAAGGAGTTTTACTC
AAGAAGACTTTTTCTTGGGGTATTAAAATGACTGTATTTGCAATGATTATTGGAGCCTTTGTAGCTGCCAGCTCTGACTTGGCATTTGATCTGG
AAGGATATGCTTTTATTCTGATAAACGATGTCCTAACAGCAGCAAATGGTGCATACGTAAAACAAAAATTAGATTCAAAAGAGCTGGGAAAA
TATGGACTGCTCTATTACAATGCACTGTTCATGATTCTGCCCACCCTGGCCATTGCGTATTTCACAGGAGATGCACAAAAGGCTGTGGAGTTTG
AAGGCTGGGCTGACACCCTCTTTCTTCTGCAGTTCACCCTCTCCTGTGTGATGGGGTTTATCTTAATGTACGCCACAGTACTCTGCACGCAGTA
TAATTCTGCTCTTACAACTACAATAGTTGGCTGTATTAAGAATATATTAATAACTTATATTGGAATGGTCTTTGGTGGAGATTATATTTTCACGT
GGACAAACTTCATTGGTTTAAATATCAGCATTGCTGGGAGCCTGGTATATTCCTATATCACTTTCACTGAAGAGCAGCTGAGCAAACAGTCAG
AGGCTAATAACAAGCTGGACATTAAGGGGAAAGGAGCAGTGTGA

SLC35B4

Native CDS:

ATGCGCCCGGCCTTGGCGGTGGGCCTGGTGTTCGCAGGCTGCTGCAGTAACGTGATCTTCCTAGAGCTCCTGGCCCGGAAGCATCCAGGATG
TGGGAACATTGTGACATTTGCACAATTTTTATTTATTGCTGTGGAAGGCTTCCTCTTTGAAGCTGATTTGGGAAGGAAGCCACCAGCTATCCCA
ATAAGGTACTATGCCATAATGGTGACCATGTTCTTCACCGTGAGCGTGGTGAACAACTATGCCCTGAATCTCAACATTGCCATGCCCCTGCATA
TGATATTTAGATCCGGTTCTCTAATTGCCAACATGATTCTAGGAATTATCATTTTGAAGAAAAGATACAGTATATTCAAATATACCTCCATTGCC
CTGGTGTCTGTGGGGATATTTATTTGCACTTTTATGTCAGCAAAGCAGGTGACTTCCCAGTCCAGCTTGAGTGAGAATGATGGATTCCAGGCA
TTTGTGTGGTGGTTACTAGGTATTGGGGCATTGACTTTTGCTCTTCTGATGTCAGCAAGGATGGGGATATTCCAAGAGACTCTCTACAAACGA
TTTGGGAAACACTCCAAGGAGGCTTTGTTTTATAATCACGCCCTTCCACTTCCGGGTTTCGTCTTCTTGGCTTCTGATATTTATGACCATGCAGT
TCTATTCAATAAGTCTGAGTTATATGAAATTCCCGTCATCGGAGTGACCCTGCCCATCATGTGGTTCTACCTCCTCATGAACATCATCACTCAGT
ACGTGTGCATCCGGGGTGTGTTTATCCTCACCACGGAATGCGCCTCCCTCACCGTCACGCTCGTCGTGACCCTACGCAAATTTGTGAGCCTCAT
CTTTTCCATCTTGTACTTCCAGAACCCCTTCACCCTGTGGCACTGGCTGGGCACCTTGTTTGTCTTCATTGGGACCTTAATGTACACAGAGGTGT
GGAACAACCTAGGGACCACAAAAAGTGAGCCTCAGAAGGACAGCAAGAAGAACTGA

MUR1 (At3g51160)

Genomic sequence from start to stop codon:

ATGGCGTCAGAGAACAACGGATCCAGATCCGATTCCGAATCCATCACCGCTCCCAAAGCTGATTCCACCGTCGTTGAACCGAGGAAGATAGC
TCTGATCACCGGAATCACCGGCCAGGACGGATCATACCTGACGGAGTTCCTTCTCGGAAAAGGCTACGAAGTTCATGGTCTGATCCGTCGATC
ATCGAATTTCAACACCCAGCGAATCAACCATATCTACATCGATCCACACAATGTCAACAAAGCTCTGATGAAACTCCACTACGCCGATCTCACC
GACGCTTCCTCTCTCCGTCGTTGGATCGATGTGATCAAACCTGACGAAGTTTATAACCTAGCTGCTCAATCTCACGTCGCTGTCTCCTTCGAGA
TCCCTGATTACACAGCCGATGTAGTCGCAACCGGTGCTCTCCGTCTCCTTGAAGCCGTCAGATCTCACACCATCGACAGTGGCCGTACCGTCA
AGTATTACCAAGCCGGATCTTCGGAGATGTTTGGATCAACTCCTCCTCCACAATCGGAGACGACGCCGTTTCACCCCAGATCTCCTTACGCAGC
TTCCAAATGCGCTGCTCATTGGTACACAGTGAATTACAGAGAGGCGTACGGTCTCTTCGCTTGTAACGGAATCTTGTTCAATCACGAGTCACCT
CGCCGTGGTGAGAATTTCGTGACGAGGAAGATCACAAGAGCATTGGGAAGGATCAAGGTTGGTTTGCAGACGAAGCTATTCCTTGGGAATT
TGCAAGCGTCAAGAGATTGGGGATTTGCAGGAGATTATGTGGAAGCAATGTGGTTGATGTTGCAGCAAGAGAAGCCAGATGATTACGTTGT
GGCAACAGAGGAAGGACACACAGTGGAAGAGTTTCTTGATGTGTCATTTGGGTATTTGGGACTCAATTGGAAAGATTATGTTGAGATTGACC
AGAGGTACTTTAGGCCTGCTGAAGTAGATAACCTTCAAGGAGATGCAAGCAAGGCAAAGGAAGTGTTGGGGTGGAAACCACAAGTAGGGTT
TGAGAAGCTTGTGAAGATGATGGTTGATGAAGATCTTGAGCTTGCTAAGAGGGAGAAAGTGCTTGTTGATGCTGGATACATGGATGCTAAG
CAGCAACCTTGA

Color	Name	Nucleotide change	Amino acid change
	axy2.4	C->T	S183F

MUR3 (At2g20370)

Genomic sequence from start to stop codon:

ATGTTTCCAAGGGTTTCTATGAGGCGTCGCTCTGCTGAAGTTTCTCCTACAGAGCCAATGGAGAAGGGAAATGGAAAAAATCAGACTAATCG
AATCTGTTTGTTGGTTGCTTTGTCACTCTTTTTCTGGGCATTGTTGTTGTACTTCCATTTTGTTGTCTTAGGGACTAGTAACATTGATAAACAGCT
CCAATTGCAACCTAGTTATGCTCAATCTCAGCCATCATCAGTCTCTCTTCGTGTTGATAAATTCCCTATAGAGCCTCATGCAGCTCCATCGAAAC
CACCTCCCAAAGAGCCTTTAGTTACTATCGATAAGCCAATACTCCCTCCTGCACCAGTAGCCAATTCTAGTTCCACCTTTAAGCCACCTAGAATT
GTTGAGAGCGGGAAGAAACAGGAGTTTTCGTTTATAAGAGCCTTGAAGACAGTAGACAACAAGAGTGACCCTTGTGGAGGGAAGTACATTT
ATGTTCACAATCTACCCTCAAAGTTCAATGAGGATATGCTTAGAGACTGTAAGAAGCTTAGTCTCTGGACTAACATGTGTAAGTTCACAACCA
ATGCTGGCCTTGGGCCGCCGCTTGAGAATGTTGAAGGTGTCTTCTCTGATGAAGGTTGGTATGCGACTAATCAGTTTGCTGTTGATGTGATAT
TCAGCAACCGGATGAAGCAGTACAAATGTTTGACAAATGATTCTTCCCTTGCTGCTGCCATCTTTGTGCCGTTCTATGCTGGATTCGACATTGC
TAGGTATCTGTGGGGATACAATATCTCGAGAAGGGATGCTGCTTCGCTTGAGTTGGTCGATTGGCTCATGAAGCGGCCTGAGTGGGATATAA
TGAGGGGAAAAGATCATTTCCTTGTGGCGGGTAGGATAACGTGGGATTTTAGGAGATTGTCTGAAGAAGAAACTGATTGGGGCAACAAGCT
TCTCTTCTTACCGGCTGCCAAGAACATGTCTATGCTTGTGGTTGAGTCGAGCCCATGGAATGCTAATGATTTTGGGATCCCGTACCCAACCTAC
TTCCACCCAGCAAAGGACTCAGAGGTTTTTGAATGGCAAGACCGGATGAGAAACCTGGAAAGGAAGTGGCTTTTCTCGTTTGCAGGAGCCCC
ACGACCTGACAATCCGAAATCTATCAGAGGGCAGATCATTGATCAATGTCGAAACTCAAATGTTGGCAAATTATTAGAATGCGATTTTGGGGA
GAGCAAATGCCATGCACCAAGCAGCATTATGCAAATGTTTCAAAGCTCTCTCTTCTGCCTGCAGCCACAGGGAGACTCTTACACCCGAAGATC
GGCATTTGACTCAATGCTTGCCGGTTGTATACCGGTCTTCTTCCATCCGGGTTCAGCCTACACGCAATACACGTGGCATCTACCCAAGAACTAC
ACAACCTACTCGGTATTCATCCCCGAGGATGATGTTCGGAAGAGAAACATAAGCATCGAGGAGCGACTCCTCCAGATTCCAGCCAAGCAGGT
CAAGATAATGAGAGAGAATGTCATCAACCTCATCCCAAGGCTGATCTACGCAGACCCGAGATCAGAACTGGAGACGCAGAAAGATGCATTTG
ATGTCTCGGTACAGGCTGTGATAGACAAGGTGACTCGGTTAAGGAAGAACATGATCGAGGGCCGAACCGAGTATGACTACTTCGTGGAGGA
GAACAGCTGGAAGTATGCGTTGCTGGAGGAAGGCCAACGGGAAGCAGGGGGCCATGTGTGGGACCCGTTCTTCTCCAAACCAAAGCCCGG
AGAAGATGGCAGCAGCGATGGTAATGGAGGCACGACTATTTCAGCAGATGCAGCTAAGAATTCATGGAAGAGTGAGCAGAGAGATAAGAC
ACAGTGA

Color	Name	Nucleotide change	Amino acid change
	8-52	G->A	C400Y

AXY3 (At1g68560)

Genomic sequence from start to stop codon (introns lowercase):

ATGGCTTCCTCTTCTTCATCTCTAGCTTTCTCTCTTTCTCTTCTTCTAGCTTTAATCCTCTGTTTCTCACCCACCCAATCCTACAAAACCATCGGCA
AAGGCTACCGTTTGGTCTCCATCGAAGAATCTCCCGACGGTGGTTTCATTGGTTATCTCCAAGTCAAACAAAAGAACAAAATCTACGGCTCTG
ATATCACCACTCTTCGTCTCTTCGTCAAgtaaaaaccaaatcctatgtttcgcaccctacattacttaccaaaactcttcttgatttctaaccgttagatttgtttactctttccaac
agGCACGAAACTGATTCACGTCTTCGTGTTCACATCACTGACGCGAAACAACAACGATGGGAAGTTCCATACAATCTCCTCCCACGAGAACAA
CCACCACAAGTCGGAAAAGTCATCGGAAAATCAAGAAAATCTCCAATCACAGTTCAAGAAATCTCCGGATCCGAACTTATCTTCAGTTACACA
ACAGATCCATTCACTTTCGCCGTGAAACGTAGATCTAACCACGAAACACTCTTCAACACAACAAGCTCATTAGTGTTCAAAGATCAATACCTAG
AGATCTCAACTTCATTACCAAAAGAAGCTTCACTCTATGGTTTAGGTGAAAACTCACAAGCTAATGGTATCAAACTCGTACCTAACGAGCCTTA
TACTCTTTACACTGAAGATGTTTCAGCTATCAATCTTAACACTGACCTTTATGGTTCGCATCCGATGTATATGGATCTAAGAAACGTTGGAGGT
AAAGCTTATGCTCATGCTGTTCTTCTTTTGAATAGTAATGGTATGGATGTTTTCTATAGAGGTGATTCTTTGACGTATAAAGTTATCGGAGGTG
TTTTCGATTTCTATTTCATCGCCGGACCGTCGCCGTTAAATGTTGTTGATCAGTATACTCAGTTAATCGGAAGACCTGCTCCTATGCCGTACTGG
TCTTTAGgtatgttcatttactgctctgttttttttgttattgttttctctctctcttgtcttaatgttttgtggtaatttaaataggattagtgattatgtttgtggtcctatagtgttttaagcat
aggttaatgtcttaatgatgagtattagtattggttcaggttcctataagtttgacttttagtcaaacgaagagtcacgtgaacatagattttgtgtatcaaccaaaccaagtgagatcgaa
aactattaaattaatgttccattagtgttgatatttcattgtaaagagtatctgattttccgacacaacttttagaatagaaaagctttaccctttctgagggatattggaaagataatgaaa
agtttcgctttctaattatggtgtttccaattaaaagataagaacgagagattggtctttgctttccatttttgttgatcttgataagaatgattattgcgattttatttggtccctataggatctt
gaaattggaccattatatagacatggaccacggttcttgaaatgagttgaaaataaatgcaaatcttgattttgccatgtgatagatttatggttatactaatcgtttatattaatgttgtgtt
gtagGATTTCATCAATGTAGATGGGGATACCATAATCTATCAGTTGTTGAAGATGTGGTGGATAATTACAAAAAGGCTAAGATCCCACTTGATG
TGATTTGGAACGATGATGATCACATGGATGGTCACAAAGATTTCACGTTGAATCCTGTGGCTTATCCTCGTGCTAAGCTATTGGCGTTCTTGG
ACAAAATCCACAAGATTGGAATGAAGTATATTGTGATCAATGATCCTGGTATTGGTGTCAACGCTAGCTACGGTACATTTCAGAGAGCCATGG
CTGCTGATGTGTTTATTAAGTATGAAGGAAAGCCATTCTTGGCTCAAGTGTGGCCTGGTCCGGTTTACTTCCCTGATTTCCTTAACCCGAAGAC
GGTTTCTTGGTGGGGTGATGAAATCAAACGGTTCCATGATTTAGTCCCTATTGATGGTTTATGGATTGACATGAACGAGGTGTCGAACTTTTG
TTCTGGTTTATGCACAATCCCTGAAGGAAAGCAGTGTCCAAGTGGGGAAGGTCCTGGTTGGGTTTGTTGCTTGGACTGCAAGAATATAACCA
AGACTAGATGGGATGATCCTCCTTACAAGATTAATGCTACTGGTGTCGTAGCTCCTGTTGGGTTCAAGACTATTGCTACAAGTGCTACTCATTA
CAATGGTGTTCGAGAGTACGATGCTCACAGTATCTACGGATTCTCTGAGACCATCGCAACTCACAAGGGCTTACTTAATGTCCAAGGCAAACG
TCCCTTCATTTTATCCCGGTCTACTTTCGTTGGTTCGGGTCAATATGCGGCTCACTGGACCGGAGATAACCAAGGAACATGGCAGAGCTTGCA
AGTGTCTATCTCAACTATGTTGAACTTCGGTATATTTGGAGTTCCCATGGTTGGTTCAGACATTTGTGGATTCTACCCACAACCAACAGAAGAA
CTCTGCAACCGTTGGATCGAAGTAGGCGCGTTTTACCCGTTTTCAAGAGATCACGCCAATTACTACTCACCAAGACAAGAGCTTTACCAATGG
GACACAGTTGCAGACTCAGCTCGTAACGCTCTTGGTATGAGATACAAGATCCTTCCTTTCCTCTACACTCTTAACTACGAAGCGCATATGACTG
GTGCACCCATCGCTAGACCGCTCTTCTTCTCATTCCCTGAATACACCGAATGTTACGGCAACAGCAGACAGTTCTTGCTTGGAAGCAGCTTCAT
GATATCTCCGGTTCTCGAGCAAGGCAAAACCGAAGTAGAAGCATTGTTCCCACCAGGTTCTTGGTACCACATGTTCGACATGACTCAAGCAGT
AGTGTCCAAGAACGGGAAACGTGTAACTCTCCCAGCTCCTTTGAACTTTGTGAACGTTCATCTTTACCAAAACACTATCTTACCAACGCAACAA
GGCGGGTTGATTTCCAAGGACGCAAGAACAACACCGTTTAGCCTCGTCATTGCTTTCCCTGCAGGAGCTTCTGAAGGTTACGCCACCGGGAA
ACTCTATCTAGACGAAGACGAGCTTCCGGAAATGAAGCTAGGAAATGGACAGTCCACGTACGTTGACTTCTACGCTTCGGTCGGAAACGGGA
CAATGAAGATGTGGTCACAAGTTAAAGAAGGTAAGTTTGCATTGAGCAAAGGATGGGTGATTGAGAAAGTGAGTGTTTTGGGACTTAGAGG
AGCAGGACAAGTGTCTGAGATTCAAATCAATGGAAGTCCAATGACAAAGAAGATTGAGGTTAGCTCGAAGGAGCATACCTATGTGATTGGTT
TGGAAGATGAAGAAGAGAACAAGAGTGTGATGGTTGAGGTTAGAGGACTTGAGATGCTTGTTGGAAAGGATTTTAACATGTCCTGGAAAAT
GGGTATCAATTAA

Color	Name	Nucleotide change	Amino acid change
	236-81	G->A	C400Y
	axy1.160	G->A	W482Stop
	axy1.3	G->A	D529N
	148-100	G->A	E630K

AXY8 (At4g34260)

Genomic sequence from start to stop codon (introns lowercase):

ATGGCGGAGAAGTCGAGCTTCTTCGTTCATTTCAGCTGTCTTCTGCTTCTTCTAACTATCATAATAACATGTGGAGAGGGAGTGAGGAATCCG
GTGAGGCCAAGATCATCGGAGAGAAGAGCATTGATGGACGGCCAAGATCTCTCGAGGCCATTGAAGCTGACATTTGGTGGGCCTTCTCGTA

ATTGGACTGACGCTATTCCCATCGGTAATGGCCGTCTCGGAGCCACGATCTGGGGCGGCGTCTCCTCTGAGATTCTCAACATCAATGgtaaatgt
taaaacggcgtcgttttttttggatttttgtccttttctctcccattctttgtgttgttttttatttttggagatttggttaataaatgctatctgctgtgttctctaaaatctattgtcagagatcaatc
aatgttcttgtttctgaacaggggattatgcttgagacttgaaataaattgtgtaaatgaaatacattgctacaaaagcataagcatttctgaatcttgtttccgctaagatgattttgtggat
tttgattgatagAGGACACAATTTGGACTGGAGTACCAGCAGACTACACGAACCAGAAGGCCCCTGAAGCATTAGCAGAAGTGAGGAGGCTTGT
TGATGAAAGAAACTATGCTGAAGCTACTTCAGAAGCTGTCAAATTATCAGGCCAGCCTTCTGATgtatgttctattcttttggtgaaaatgttgctttcacttct
tgtttttttgtttgttatgtgtaaagactaatccttttgtttgtggcctctaaagGTTTACCAGATTGTGGGAGATTTGAATTTAGAGTTTGACAGTTCGCATCGGAAG
TACACTCAAGCCTCTTATCGCCGTGAACTCGACTTGGAAACTGCTGTAGCTAAAGTAAGTTACTCGGTTGGGGCTGTGGACTTCTCAAGGGAG
TTTTTCGCTTCTAATCCGGATCAAGTAATAATAGCCAAGATCTATGCAAGCAAACCGGGATCATTGTCTTTTAAGGTGTCTTTTGACAGCGAGT
TGCACCATCATTCAGAGACAAATCCCAAGGCTAATCAGATTTTGATGCGGGGAAGTTGTCGCCCCAAGCGACTTCCTGTGAACTTGAAGAAAA
GTATCAACGCGACAAATATACCGTATGATGATCACAAAGGATTGCAGTTTGCTTCGATTCTTGAAGTGAGAGTCAGCAACGGTGGCTCTGTGT
CTTCCTTAGGTGGCAAGAAACTAAGTGTTGAAAAGGCAGACTGGGCGGTTCTGCTTCTGGCGGCTTCATCTAACTTTGATGGGCCATTTACTA
TGCCTGTTGATTCTAAGATAGACCCTGCAAAGGAATGCGTTAACAGAATCAGCTCAGTCCAGAAATACTCGTACTCTGATCTTTATGCTCGTCA
TTTGGGTGACTATCAGAAACTTTTCAACCGGGTCTCTTTACACCTTTCTGGAAGCTCAACAAATGAAACTGTGCAACAGGCAACTTCAACTGCT
GAAAGGGTGCGATCATTTAAAACCGATCAAGACCCTTCATTGGTGGAACTTCTATTCCAATATGGTCGATACCTGCTTATATCTTCTTCAAGAC
CCGGGACTCAGGTGGCTAACTTGCAGGGTATATGGAACAGGGACATTCAGCCACCATGGGAgtgcgttttcgccttactatcttttcttggactttcatagcttt
tgttgtatcactaactccacttcttcttctgcggtattgcagTGGTGCTCCTCACCTGAACATTAATCTGCAGATGAACTATTGGCATTCTCTTCCTGGTAATATC
CGGGAATGTCAGGAACCATTGTTTGATTATATGTCTGCTTTAGCAATTAATGGCCGCAAGACAGCACAAgtaagtttctatccttttttctaatgatgaccttt
tgcaatgcatttgtgctccatcactagagacattagtctaagtatcattttgttgcgctgtatagGTGAACTACGGAGCAAGTGGTTGGGTGGCACATCAAGTATCAGA
CATATGGGCAAAAACTTCACCAGATAGAGGTGAGGCAGTGTGGGCACTCTGGCCCATGGGTGGAGCTTGGCTCTGCACTCATGCATGGGAG
CATTATACATACACAATGGACAAGgtatggacctcaattttctgattttcaaatcagtttctttaaccggtactatgtttaatcagcttttacttgttgattcatttctagGAATTT
CTTAAAAAGAAAGGATATCCATTGTTGGAAGGATGTACGTCGTTTCTTTTGGACTGGCTGATCAAAGGTAAAGATGGATTTCTCCAGACCAAT
CCATCGACTTCCCCTGAGCATATGTTTACTGCTCCCATTGGCAAACCTGCTAGTGTTAGCTACTCATCAACCATGGACATTGCCATTATAAAGG
AAGTCTTCGCTGATATCGTTTCAGCTTCAGAGgtaacatgaagcaacaaaaacaacaaaacaatcatatcacctcatggatactgagcttatggtttctggtttatcaaca
gATTTTGGGAAAAACAAATGATACTCTTATTGGAAAAGTCATTGCTGCTCAAGCAAAACTTCCACCCACTAGAATATCAAAGGACGGTTCCATT
AGGGAATGGGCAGAAGACTTTGAGGACCCAGAGGTACATCATAGACATGTTTCACATCTCTTTGGCCTCTTCCCGGGTCACACAATCACTGTT
GAGAAGTCTCCAGAACTTGCTAAAGCTGTGGAAGCAACTCTTAAAAAAAGAGgtttgtgtgattctttcttctctcccattactatttccacaatttcatcgacgtgg
ttggagcaatatataatgttccactgattgacgcagGAGAGGAAGGACCAGGATGGTCAACCACCTGGAAAGCCGCCTTGTGGGCAAGGCTTCACAATAG
CGAGCACGCATATCGGATGGTTACACATATTTTTGATCTTGTGGATCCGTTGAACGAACGTAACTATGAAGGAGGACTGTATAGTAACATGTT
CACTGCACACCCGCCATTCCAGATCGATGCTAATTTTGGgtaagttgttgcaaacacttacctataagatatgtggaatgtttacattcacattccaacgcactataccat
ttgatgactagtctaaacttacacatgagctttttttccttggcagATTCGCAGCAGCGGTAGCAGAAATGTTAGTGCAGAGCACTACGAAGGATCTATATTTGC
TCCCGGCACTTCCAGCAGATAAATGGCCTAACGGCATTGTGAATGGGCTAAGGGCAAGGGGCGGTGTGACGGTTAGTATAAAGTGGATGGA
AGGGAACCTTGTGGAGTTTGGTCTTTGGTCGGAACAGATTGTTTCCACCAGAATAGTGTACAGAGGAATCTCTGCGGCTGCCGAATTGTTGCC
AGGGAAAGTCTTTACTTTTGACAAGGATTTGAGATGCATTCGGACCGACAAGTTATAG

Color	Name	Nucleotide change	Amino acid change
	148-69	G->A	W95Stop
	129-16	G->A	splice site disruption
	142-71	G->A	W441Stop

AXY9 (At3g03210)

Genomic sequence from start to stop codon:

ATGCTAGGAGCGATTCATTTGGGAGTATTAGCAGCTTGTTTCGTCCTCTTCGTTCCCATGGCAATGGCTGGTTGGCATCTGAGCAGGAACAAG
ATGCTCTTCTTTAGTGGCGCTTTGTTTATTTCTCTCGCGGTTTGTGTTCATCTCACGCCTTATTTCCCTTCCGTTTCCGACATCGTTGCCTCTGTTT
CCTCTGTTGTTGTCTACGATCATCGTATCTCTTGTATCAACGAGGTTAATCAGATTGTCTGGGATGTTAAACCGGTACCGAATCCGGAGTCTGT
TCGTAGGAACAATGGCTCAACAAAGCTTGATTATTTCGTGAAGAACTGGGATTGGATGAAATCGAGGAAAGTTTTGAGTTGTGAGTTTCAGA
AATTGGATAAGTTCGATGTTTCGGATTTGTTGAATGGATCGTGGGTTGTTGTTGCTGGTGATTCTCAGGCGAGATTCGTGGCTTTGTCTTTGTT
GAATCTTGTTTTGGGTTCGGATTCCAAGGCTATGGATTCGGTTAGAGGGGATTTGTTTAGGAGACATAGTGATTACAGCATTGTGGTTAAGG
AGATTGGGATGAAGCTTGATTTTGTTTGGGCTCCTTATGAGAAAGATTTGGATGATCTTGTGGTATCTTATAAGAAGATGAAAAAGTATCCTG

ATGTTGTGATCATGGGAACTGGGTTATGGCATATGCTTCATGTGAACAATGCTTCAGATTTTGGTTTTAGGTTGAGGCAATTGAGTAGTCATG
TGGAGTCTCTAGTGCCCCTTACGCCAAAGGAGCAAGAAGGAGGTGGGTCGGTCTCCGGTAGATCTGTGCATCTGTTCTGGATTGGGATGCCG
GTTTTGATCAACGGGATGTTGAATACGGATGAGAAGAAGGAGAAGATGAGTGATACGGTGTGGCATGAGTATGATAGATCACTCGGGGAG
AGTAAGATCTTACGTCAAATGGGTGGTCCTCTTATCTTGCTTGATATCCAGTCGTTTACCTGGAACTGTGGACCGCAATGTACACTTGACGGG
ATGCATTATGACTCTGCGGTCTATGATGCCGCTGTCCATGTCATGCTCAATGCGTTGCTTATTGAATCTCATCAATCTTTATGA

Color	Name	Nucleotide change	Amino acid change
	axy9.1	G->A	W276Stop
	axy9.2	T-DNA insertion	

Genomic sequence from start to stop codon in *axy9.2*:

ATGCTAGGAGCGATTCATTTGGGAGTATTAGCAGCTTGTTTCGTCCTCTTCGTTCCCATGGCAATGGCTGGTTGGCATCTGAGCAGGAACAAG
ATGCTCTTCTTTAGTGGCGCTTTGTTTATTTCTCTCGCGGTTTGTGTTCATCTCACGCCTTATTTCCCTTCCGTTTCCGACATCGTTGCCTCTGTTT
CCTCTGTTGTTGTCTACGATCATCGTATCTCTTGTATCAACGAGGTTAATCAGATTGTCTGGGATGTTAAACCGGTACCGAATCCGGAGTCTGT
TCGTAGGAACAATGGCTCAACAAAGCTTGATTATTTCGTGAAGAACTGGGATTGGATGAAATCGAGGAAAGTTTTGAGTTGTGAGTTTCAGA
AATTGGATAAGTTCGATGTTTCGGATTTGTTGAATGGATCGTGGGTTGTTGTTGCTGGTGATTCTCAGGCGAGATTCGTGGCTTTGTCTTTGTT
GAATCTTGTTTTGGGTTCGGATTCCAAGGCTATGGATTCGGTTAGAGGGGATTTGTTTAGGAGACATAGTGATTACAGCATTGTGGTTAAGG
AGATTGGGATGAAGCTTGATTTTGTTTGGGCTCCTTATGAGAAAGATTTGGATGATCTTGTGGTATCTTATAAGAAGATGAAAAAGTATCCTG
ATGTTGTGATCATGGGAACTGGGTTATGGCATATGCTTCATGTGAACAATGCTTCAGATTTTGGTTTTAGGTTGAGGCAATTGAGTAGTCATG
TGGAGTCTCTAGTGCCCCTTACGCCAAAGGAGCAAGAAGGAGGTGGGTCGGTCTCCGGTAGATCTGTGCATCTGTTCTGGATTGGGATGCCG
GTTTTGATCAAATTGACGCTTAGACAACTTAATAACACATTGCGGACGTTTTTAATGTACTGGGGTGGTTTTTCTTTTCACCAGTGA

10 Appendix 2: Vector maps

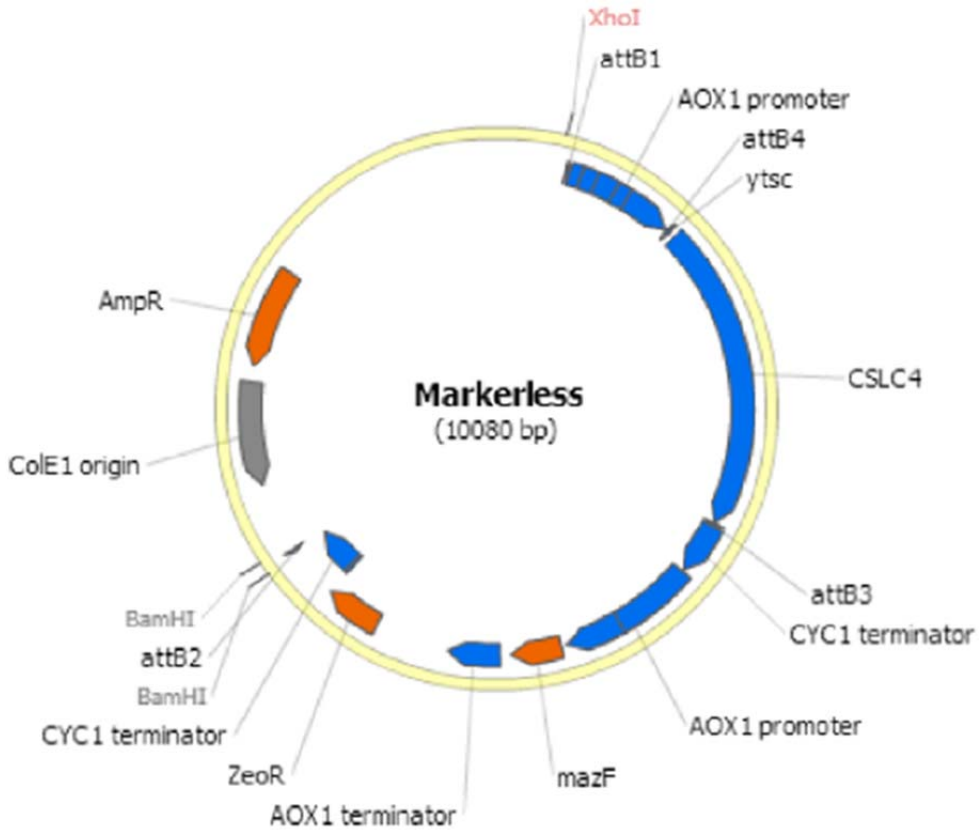


Figure 10-1. Vector map for *P. pastoris* constructs with the excisable markerless cassette. This vector was used for expression of CSLC4 (complete construct shown above), XXT2, UGD and UXS in *P. pastoris*. The selectable (ZeoR) and counterselectable (MazF) markers are flanked by direct repeats of the CYC1 terminator, allowing for excision of the cassette which is selected for by growth on methanol. The vector was linearized with XhoI and BamHI prior to transformation. The CSLC4 and XXT2 constructs contained an N-terminal T7 tag (ATGGCTAGCATGACTGGTGGACAGCAAATGGGTACT). The UGD and UXS constructs contained an N-terminal myc tag (ATGGAACAGAAGCTGATTTCTGAGGAAGATTTG) and FLAG tag (ATGGATTACAAAGACGACGATGACAAA) respectively. ytsc – yeast translational start context (TTCAAACAAA).

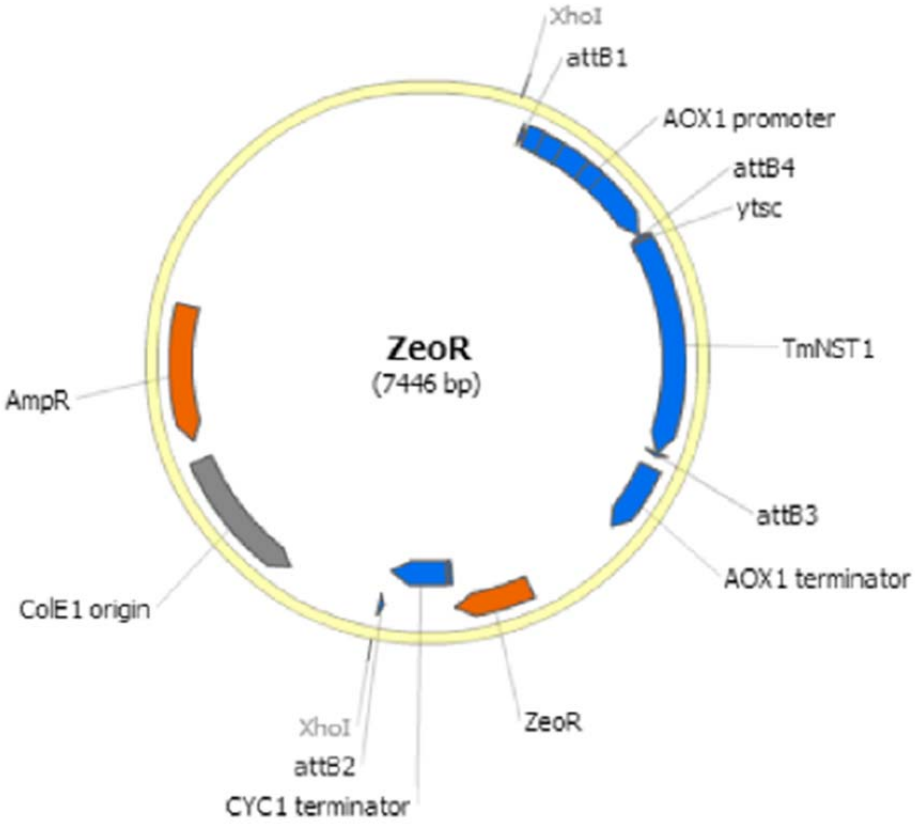


Figure 10-2. Vector map for *P. pasto*ris expression plasmids with the ZeoR cassette. This vector was used for expression of TmNST1 (complete construct shown above), TmNST2, TmNST3, TmNST4, SLC35B4 and SLC35D1 in *P. pastoris*. The construct lacks the counter-selectable marker so the ZeoR cannot be excised but transgene expression can be done immediately following the selection and verification of transformants. The vector was linearized with XhoI prior to transformation. ytsc – yeast translational start context (TTCAAACAAA).

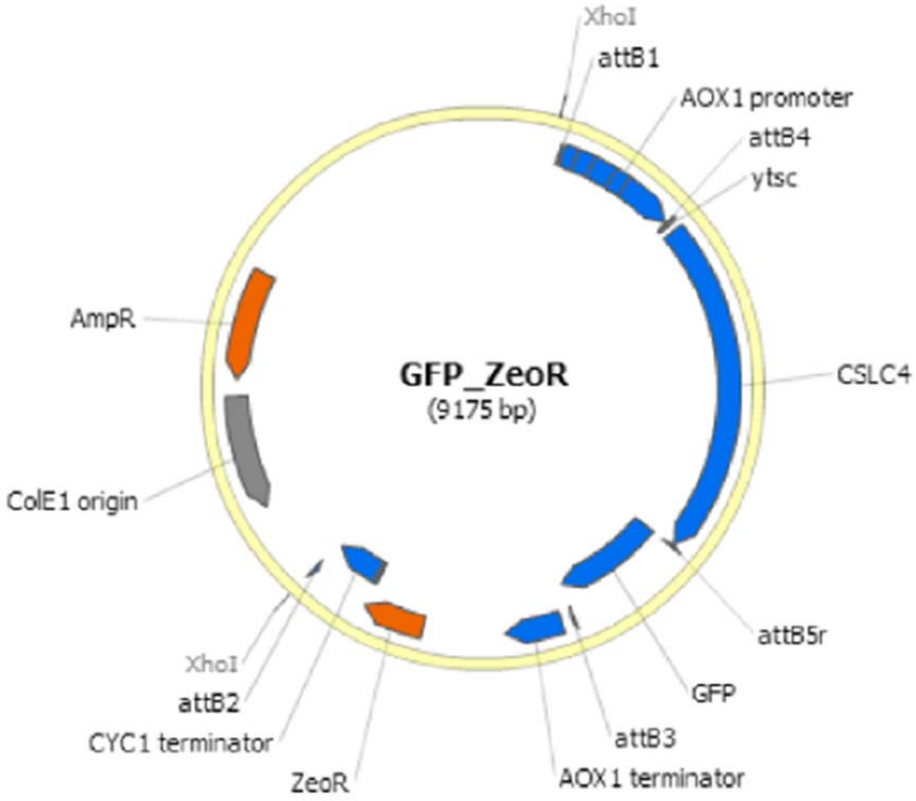


Figure 10-3. Vector map for *P. pastoris* constructs with GFP fusions and the ZeoR cassette. This vector was used for expression of CSLC4 (complete construct shown above) and XXT2 with C-terminal GFP fusion in *P. pastoris*. The construct lacks the counter-selectable marker so the ZeoR cannot be excised but transgene expression can be done immediately following the selection and verification of transformants. The vector was linearized with XhoI prior to transformation. ytsc – yeast translational start context (TTCAAACAAA).

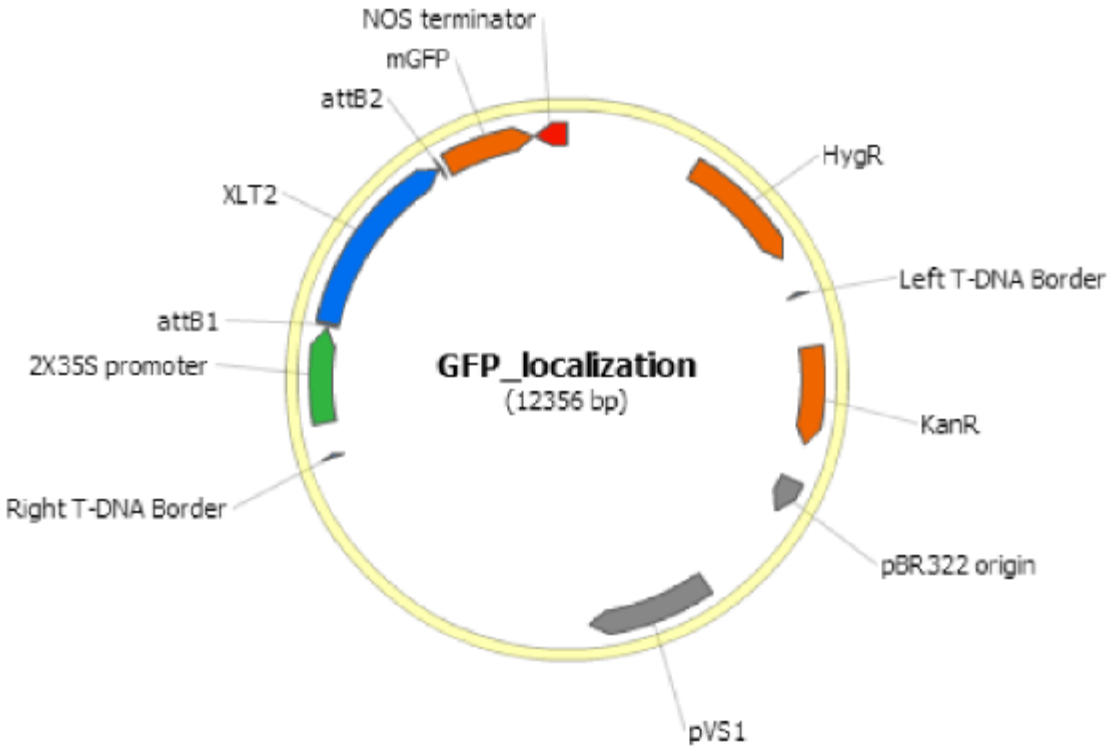


Figure 10-4. Vector map for GFP localization in *N. benthamiana* and *A. thaliana*. This vector was used for expression of C-terminal GFP fusions with XLT2 (shown above) and AXY9. The vector was transformed into *A. tumefaciens*, which was subsequently used for transient expression in *N. benthamiana* or stable expression in *A. thaliana*.

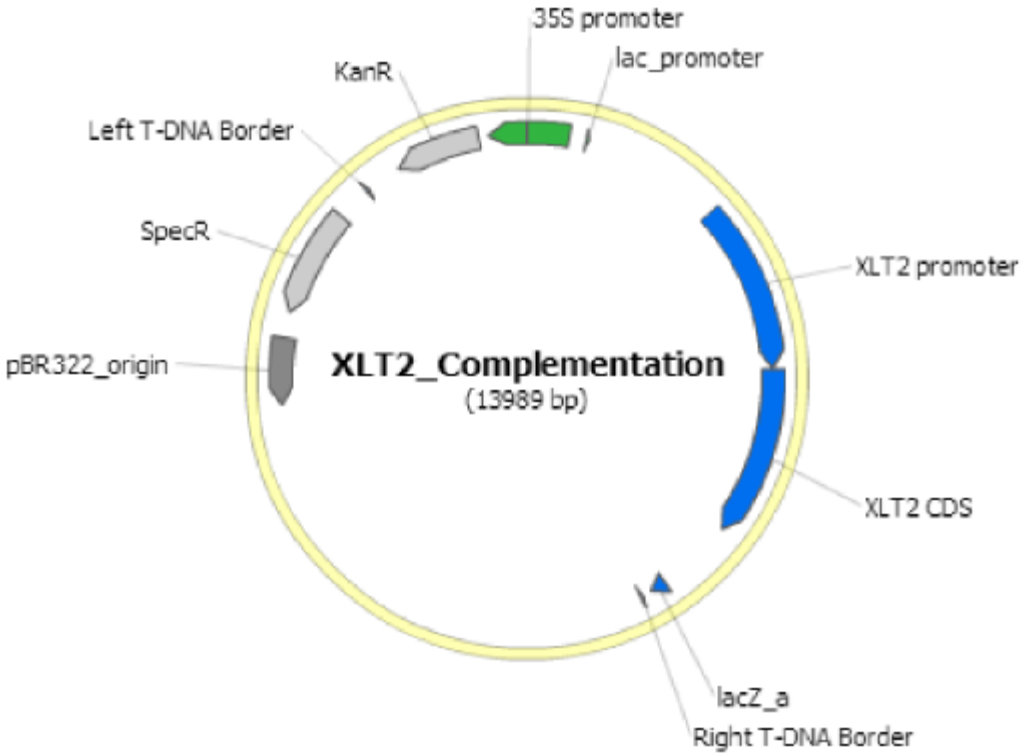


Figure 10-5. Vector map for plasmid used for complementation of the *xlt2* mutant. The XLT2 gene, including the promoter, coding sequence and terminator, was cloned into pPZP211 for transformation into *A. thaliana* (Hajdukiewicz et al. 1994).

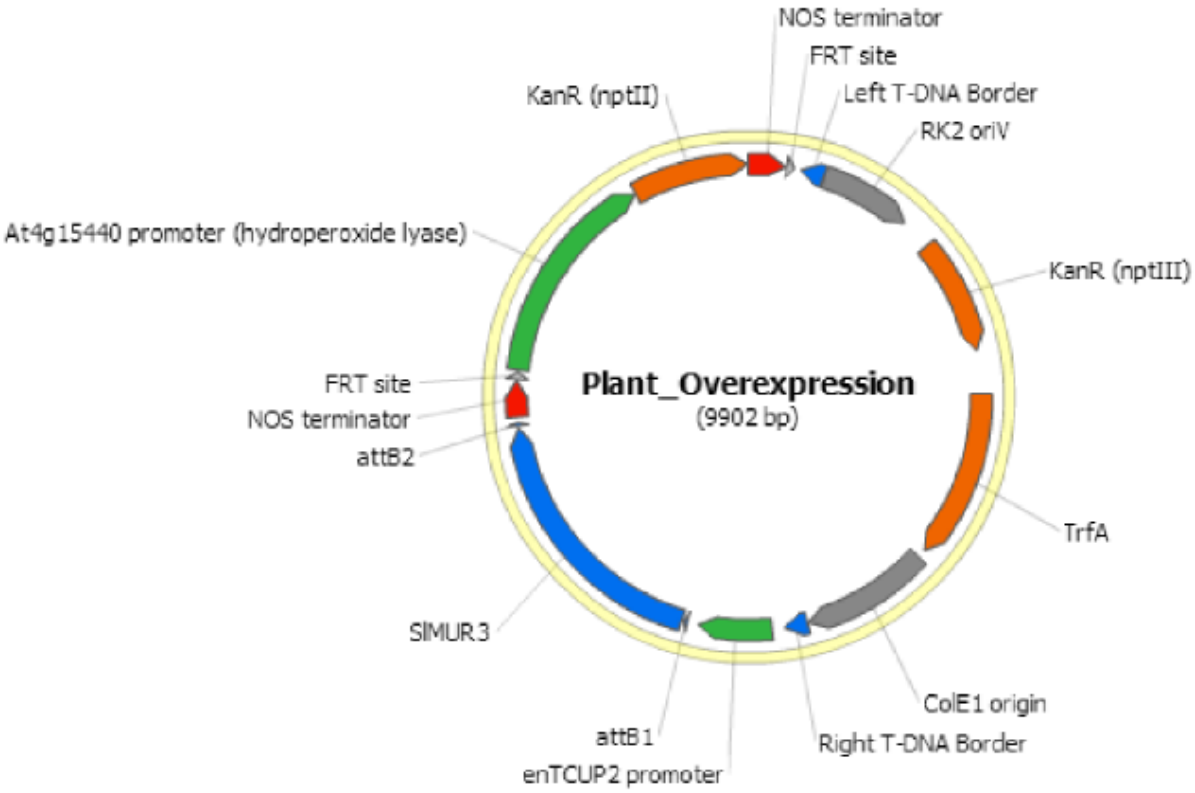


Figure 10-6. Vector map for a plasmid used for overexpression of genes in *A. thaliana*. This vector, a gateway compatible version of PORE E4 (Coutu et al. 2007),was used for expression of SlMUR3 (shown above), XST1, XST2 and Sl02g092840.

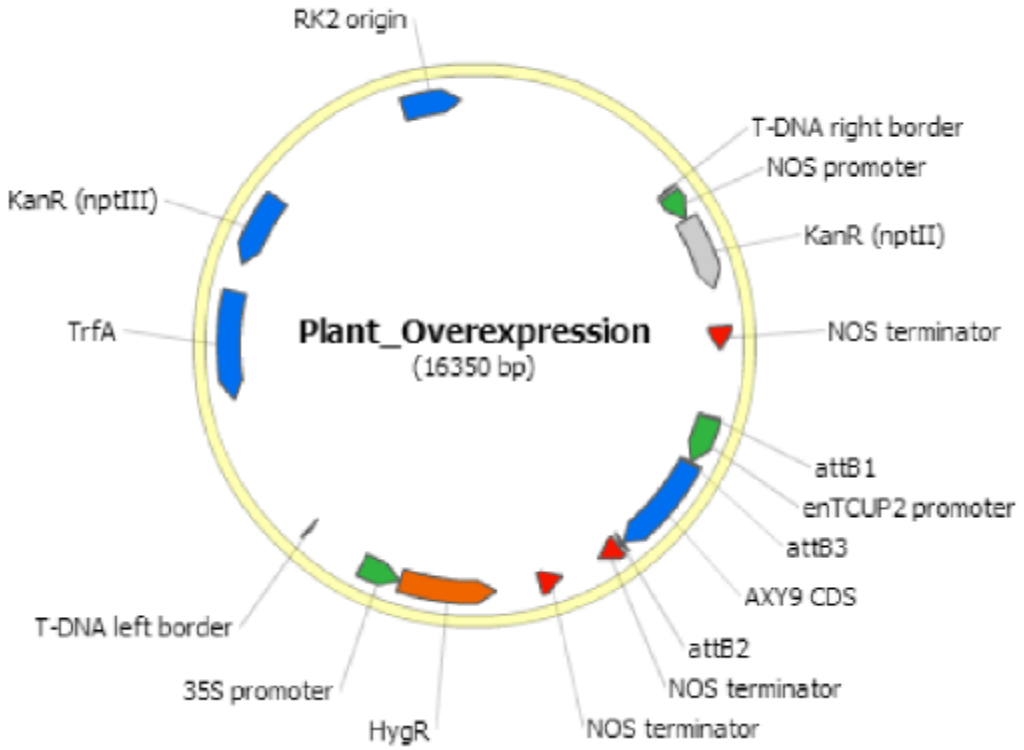


Figure 10-7. Vector map for a plasmid used for plant overexpression. The Gateway-compatible pGWB1 vector (Nakagawa et al. 2007) was used for overexpression of AXY9 under the control of the enTCUP2 promoter. A two-piece Gateway system was used for fusion of the promoter and protein coding sequence (Invitrogen). This vector provides transformed plants with resistance to both kanamycin and hygromicin.

11 Appendix 3: Genotyping primer combinations

Name	Position (Mb)	Primer sequence Forward	Reverse	PCR product length Col0	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					
NGA172	0.8	CATCCGAATGCCATTGTTC	AGCTGCTTCCTTATAGCGTCC	162	136
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

Table 11-1. Simple Sequence Length Polymorphism (SSLP) Markers. Simple sequence length polymorphism (SSLP) markers used for rough mapping *axy* mutants. Markers from www.arabidopsis.org, tested by Sascha Gille, a member of the Pauly Lab.

Name	Position (bp)	Enzyme	Primer sequences		Product length	
			Forward	Reverse	Col0	Ler
axy6 (chromosome 1)						
6fm1	6,132,583	EcoRI	AATCATCTTTTTCAGCAATC	CTGAACAAGTTTCTCAGGTT	243, 309	552
6fm2	7,370,859	EcoRI	CTCGATTTGGAAGTTTGTAT	TAACAACTAGGTAAGTGGGG	192, 325	517
6fm4	9,563,305	EcoRI	AATTGAATCATCACACACAA	GGTATGTTCACAACCTTCTC	224, 353	577
6fm6	11,325,127	EcoRI	AGAAGGTCCATCTAATTTTG	TCTCAAATAAGAACTTTGGG	230, 275	505
6fm7	12,880,194	EcoRI	ATTTATGAGAATGAACGAGC	CGAGTGACCAAATAAGAAAT	471, 306	777
6fm10	13,340,164	EcoRI	CACTTCTGGACAGAACTCTT	GTGAAAGCTTCTCATTTTTC	324, 166	490
6fm12	15,909,937	EcoRI	GAATGAGAGTGTCAGAGCTT	ACCATAATTTGTTTAGGTCG	256, 227	483
6fm13	16,788,779	EcoRI	GTGGAGTATATTGGGTTCAG	GACTGTTACCTCTCTCATGG	161, 392	553
axy9.1 (chromosome 3)						
9fm1	88,634	EcoRI	GACAAATCACAATAAGCACA	TTCAAACCCTAATTATCGAC	482, 135	617
9fm2	651,406	EcoRI	CAAAGATCGGTAGTCAAACT	TTTGTATAGGCCCTAAGGTA	457, 389	846
9fm4	1,592,504	EcoRI	CCTTCCTTTATTAGCTGTTG	CCAATATTACGAGTAAAGGC	322, 404	726
9fm5	1,780,116	EcoRI	GAGCTAATTCCCTCCTAAAT	GAAAACTTGGTTGGTAAGAA	345, 317	662
9fm8	4,368,941	EcoRI	ATCAGTAATGCTTCTATGGC	AAATGCTAGTTATGAGTGGG	304, 489	793
axy2.4 (chromosome 3)						
2fm1	12,974,235	EcoRI	TGGTAGTATGTCGTCGATGT	CGAGTTGCGAATTTTAACGT	478, 385	863
2fm2	14,172,292	EcoRI	CCATCTCTTTGAACCACTCC	CCAAAAAGCCCTAGGAACAA	427, 391	818
2fm4	16,034,096	EcoRI	ACTAGTTGGCGACTGAAAAG	GTGCAGTGAACTAGCAGAAT	463, 388	851
2fm5	16,990,400	EcoRI	CCGGTTAATCCGGTTATTCA	TTCTTCTGCAATCGAATCCC	439, 321	760
2fm7	18,725,094	EcoRI	TCGAACCTCTCGTCATAACA	GAGTGAGTATGAGAGCTGGA	477, 403	880
2fm8	19,572,056	EcoRI	GAGATTGATTTACCGGCGAT	GTTAACCAGTCTCTTAGCCG	492, 171	663
2fm9	20,858,475	EcoRI	GCTGCCTTTTCAAAGTAAGC	TCTAAAACTTGCCGACTGTC	414, 247	661
2fm10	21,997,821	EcoRI	TTGCAATTCACCGGAACTTA	CGCCAAGTAATGAACAGACT	492, 295	787
2fm11	22,974,689	EcoRI	ATGTATGGTGGATGAATGCC	GCCTCCTCTGAAGAAGGAATA	352, 272	624

Table 11-2. Cleaved Amplified Polymorphic Sequence (CAPS) Markers. Cleaved amplified polymorphic sequence markers. Following PCR, the product is digested with the indicated enzyme. The product from one of the ecotypes (Col0) will be cut whereas Ler will not. These CAPS markers were used for fine mapping *axy* mutants. Markers were obtained by automated searching using the Col0 reference genome, a Ler SNP list, NCBI BLAST, and primer3 as described (2.2.3).

Line	T-DNA Accession	Forward primer	Reverse primer	Mutant Forward / Notes
xlt2		AS-82	AS-83	GABI-Lb
xlt2_Comp		M13F	AS-120	
mur3.1		mur3_FOR	mur3_REV_MISM	digest with TaqI
SlMUR3		AS-139	AS-138	
XST1		AS-145	AS-144	
XST2		AS-447	AS-147	
Sl02g0928401		AS-142	AS-141	
axy9.2	SALK_090612	AS-289	AS-290	Salk-LBa1
axy3.2	GABI_749G08	GABI_749G08-LP	GABI_749G08-RP	GABI-Lb
axy9 OE		AS-379	AS-440	
axy9-GFP		M13F	AS-440	
	SALK_124729	AS-287	AS-288	Salk-LBa1
	SALK_065776C	AS-291	AS-292	Salk-LBa1
	SALK_077065	AS-285	AS-286	Salk-LBa1
MUR3 T-DNA	CS1007414	AS-133	AS-134	Ski3

Table 11-3. Genotyping primer combinations for *A. thaliana* lines. The primer combinations listed above were used to check for the presence of mutations / T-DNA insertions using PCR.

Gene	Construct	Primer 1	Primer 2	Product Length
Actin	Actin (native)	AS-122	AS-123	632
CSLC4	CSLC4-Markerless	AS-18	AS-21	576
	CSLC4-Excised	AS-18	AS-74	505
XXT2	XXT2-Markerless	AS-19	AS-21	696
	XXT2-Excised	AS-19	AS-74	625
UGD	UGD-Markerless	AS-111	AS-21	572
	UGD-Excised	AS-111	AS-74	501
UXS	UXS-Markerless	AS-113	AS-21	652
	UXS-Excised	AS-113	AS-74	581
TmNST1	TmNST1-ZeoR	AS-186	AS-22	362
TmNST2	TmNST2-ZeoR	AS-187	AS-22	470
TmNST3	TmNST3-ZeoR	AS-188	AS-22	329
TmNST4	TmNST4-ZeoR	AS-189	AS-22	422
TmXXT5	TmXXT5-ZeoR	AS-218	AS-22	857
SLC35B4	SLC35B4-ZeoR	AS-279	AS-22	580
SLC35D1	SLC35D1-ZeoR	AS-280	AS-22	318
CSLC4-GFP	CSLC4-GFP-ZeoR	AS-18	AS-22	1007
XXT2-GFP	XXT2-GFP-ZeoR	AS-19	AS-22	1126

Table 11-4. Genotyping primer combinations for *P. pastoris*. The specified primer pairs were used in PCR reactions with P. pastoris cells to assay for the presence of the specified construct.

12 Appendix 4: Primer sequences.

Name	Purpose	Sequence
M13F	general sequencing/genotyping	TGTAAAACGACGGCCAGT
M13R	general sequencing/genotyping	GGAAACAGCTATGACCATG
GABI-LB	genotyping T-DNA lines	CCCATTTGGACGTGAATGTAG
mur3_F	genotyping MUR3.1	CGAAGATCGGCATTTGACTC
mur3_RM	genotyping MUR3.1	CATCCTCGGGGATGAATATC
SALK-LBa1	genotyping T-DNA lines	TGGTTCACGTAGTGGGCCATCG
axy3.2-LP	genotyping *axy3.2*	CATGGCAAAATCAAGATTTGC
axy3.2-RP	genotyping *axy3.2*	GTTCAAGAAATCTCCGGATCC
Ski3	genotyping T-DNA lines	TGATCCATGTAGATTTCCCGGACATGAA
AS-18	sequencing/genotyping CSLC4	GCTACTTACGGCTGCTGGTA
AS-19	sequencing/genotyping XXT2	GTGGGATGTAAACCCTGTG
AS-20	sequencing/genotyping *P. pastoris* constructs	ACAGGAAACAGCTATGACCA
AS-21	sequencing/genotyping *P. pastoris* constructs	ATGTGTGAGAATGGACCTGT
AS-22	sequencing/genotyping ZeoR	CCCAACTTGAACTGAGGAAC
AS-23	sequencing *P. pastoris* constructs	ATCTAATCTAAGGGCGGTGT
AS-35	sequencing *P. pastoris* constructs	GTGGACAGCAAATGGGTACT
AS-36	sequencing CSLC4	GCGAGAAGTCTACGAACAGTC
AS-37	sequencing CSLC4	CAAGATATCCGCTTGGAAA
AS-38	sequencing XXT2	TGGCTTTTGAACTACCTTGG
AS-39	sequencing *P. pastoris* constructs	AAGAGGAGCAACGAAGAAAG
AS-28	attB1-AOX1 for Gateway	GGGGACAAGTTTGTACAAAAAAGCAGGCTcgctcattccaattccttc
AS-29	attB4-AOX1 for Gateway	GGGGACAACTTTGTATAGAAAAGTTGGGTGctcgtttcgaataattagttg
AS-30	attB4r-CSLC4/attB4r-XXT2 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAATGGCTAGCAT
AS-31	attB3r-CSLC4 for Gateway	GGGGACAACTTTATTATACAAAGTTGTCTAAGACATTTGTTCTCCAA
AS-32	attB3r-XXT2 for Gateway	GGGGACAACTTTATTATACAAAGTTGTTCAACCACTAGCAACCTT
AS-33	attB3-Markerless for Gateway	GGGGACAACTTTGTATAATAAAGTTGATTGACACGTCCGACGG
AS-34	attB2-Markerless for Gateway	GGGGACCACTTTGTACAAGAAAGCTGGGTAAGTGGATCCCCCTGTTGG
AS-48	sequencing MUR3	TGCTGGATTCGATATTGCTC
AS-49	sequencing XLT2	GAAGCAGTGTGACAATCTGA
AS-70	attB4r-Forward for Gateway	GGGGACAACTTTTCTATACAAAG
AS-71	attB3r-Reverse for Gateway	GGGGACAACTTTATTATACAAAG
AS-72	attB4r-CSLC4F/attB4r-XXT2 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAA
AS-73	attB3r-XXT2R for Gateway	GGGGACAACTTTATTATACAAAGTTGTTCAACCACTAG
AS-74	genotyping excision of Markerless cassette	AGAAAGCTGGGTAAGTGGATC
AS-75	sequencing CSLC4	CCCCTAACTCCGTTGTTGTT
AS-76	sequencing XXT2	AGAAAGATCCAGCGTGCTTT
AS-77	sequencing *P. pastoris* constructs	ACAGGTCCATTCTCACACAT
AS-90	attB1-XLT2 for Gateway	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGCTTCCTGTCTCAAATCC
AS-91	attB2-XLT2 for Gateway	GGGGACCACTTTGTACAAGAAAGCTGGGTtctccatttgtaccatttctcttg
AS-92	cloning XLT2 for complementation	GGTACCCCGGTAAGAGTGTTATCAGC
AS-93	cloning XLT2 for complementation	CCCGGGAGGTGAACTTATCTCCCGTT
AS-97	RT-PCR for XLT2	GGGATTTTCGCAGGTCTAAA
AS-98	RT-PCR for XLT2	TTGCTATACCGTTCCAACAC
TH-01	sequencing XLT2	TATCCTTCGCAAGACCCTTC
AS-103	sequencing XLT2	GTCCGGTTTGGGATAATTGT
AS-104	sequencing XLT2	TCACATACCACTATGGCACT
AS-105	sequencing XLT2	AGACTGAACACAAACCGATC
AS-106	sequencing XLT2	TTTCCCAAGTTTGTAACGCT
AS-107	sequencing XLT2	AGCATGTGCTATTTGACTCC
AS-108	sequencing XLT2	GAGAACCACACAAGAACACA
AS-111	genotyping/sequencing UGD	TTCGTATTTGACGGCAGAAA
AS-113	genotyping/sequencing UXS	ACACAGAGGATGATCCACAT
AS-114	sequencing CSLC4	AACCAAGCCATATAGGACCA
AS-115	sequencing UGD	AGGACTTACTGACATCTGCA

AS-120	sequencing XLT2	AGTCTCGATGTTGATAGGCT
AS-121	sequencing XLT2	ACACAGCAATAAAATCGCTAC
AS-122	genotyping *P. pastoris* (actin control)	AGGTTCCCACTTATTTCCCA
AS-123	genotyping *P. pastoris* (actin control)	CCTTCAATCCTGGGTTCATC
AS-124	genotyping XLT2 complementation	AAGCAACGTGATCCCTAAAG
AS-125	genotyping XLT2 complementation	TACACTTGGCACAACACAAT
AS-133	genotyping CS1007414 (MUR3 T-DNA)	ATTGACGGTGATGAAGACCAG
AS-134	genotyping CS1007414 (MUR3 T-DNA)	CAGCAAGGGAAGAATCATTTG
AS-137	cloning Sl09g064470 (SlMUR3)	ATGAGGCGCCGTTCTCCG
AS-138	cloning Sl09g064470 (SlMUR3)	TTATTGTTGTCCTTGTTCATTTTTC
AS-139	sequencing/genotyping (SlMUR3)	ATACTGGATGGTATGCGACT
AS-140	cloning Sl02g092840	ATGCTGTCACCTTACTCCG
AS-141	cloning Sl02g092840	TCAACTGCCCATATCCACG
AS-142	sequencing and genotyping Sl02g092840	AGCGTGGTATCGTACAAATC
AS-143	cloning Sl07g044960 (XST1)	ATGTTGCCATCTGAAAATTCTTC
AS-144	cloning Sl07g044960 (XST1)	CTATAGTTTTTGTTGCTTGAATC
AS-145	sequencing/genotyping Sl07g044960 (XST1)	CAATCGTACCGGAGAATCTC
AS-146	cloning Sl07g049610 (XST2)	ATGCTGCAGCTGAATCTCT
AS-147	cloning Sl07g049610 (XST2)	TTACGAATCAAAATACTGATCTC
AS-148	sequencing Sl07g049610 (XST2)	TTGGTACTGGACTGACATGT
AS-154	attB4r-TmNST1 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAatgaccgagggccaaact
AS-155	attB3r-TmNST1 for Gateway	GGGGACAACTTTATTATACAAAGTTGtcatgcatgcagatcctt
AS-156	attB4r-TmNST2 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAatggattcggagaagaaaca
AS-157	attB3r-TmNST2 for Gateway	GGGGACAACTTTATTATACAAAGTTGctaagctttattttctccgag
AS-158	attB4r-TmNST3 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAatgggtaagggtggatca
AS-159	attB3r-TmNST3 for Gateway	GGGGACAACTTTATTATACAAAGTTGttaattctgagattcatttctc
AS-160	attB4r-TmNST4 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAatgggaaagggcggctca
AS-161	attB3r-TmNST4 for Gateway	GGGGACAACTTTATTATACAAAGTTGttagttttgagaatcatttctc
AS-162	attB4r-UGD for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAATGGAACAGAAG
AS-163	attB3r-UGD for Gateway	GGGGACAACTTTATTATACAAAGTTGTCAAGCGACGGCTGGCATATC
AS-164	attB4r-UXS for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAATGGATTACAAAGACGACGATG
AS-165	attB3r-UXS for Gateway	GGGGACAACTTTATTATACAAAGTTGTCATGATGAGTTTTTCGTGG
TBL27s-f	sequencing AXY4	acgcaagggaactgggtcagagacga
TBL27s-r	sequencing AXY4	agtaaaccgagacagtgacgttgtgtg
AS-166	sequencing AXY4	ATTATTGGAAGCTGGTGAGC
AS-167	sequencing AXY4	gtgtagactagcatttcactg
AS-168	sequencing AXY4	AGCAAAGCCAGAAAGAATGT
AS-169	sequencing AXY4	aaaaccattgaaccactcgtc
AS-170	sequencing AXY4	GCAGTCCGTTATACAACTGG
AS-171	sequencing AXY4	ACTACGACCGTATCAAAACG
AS-172	sequencing AXY4	AACGAATGCGATATACCGAG
AS-173	cloning TmNST2	atggattcggagaagaaaca
AS-174	cloning TmNST2	ctaagctttattttctccgag
AS-179	attB4r-TmXXT5 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAatgggtcaggagattttcac
AS-180	attB3r-TmXXT5 for Gateway	GGGGACAACTTTATTATACAAAGTTGctatgatccaacttttccaa
AS-181	sequencing TmXXT5	TTCTGTTCCGGAATTGTCAG
AS-184	construction of Gateway-compatible pORE E4	AATGAGCTCACAAGTTTGTACAAAAAAGCT
AS-185	construction of Gateway-compatible pORE E4	AATGATATCACCACTTTGTACAAGAAAGC
AS-186	genotyping, sequencing TmNST1	TTCTGGGTATACTTGTCGCT
AS-187	genotyping, sequencing TmNST2	TAGCAGTATTCTGCAACGTG
AS-188	genotyping, sequencing TmNST3	TTAACCTGGTTGGGTATGGT
AS-189	genotyping, sequencing TmNST4	AGACTTCTGCATTGACCATG
AS-205	sequencing *P. pastoris* constructs	GTCGGACGTGTCAATcaac
AS-208	sequencing plant overexpression constructs	GGGAGAATTCGTCGACTTTG
AS-209	sequencing plant overexpression constructs	CGTAGCGGATAACAATTTCAC
AS-216	cloning ZeoR cassette attB3 for Gateway	GGGGACAACTTTGTATAATAAAGTTGGCCGCCAGCTTTCTAGA
AS-218	genotyping, sequencing TmXXT5	ATCCATCCACCAAATCCACT

AS-219	attB4r-SLC35B4 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAatgcgcccggccttggc
AS-220	attB3r-SLC35B4 for Gateway	GGGGACAACTTTATTATACAAAGTTGtcagttcttcttgctgtcct
AS-221	attB4r-SLC35C1 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAatgaacagggcgcctctg
AS-222	attB3r-SLC35C1 for Gateway	GGGGACAACTTTATTATACAAAGTTGtcacaccctgatagcactc
AS-223	attB4r-SLC35D1 for Gateway	GGGGACAACTTTTCTATACAAAGTTGTTCAAACAAAatggcggaagttcatagac
AS-224	attB3r-SLC35D1 for Gateway	GGGGACAACTTTATTATACAAAGTTGtcacactgctccttttccc
AS-260	P3r to P5r to make pDONR221 P4r-P5r	/5Phos/AACTTTGTATAcaAAAGTTGAACGAGAAAC
AS-261	P3r to P5r to make pDONR221 P4r-P5r	/5Phos/GGCATTATAAAAAAGCATTGCTCATCAATT
AS-262	P1 to P5 to make pDONR221 P5-P3r	/5Phos/GCCAACTTTGTAtAcAAAAGtTGAACGAGAAACGT
AS-263	P1 to P5 to make pDONR221 P5-P3r	/5Phos/ATTATAAAAAAGCATTGCTCATCAATTTGTTGC
AS-264	sequencing AXY8	acaacgaatacaatcagcgt
AS-265	sequencing AXY8	AAACTCCCTTGAGAAGTCCA
AS-266	sequencing AXY8	gtttgtggcctctaaagGTT
AS-267	sequencing AXY8	CCAATGGGAGCAGTAAACAT
AS-268	sequencing AXY8	ATGGACAAGgtatggacctc
AS-269	sequencing AXY8	gacctagttgtctgctccta
AS-270	sequencing AXY8	atgcactcaactctgtttgg
AS-273	attB5r-CSLC4 for Gateway (no stop)	GGGGACAACTTTTGTATACAAAGTTGTAGACATTTGTTCTCCAATTAGGT
AS-274	attB5r-XXT2 for Gateway (no stop)	GGGGACAACTTTTGTATACAAAGTTGTACCACTAGCAACCTTTACGG
AS-277	attB5-GFP for Gateway	GGGGACAACTTTGTATACAAAAGTTGgtgtgagcaagggcgaggag
AS-278	attB3r-GFP for Gateway	GGGGACAACTTTATTATACAAAGTTGTttacttgtacagctcgtccat
AS-279	genotyping SLC35B4	ATTTGGGAAACACTCCAAGG
AS-280	genotyping SLC35D1	ATTGGAATGGTCTTTGGTGG
AS-285	SALK_077065-LP	TTTCATTCCTCTCATGGCATC
AS-286	SALK_077065-RP	TAGGATCACCTCCAAATGCAG
AS-287	SALK_124729-LP	CCCACAGAAGTGCTCTCACTC
AS-288	SALK_124729-RP	TTATCGGGTAAGCAATTCGTG
AS-289	SALK_090612-LP	TGGGTCAGACTTTTGCATTTC
AS-290	SALK_090612-RP	TTGAATCTTGTTTTGGGTTCG
AS-291	SALK_065776C-LP	ACTGTGAGTGCGTTCTGTGTG
AS-292	SALK_065776C-RP	CATGGTGAAGGTCTTTGGAAC
AS-329	sequencing AXY3	TGTTAATTTTCCTTAAACCGCA
AS-330	sequencing AXY3	AGCAAAGACCAATCTCTCGT
AS-331	sequencing AXY3	CCAAACCAAGTGAGATCGAA
AS-332	sequencing AXY3	GAGTCATGTCGAACATGTGG
AS-333	sequencing AXY3	CTACGAAGCGCATATGACTG
AS-334	sequencing AXY3	CCCAAACTGTAATGCGTGAA
AS-337	sequencing AXY9	CCCAAGACTTAAGACCGTGA
AS-338	sequencing AXY9	GATGCTCTTCTTTAGTGGCG
AS-339	attB1-AXY9 for Gateway	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGCTAGGAGCGATTCATTTG
AS-340	attB2-AXY9 for Gateway (no stop)	GGGGACCACTTTGTACAAGAAAGCTGGGTTAAAGATTGATGAGATTCAATAAG
AS-343	attB1-entcup2 for Gateway	GGGGACAAGTTTGTACAAAAAAGCAGGCTgggatcttctgcaagca
AS-369	sequencing plant GFP constructs	CACAATCCCACTATCCTTCG
AS-370	sequencing plant GFP constructs	ATGACTCTGCGGTCTATGAT
AS-379	sequencing plant overexpression constructs	CAAACTCTCTCGAACCTTCC
AS-380	sequencing plant overexpression constructs	GGAAAAGCCAAGTAACCACT
AS-401	sequencing plant overexpression constructs	ACGGGGGACTCTAGAGTT
AS-402	sequencing plant overexpression constructs	GGGAAATTCGAGCTCTAA
AS-425	attB3r-entcup2 for Gateway	GGGGACAACTTTATTATACAAAGTTGtccggtgggttttgagg
AS-428	attB3-AXY9 for Gateway	GGGGACAACTTTGTATAATAAAGTTGATGCTAGGAGCGATTCATTTG
AS-429	attB2-AXY9 for Gateway	GGGGACCACTTTGTACAAGAAAGCTGGGTATCATAAAGATTGATGAGATTCAATAAG
AS-432	sequencing plant overexpression constructs	ATAACAATTTCACACAGGAA
AS-438	sequencing plant overexpression constructs	CTCTGTCTTGTGAAATGAA
AS-440	sequencing AXY9	TTGATACAAGAGATACGATG
AS-447	genotyping XST2	GGATCGAGTTTGCTCTACAT
AS-449	SlMUR3 for qRT-PCR	TTATGCCGATCCTCGCTCTAAA
AS-450	SlMUR3 for qRT-PCR	ATTTCCAGCTGAGCTCCTCAAT

AS-451	Sl02g092840 Forward	TATAGACCCGGAAGCAGTGAGA
AS-452	Sl02g092840 Reverse	CCTCCACTTGGCCTTGCATA
AS-453	XST1 for qRT-PCR	TCGCCATCGGAGGATTATTCAG
AS-454	XST1 for qRT-PCR	TCATTTTTCTCACGTCGTCTTTGG
AS-455	XST2 for qRT-PCR	TGCAAGTTTTGCAACAGATTCAT
AS-456	XST2 for qRT-PCR	TACCTGGAGTTGCGTAAACAAA
AS-457	PTB for qRT-PCR	GATGAGCCGAACCGTATTCTCT
AS-458	PTB for qRT-PCR	AGTGACGAGCTTCTCGACAAAT
mur1-s1	sequencing MUR1	GTTGAGTACAAAATAGTAATTGCTTC
mur1-r	sequencing MUR1	GACCTAAGAATGCTTCACATGC
mur1-s2	sequencing MUR1	TTTGTTTGGACAGTTCATAAAATGAT
mur1-s3	sequencing MUR1	TTATCCTGAAATTGACTAAAATACCC
mur1-s4	sequencing MUR1	CAAAGCTGATTCCACCGTCGT
mur1-s5	sequencing MUR1	TGAAACTCCACTACGCCGATCTC
mur1-s6	sequencing MUR1	TGGCCGTACCGTCAAGTATTACC
mur1-s7	sequencing MUR1	GCCGTGGTGAGAATTTCGTGA
mur1-s8	sequencing MUR1	GAGGAAGGACACACAGTGGAAGAGT
mur1-s9	sequencing MUR1	TTGAGCTTGCTAAGAGGGAGAAAGT
mur1-s10	sequencing MUR1	AGTATACTAAATTGTTTCCCTCCAA

13 Appendix 5: Miscellaneous data.

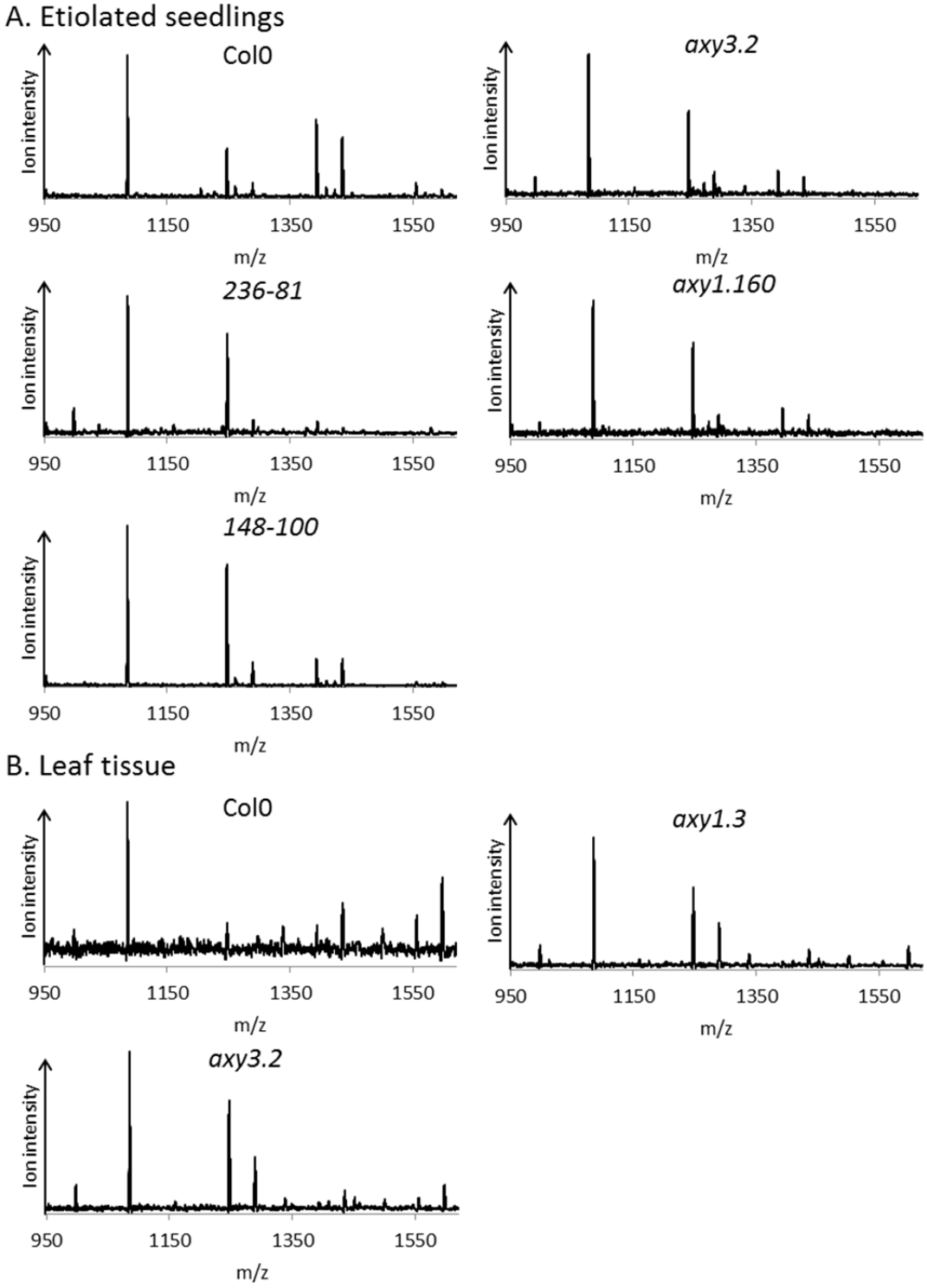


Figure 13-1. XyG OLIMP of *axy3*-like mutants. XyG OLIMP performed on AIR material from (A) etiolated seedlings and (B) leaves. The *axy3*-like phenotype is observed in both tissues by a

decreased relative abundance of the fucosylated oligosaccharides XXFG and XXFG with mass to charge ratios of 1393 and 1435 respectively.

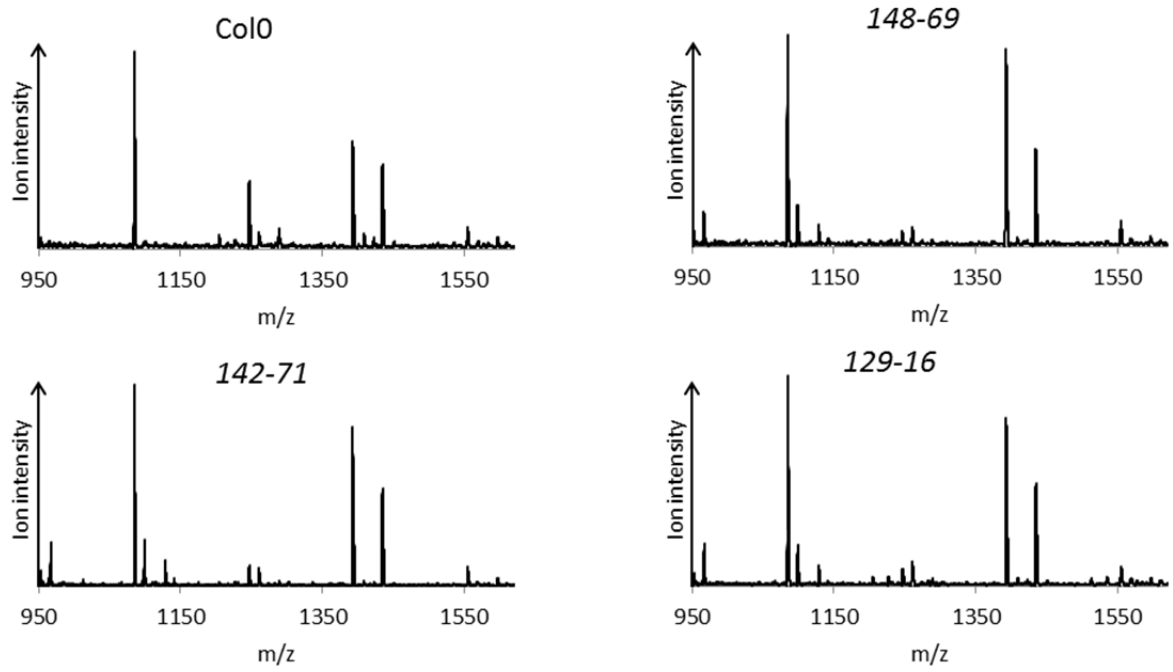


Figure 13-2. MALDI-TOF MS analysis of XEG-released material from *axy8*-like mutants. XyG OLIMP was performed on AIR material prepared from etiolated seedlings. Mutants 148-69, 142-71 and 129-16 displayed the *axy8* phenotype as observed by increased relative abundance of XXFG and XXFG as well as novel oligosaccharides GFG and XFG (m/z 967 and 1099) (Gunl et al. 2011).

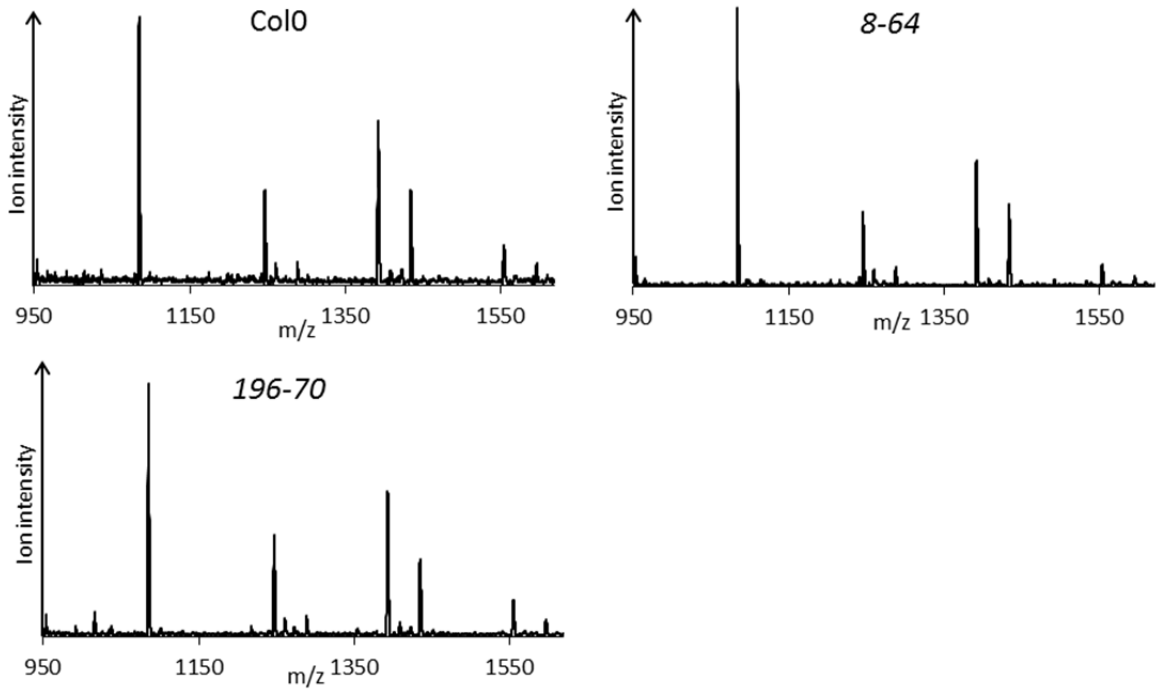


Figure 13-3. XyG OLIMP of candidate mutants *196-70* and *8-64*. OLIMP was performing on AIR prepared from etiolated seedlings using the XEG enzyme. Mutant phenotypes were not observed for any of the sixteen plants examined per line. Representative spectra are shown above.

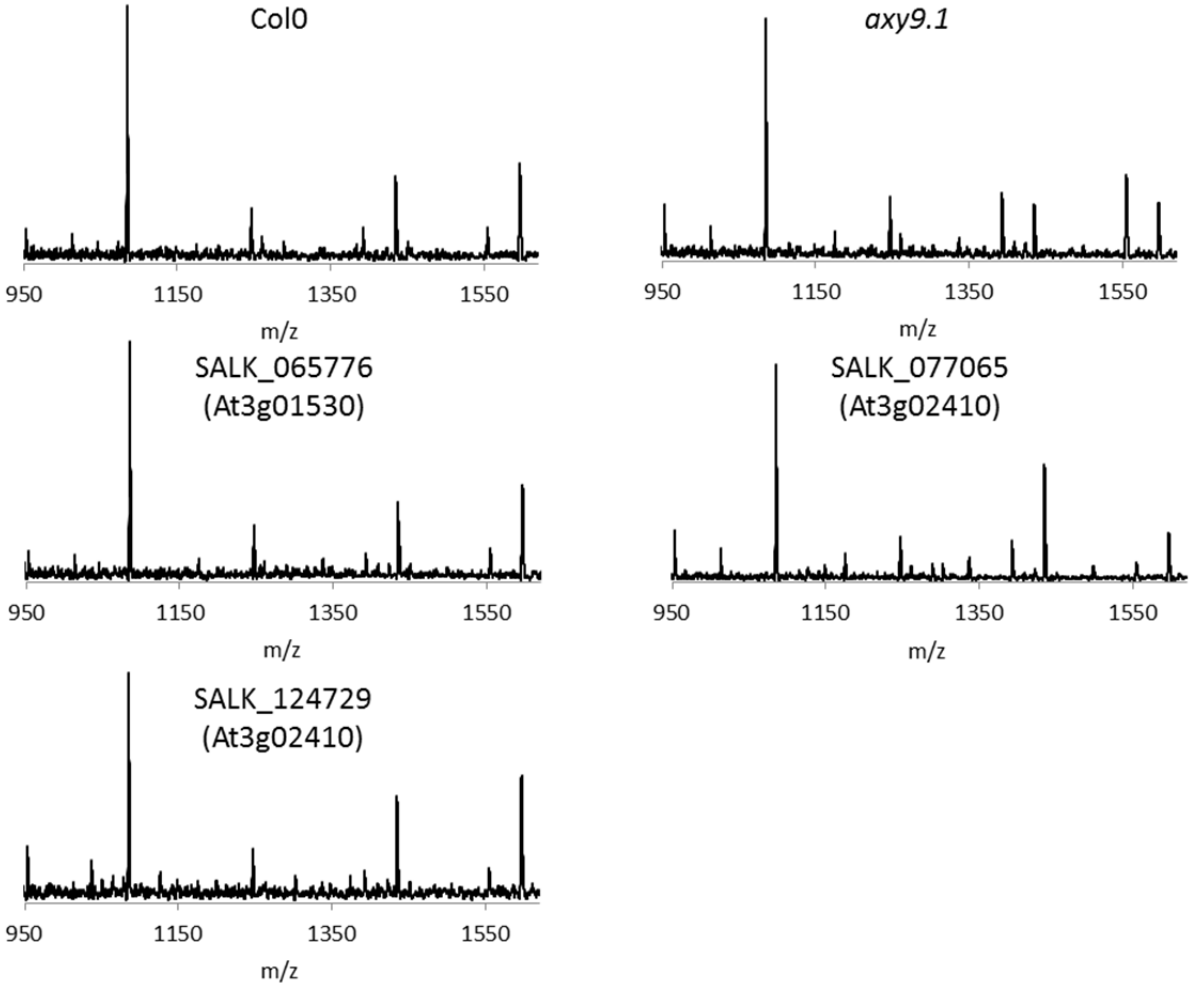


Figure 13-4. XyG OLIMP of candidate *axy9* alleles. Homozygous T-DNA insertional lines were obtained for several candidate genes (two for At3g02410 and one for At3g01530). These lines did not display a reduced XyG acetylation phenotype based on XyG OLIMP of leaf tissue. A T-DNA line with a disruption in At3g03210 did show reduced XyG acetylation (Figure 2-19).

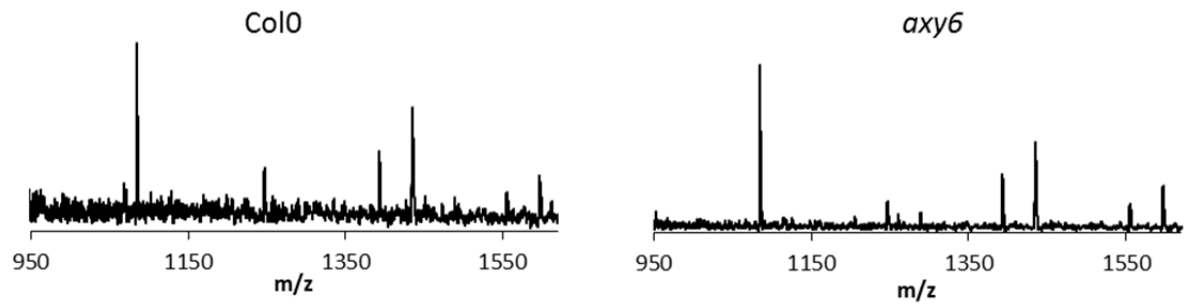


Figure 13-5. XyG OLIMP of *axy6* leaves. A XyG OLIMP phenotype was not observed in *axy6* leaf tissue.

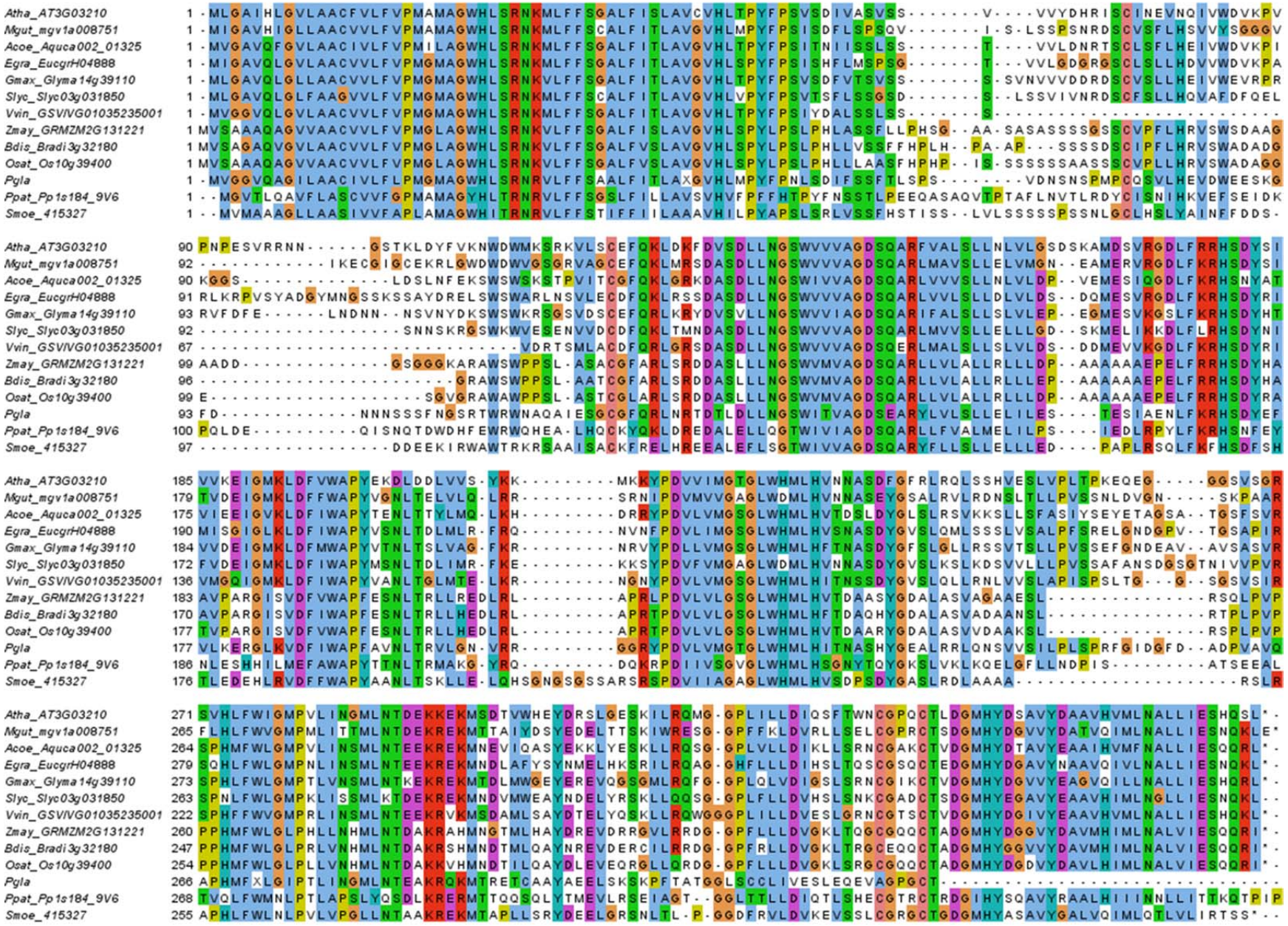

Atha_AT3G03210
Mgut_mgv1a008751
Acoe_Aquca002_01325
Egra_EucgrH04888
Gmax_Glyma14g39110
Slyc_Slyc03g031850
Vvin_GSVIVG01035235001
Zmay_GRMZM2G131221
Bdis_Bradi3g32180
Osat_Os10g39400
Pgla
Ppat_Pp1s184_9V6
Smoe_415327

Figure 13-6. AXY9 protein sequence alignment. Select plant AXY9 sequences were obtained (www.phytozome.org, NCBI Sequence Read Archive) and aligned. The alignment shows many conserved residues, notably two predicted N-terminal transmembrane domains and the GDS motif.

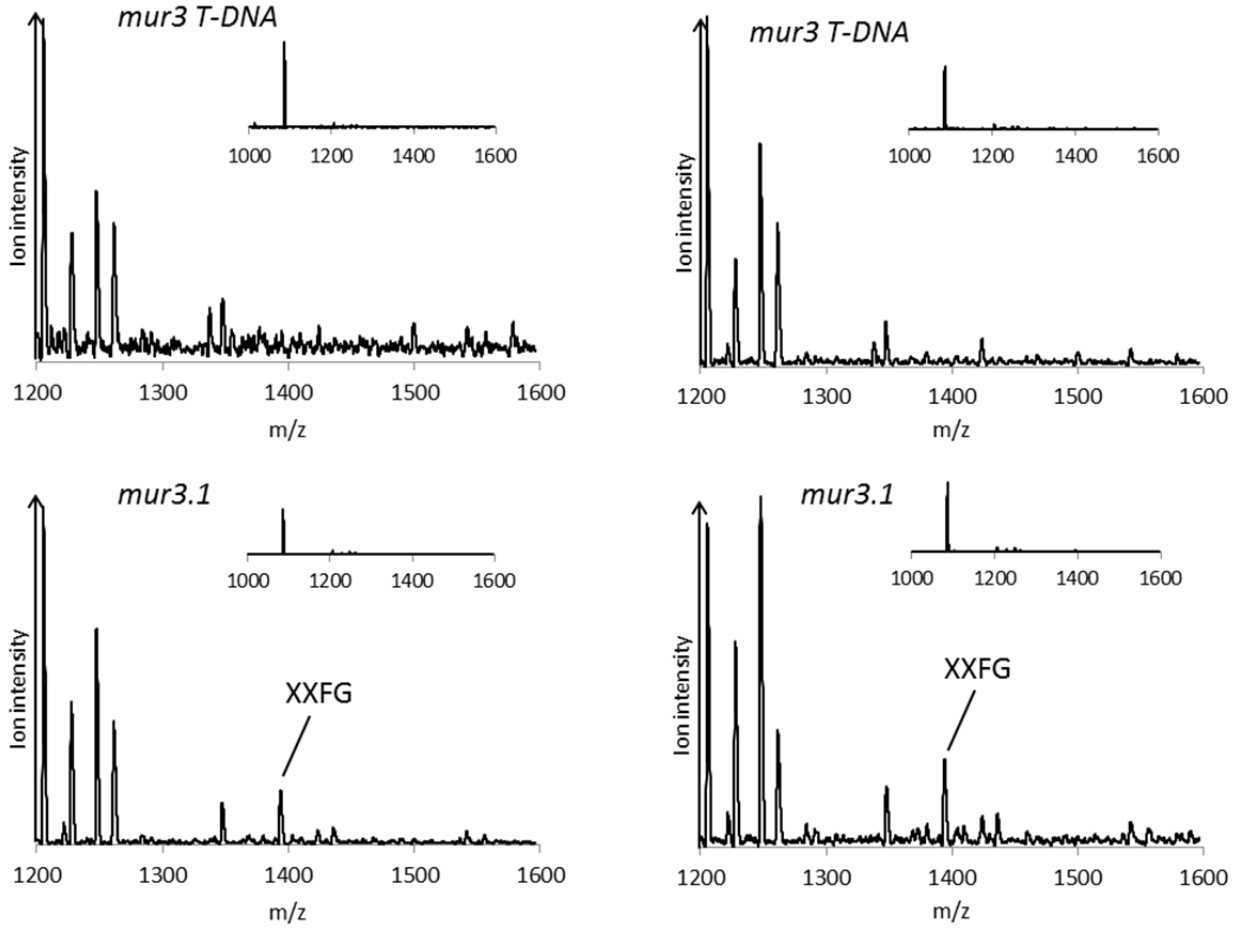


Figure 13-7. XyG OLIMP of *mur3.1* and a *mur3 T-DNA* allele. Cell wall material was isolated from etiolated seedlings of *mur3.1* and *mur3 T-DNA* plants and digested with XEG. The released oligosaccharides were analyzed by MALDI-TOF MS. An ion with an m/z of 1393, consistent with XXFG, was observed in *mur3.1* plants but not in *mur3 T-DNA* material. This suggests that *mur3.1* is a weak allele and that the galactosyltransferase of MUR3 in this line is not completely eliminated. Two replicates are shown for each genotype. The insets show the spectra from 1000-1600 m/z, whereas the larger graphs show the spectra from 1200-1600 m/z.