

**Functional characterisation of genes
involved in cell wall metabolism from the
flowering plant *Arabidopsis thaliana***

Issariya Chairam

Imperial College London, Department of Life Sciences

A thesis submitted for the degree of Doctor of Philosophy of

Imperial College London

Statement of Originality

I declare the work described in this thesis to be original work of my own. The work was carried out in the Department of Life Science of Imperial College London, under the supervision of Dr Thorsten Hamann. This thesis has not previously been submitted for examination and contains no material previously published or written by another person, except where indicated and acknowledged.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Acknowledgments

I gratefully acknowledge the financial support from National Science and Technology Development Agency (NSTDA) via a Thai Government Scholarship.

I thank people in the Plant and Microbial Science (PMS) section with whom I got to meet and discuss for all technical support and useful advices. Thanks to all the friends here that make SAF such a lovely place to work. Thanks to people in Dr David Leaks and Bishops labs for their help on *Pichia* system and molecular cloning. I would like to thank Dr Bart Feys, Dr David Leak and Dr John Rossiter for their helpful suggestions and advices on my MPhil transfer.

Special thanks to my supervisor, Dr Thorsten Hamann, for his dedicated supervision, his support and his patience to teach me to be a scientist. For the inspiration and all the helpful support and fruitful discussion are acknowledged to everybody in Hamann Lab; Dr Priyadharshini Madhou for her useful advices since I have done my MRes project; Dr Lucinda Denness for her helpful guidance and all the scientific, technical and personal support; Dr Alexandra Wormit for her dedicated help and guidance in science and also in mental and physical strength; Dr Izabela Szostkiewicz for her help with the protein works and her lovely friendship. Thanks to Lars for sharing all his interesting stories. Thanks to my little Joe for being awesome in the lab and making science fun. Thanks to Luis and Vin for the great friendships, especially Luis for your warmly helping hands. Obrigado! Thanks to David and the family; Maria the Super Mum, Claire and Christine for their kind moral support, for making me feel super special and for lighting up everyday with your lovely sweetness. I am so lucky I found you.

The most important people that I would like to thank are my lovely amazing family: Mummy, Dad and my two sisters: Reo-Farng and Meena for their unconditional loves and all their supports. Thanks for the patient and time, the opportunities and freedom you give me, the chances for me to learn from the outside world, with all your protection while letting me experience and live this life, the life you gave me.

Abstract

The plant cell wall is a complex and dynamic structure, involved in interaction with the environment, cell-cell interactions as well as plant growth and development. Although structure and architecture of the walls have been studied extensively, knowledge regarding the mechanisms required to generate the complexity and control dynamic changes of the walls is limited. In order to improve our understanding, this study employs a combination of genetic, molecular and biochemical methods to characterize the functions of three candidate genes, *At1g65610*, *At3g44990* and *At1g07260*. They encode a putative membrane-bound endo-1,4- β -glucanase (KORRIGAN2), a putative xyloglucan endo-transferase/hydrolase (XTH31) and a putative UDP-glycosyltransferase (UGT71C3) from *Arabidopsis thaliana*. The expression of the candidate genes changes in response to treatment with the cellulose biosynthesis inhibitor isoxaben. Plants with loss of gene activity of the target genes show only subtle growth phenotypes under standard growth conditions. However, cell wall analysis of stems from mature plants found distinct differences between the mutant plants and their wild type controls. In addition, my research showed that during early stage phenotypes, germination, root growth and isoxaben hypersensitivity phenotypes were observed in *korrigan2* and *xth31* seedlings. Interestingly, *ugt71c3* seedlings exhibited early germination phenotypes but no effects on root growth. In parallel, it was found that all mutant seedlings tested produce more ectopic lignin upon isoxaben treatment than wild type controls. These results show that the genes characterized are involved in cell wall metabolism and are required for the growth as well as the response to cell wall damage simulated by isoxaben treatment.

Table of contents

Abstract.....	3
List of figures.....	8
List of tables	19
Abbreviations	20
Chapter 1 Introduction.....	23
1.1 General introduction.....	23
1.2 Architecture of the plant cell wall	24
1.2.1 Nucleotide sugar inter-conversion enzyme family.....	25
1.2.2 Cellulose	27
1.2.3 Hemicelluloses.....	29
1.2.4 Pectins	30
1.2.5 Lignin	33
1.2.6 Minor components of the walls.....	34
1.3 Cell wall integrity maintenance in yeast and plants.....	35
1.4 Enzymes modifying plant cell wall	39
1.5 Glycosyl hydrolases (GHs).....	40
1.6 Selection of candidate genes for functional analysis	41
1.7 Aims of the project.....	45
Chapter 2 Materials and methods.....	47
2.1 Bioinformatics.....	47
2.2 Materials and methods for characterisation target genes	47
2.2.1 Plant materials and growth conditions.....	47
2.2.2 Isolation of T-DNA insertion lines.....	47
2.2.3 Seed sterilization and media preparation for plant growth	48
2.2.4 Isolation of homozygous insertion lines	48
2.2.4.1 DNA extraction.....	48
2.2.4.2 RNA isolation	49
2.2.4.3 Polymerase chain reaction (PCR) - based genotyping.....	49
2.2.4.4 Reverse transcriptase PCR (RT-PCR).....	49
2.2.4.5 Agarose gel electrophoresis and DNA sequencing.....	49
2.2.5 Phenotypic characterization of the mutants	50
2.2.5.1 Germination assay	50

2.2.5.2	Growth stage based analysis.....	50
2.2.5.3	Root growth assay	51
2.2.5.4	Phloroglucinol staining for lignin deposition analysis	52
2.2.5.5	Cell wall analysis	52
2.3	Molecular cloning for heterologous expression	53
2.3.1	<i>Escherichia coli</i> strains employed	53
2.3.2	Molecular cloning procedures	53
2.3.3	<i>E. coli</i> transformation competent cells preparation.....	54
2.3.3.1	Electro-competent cells.....	54
2.3.3.2	Chemically competent cells	54
2.3.4	Transformation methods	55
2.3.4.1	Electroporation.....	55
2.3.4.2	Heat-shock transformation	55
2.3.5	Protein expression and purification.....	55
2.3.5.1	Conditions for expression of the recombinant proteins	55
2.3.5.2	SDS-PAGE	56
2.3.5.3	Immunoblotting.....	56
2.3.5.4	His-tagged protein purification using Ni-TED column	57
Chapter 3	Bioinformatics analysis	58
3.1	Functional analysis of candidate genes and proteins	58
3.1.1	Adressing possible functional redundancies.....	62
3.1.2	Analysis of native signal peptides	63
3.2	Bioinformatics based analysis of <i>KORRIGAN2</i>	64
3.3	Bioinformatics based analysis of <i>XTH31</i>	65
3.4	Bioinformatics based analysis of <i>UGT71C3</i>	66
3.5	Discussion and summary.....	68
3.5.1	Functional analysis of candidate genes and proteins.....	68
3.5.2	Bioinformatics based analysis of <i>KORRIGAN2</i>	69
3.5.3	Bioinformatics based analysis of <i>XTH31</i>	71
3.5.4	Bioinformatics based analysis of <i>UGT71C3</i>	73
Chapter 4	Functional characterization of <i>KORRIGAN2</i>	75
4.1	Glycosyl hydrolase family 9	75
4.2	<i>Arabidopsis</i> membrane bound EGASEs	76
4.3	Heterologous expression of <i>KORRIGAN2</i>	77
4.4	Isolation of homozygous T-DNA insertion plants.....	81
4.5	Analysis of gene expression in the <i>kor2</i> alleles.....	83

4.6 Phenotypic characterization of <i>kor2</i> plants	83
4.6.1 Characterisation of germination of mutant and wild type plants	86
4.6.2 Root growth of <i>kor2</i> mutants and wild type seedlings	87
4.6.3 Lignin deposition in seedlings upon exposure to isoxaben	89
4.6.4 Cell wall in <i>kor2</i> and wild type plants.....	93
4.7 Discussion and summary	95
4.7.1 Heterologous expression of KORRIGAN2	95
4.7.2 Isolation of homozygous T-DNA insertion lines for <i>KORRIGAN2</i>	96
4.7.3 Analysis of <i>KOR2</i> gene expression in the insertion lines.....	98
4.7.4 Phenotypic characterisation of the mutant plants.....	99
Chapter 5 Functional characterization of <i>XTH31</i>	104
5.1 Plant xyloglucan endrotransglycosylase/ hydrolases (XTHs).....	104
5.2 <i>Arabidopsis</i> XTHs.....	105
5.3 Heterologous expression of <i>XTH31</i>	105
5.4 Isolation of homozygous T-DNA insertion plants.....	109
5.5 Analysis of <i>XTH31</i> expression in the T-DNA insertion lines	110
5.6 Phenotypic characterization of the mutant plants	111
5.6.1 Germination assay.....	113
5.6.2 Root growth and isoxaben response phenotypes	114
5.6.3 Lignin deposition after isoxaben treatment	115
5.6.4 Analysis of cell wall composition in mutant and wild type plants.....	119
5.7 Discussion and summary	121
5.7.1 Heterologous expression of <i>XTH31</i>	121
5.7.2 Isolation of the T-DNA insertion lines for <i>xth31</i> mutations.....	123
5.7.3 Analysis of <i>XTH31</i> expression disruption using RT-PCR	123
5.7.4 Phenotypic characterisation of <i>xth31</i> mutants.....	124
Chapter 6 Functional characterisation of <i>UGT71C3</i>	128
6.1 Functions of plant UDP-glycosyltransferases (UGTs)	128
6.2 UGTs in <i>Arabidopsis</i>.....	128
6.3 Heterologous expression of <i>UGT71C3</i>	131
6.4 Isolation of homozygous insertion lines	135
6.5 Analysis of <i>UGT71C3</i> expression using RT-PCR.....	135
6.6 Phenotypic characterization of the mutant plants	136
6.6.1 Characterisation of germination rates of <i>ugt71c3</i>	138
6.6.2 Analysis of root growth of mutant plants	139

6.6.3	Lignin deposition upon exposure to isoxaben.....	140
6.6.4	Analysis of cell wall composition in mutant and wild type plants.....	145
6.7	Discussion and summary.....	147
6.7.1	Heterologous expression of UGT71C3	147
6.7.2	Isolation of homozygous insertion lines	148
6.7.3	Analysis of <i>UGT71C3</i> expression using RT-PCR	148
6.7.4	Phenotypic characterisation of the mutants	149
Chapter 7	General discussion and summary	153
7.1	Phenotypic analysis and lignification of mutant seedlings.....	153
7.2	KORRIGAN 2	158
7.3	Xyloglucan transglycosylase/ hydrolase 31.....	160
7.4	UDP- glycosyl transferase 71C3.....	162
7.5	Perspectives and possible contributions to facilitated lignocellulosic biofuel production.....	165
Appendices		168
Appendix 1	Primers: annealing sites and sequences.....	168
Appendix 2	Results from Prosite analysis for the candidates	171
Appendix 3	Preliminary studies of eight candidate genes.....	173
Appendix 3.1	T-DNA insertion mutants for the candidate genes	173
Appendix 3.2	Heterologous expression of target enzymes in <i>Pichia</i>	174
References		176

List of figures

- Figure 1** Simplified schematic structure of primary plant cell wall where cellulose and matrix components are showing. Cellulose microfibrils (purple rods) are synthesised by cellulose synthase complexes embedded in the plasma membrane. Matrix polysaccharides like hemicelluloses and pectins (consisting of glycans synthesised in the golgi apparatus) are transported to the cell surface by vesicles. The schematic diagram shows hemicelluloses, xyloglucan and arabinoxylan tethering the cellulose microfibrils together. Pectins are shown on the right. Notably, other components such as lignin and structural proteins are not shown (adapted from Cosgrove, 2005). 25
- Figure 2** An overview of nucleotide sugar inter conversion reactions in plants, showing related carbohydrate metabolism pathways (redrawn from Reiter, 2008). 27
- Figure 3** A) The structure of cellulose consists of β -(1,4) linked-D-glucan molecules. B) The intra- and inter-chain hydrogen bond network in cellulose (orange lines) (adapted from Cosgrove, 2005; Himmel et al., 2007). 28
- Figure 4** A schematic diagram of the CESA complex. (A) an individual CESA protein, (B) a rosette subunit containing six CESA proteins, (C) the rosette producing a cellulose microfibril and (D) a cross-view of the rosette consisting of six subunits with three different CESAs (blue, brown and orange) (adapted from Cosgrove, 2005). 28
- Figure 5** Hemicellulose structures, (A) xyloglucan and (B) arabinoxylan. Both types of hemicelluloses have β -(1,4)-D-glucan backbones but differ in their branches. Xyloglucan has xylose branches while arabinoxylan can have arabinose (green), glucuronic acid (red) and ferulic acid ester (FAE) branches (adapted from Cosgrove, 2005). 29
- Figure 6** Schematic structure of pectin. Pectin consists of four different types of polysaccharides as their structures shown (from left to right), RhamnogalacturonanII (RG-II), Homogalacturonan (HG), Xylogalacturonan (XG) and RhamnogalacturonanI (RG-I). RGII is a complex domain, containing a galacturonic acid backbone (GalA, yellow hexagons), which can form dimers through borate esters. HG is a linear chain of GalA residues, where the carboxyl groups of some residues are esterified. Similar to HG, XG consists of the GalA backbone with esterification but differs in the xylose branches. RG-I consists of a backbone with the repeating units of GalA and rhamnose (adapted from Harholt et al., 2010). 31
- Figure 7** The structure of the RG-II boron diester in the walls (adapted from Cosgrove, 2005). ..32
- Figure 8** The structure of homogalacturonan (HG) as a linear polymer of α -(1,4)-linked galacturonic acid (GalA) residues. Methylesterification at the C-6 and O-acetylation at the C-2 or C-3 of the GalA residues are shown on the left. The egg-box model of calcium crosslinking in HG is shown on the right (adapted from Caffall & Mohnen, 2009). 32
- Figure 9** Schematic pathway of monolignol biosynthesis showing enzymes involved in the pathway are phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-(hydroxy) cinnamoyl CoA ligase (4CL), hydroxycinnamoyl CoA shikimate:quininate

hydroxycinnamoyl transferase (HCT), *p*-coumaroylshikimate 3'-hydroxylase (C3'H),
 caffeoyl CoA *O*-methyl transferase (cCoAOMT), (hydroxy)cinnamoyl CoA reductase (CCR),
 ferulic acid/ coniferaldehyde/ coniferyl alcohol 5-hydroxylase (F5H); caffeic acid/ 5-
 hydroxyferulic acid *O*-methyltransferase (COMT), (hydroxy)cinnamyl alcohol
 dehydrogenase (CAD), peroxidase (PER) and laccase (LAC) (adapted from Weng & Chapple,
 2010). 34

Figure 10 Schematic diagram for cell wall integrity pathway in yeast. Under cell wall stress, the
 signals are initiated at the plasma membrane. There are five putative mechanosensors
 which recruit Rom1/2 GEFs to the membrane and stimulate nucleotide exchange in Rho1
 G-protein. Activated Rho1 will trigger the effectors in mitogen activated protein kinase
 cascade whose targets are transcriptional factors for genes encode cell wall modifying
 enzymes (OCH1=mannosyltransferase, FSK1= β -1,3-glucan synthase). Under hypo-
 osmolarity condition, a part of the high osmolarity pathway including the Sln1 sensor
 kinase is autophosphorylated. It could phosphorylate a histidine phosphotransfer protein,
 Ypd1 can phosphorylate Skn1 transcription factor. In calcium pathway, under stress
 conditions, Mpk1 can activate a Mid1/Cch1 calcium channel. This can lead to an active form
 of calcium-calmodulin dependent protein phosphatase, calcineurin that will dephosphorylate
 Crz1 transcription factor into active form, allowing transcription of Fsk1 gene. (redrawn
 from Levin, 2005) 36

Figure 11 Candidate gene expression profiles from *Arabidopsis* seedlings isoxaben or mock
 treated (represented by bold and dash lines). The Y-axis represents the expression level,
 while the X-axis represents the treatment period (hours). Samples with isoxaben treatment
 are labelled as 'IXB' (expression data derived from Hamann *et al.* 2009). 43

Figure 12 Expression profiles of the candidate genes derived from Genevestigator. (A) shows the
 expression profiles of the genes in different types of plant organs and tissues, while (B)
 shows the profiles from different plant developmental stages. Heat map represented
 percentage of expression potential for candidate genes were derived based on microarray
 expression profiles deposited on GeneVestigator database. 44

Figure 13 Candidate gene expression profiles of *Arabidopsis* seedlings in response to isoxaben
 and mock treatment (represented by bold and dash lines). The Y-axis represents the
 expression level, while the X-axis represents the period of treatment (hours). Samples with
 isoxaben treatment are labelled as 'IXB' (expression data derived from Hamann, 2009). .46

Figure 14 Schematic depiction for the arrangement of *Arabidopsis* seeds sown on a round plate
 of ½ MS with 1% sucrose supplement for germination assay. 50

Figure 15 An example for the arrangement of plant pots grown for growth stage based analysis.
51

Figure 16 Schematic depiction for the arrangement of *Arabidopsis* seedlings grown on a square
 plate of ½ MS with 1% sucrose supplement for root growth assay.....51

Figure 17 Schematic depiction for the assembly of the protein gel and the membrane for
 Western blotting (From Bio-Rad Mini Trans-Blot® Electrophoretic Transfer Cell manual). .57

- Figure 18** Predicted secondary structure of putative KOR2. KOR2aa = amino acid sequences, KOR2Profsec = predicted secondary structure for KOR2 with residues in α -helix regions (red H) and extended β -strand regions (blue E) colour coded. KOR2PHDhtm = profile network prediction Heidelberg where the helical trans-membrane region (purple M) is shown. Highlighted in yellow is a cell adhesion motif. Low complexity regions (LCRs) are highlighted in grey. Two glycosyl hydrolase family 9 active signature regions are highlighted in green (www.predictprotein.org; www.uniprot.org) 64
- Figure 19** Predicted secondary structure of putative XTH31. XTH31aa = amino acid sequences where low complexity regions (LCRs) are highlighted in grey. XTH31Profsec = predicted secondary structure for XTH31 that the residues in α XTH31P regions (red H) and extended β -strand regions (blue E) are predicted. XTH31PHDhtm = profile network prediction Heidelberg where the helical trans-membrane region (purple M) is shown. XTHDBbond = predicted disulfide bonds. Predicted β -glucanase active region is highlighted in green (www.predictprotein.org; www.uniprot.org) 66
- Figure 20** The predicted secondary structure of UGT71C3. UGT71C3 aa = amino acid sequences where low complexity regions (LCRs) are highlighted in grey. UGT71C3Profsec = predicted secondary structure for UGT71C3 with residues α -helix regions (red H) and extended β -strand regions (blue E) predicted. UGT71C3PHDhtm = profile network prediction Heidelberg where the helical trans-membrane region is highlighted with (purple M) is shown; and the UDP-glycosyltransferase signature region being highlighted in green (www.predictprotein.org; www.uniprot.org).....67
- Figure 21** Structural variations among 3 subclasses of the plant GH9 family. The schematic illustrates the modular structure: cytoplasmic domain (dark grey), transmembrane domain (white), signal sequence (black), GH9 catalytic domain (light grey), linker region (thick black line), carbohydrate binding module (CBM) (hexagon). Structural subclasses are represented by GH9 members from tomato; SlGH9A1 (Subclass A, U78526), SlGH9B1 (Subclass B, U13054) and SlGH9C1 (Subclass C, AAD08699) (adapted from Urbanowicz et al., 2007^b). 75
- Figure 22** Schematic model showing plant GH9 enzymes that are membrane bound (Class A), associated with the cell wall through a CBM (Class C) or more freely soluble (Class B) (adapted from Lopez-Casado et al., 2008). 76
- Figure 23** PCR to amplify extracellular domain for KOR2 flanking with *Bam*HI and *Not*I restriction sites using Phusion® DNA polymerase. DNA ladder shown as kb.....78
- Figure 24** Colony PCR to isolate DH5 α harbouring pCR2.1KOR2 (1-9) and DH5 α with empty plasmid (10). Size of DNA ladder fragments shown in kb.....78
- Figure 25** *Bam*HI-*Not*I double digest for pCR2.1KOR2 construct. Size of DNA ladder fragments shown in kb 78
- Figure 26** Colony PCR for DH5 α pET28KOR2 (1-11) and DH5 α harbouring pET28 empty plasmid (12) are shown. Size of DNA ladder fragments shown in kb 79

Figure 27 <i>Bam</i> HI- <i>Not</i> I double digest for pET28KOR2 construct. Size of DNA ladder fragments shown in kb	79
Figure 28 SDS-PAGE analysis for KOR2 expressed in <i>E.coli</i> strain BL21 (A), SoluBL21™ (B) and BL21 (DE3) (C). Four different clones were tested; clone1 = lane 1 and 2, clone2 = lane 3 and 4, clone3 = lane 5 and 6 and clone4 = lane 7 and 8. Samples from 0 hour (lane 1, 3, 5 and 7) and 6 hours (2, 4, 6 and 8) of IPTG induction are shown. Size of protein marker shown in kD.....	80
Figure 29 Immuno dot blot for KOR2 expressed under 0, 0.25, 0.5 and 1 mM IPTG induction for 0, 3, 6 and 18 hours.....	81
Figure 30 SDS-PAGE for KOR2 after purification for His-tagged proteins. Shown above are cell lysate (1), soluble proteins (2), Flow-through (3), 1 th Wash (4), 2 nd Wash (5), 1 st Elution (6), 2 nd Elution (7), 3 rd Elution (8) and 4 th Elution (9). Size of protein marker shown in kD.....	81
Figure 31 PCR-based screening to isolate homozygous <i>kor2</i> mutant plants. Gene specific fragments and T-DNA fragments were amplified from 10 independent mutant plants (1-10) for each T-DNA insertion line and Col-0 (11). Size of DNA ladder fragments shown in kb ..	82
Figure 32 Semi-quantitative RT-PCR results for <i>KORRIGAN 2</i> expression in Col-0 and <i>kor2</i> seedlings. <i>UBIQUITIN 10</i> was used as a control. Col-0 shown in (1, 2), <i>kor2-1</i> in (3, 4), <i>kor2-2</i> (5, 6) and <i>kor2-3</i> (7, 8) with seedlings in 1, 3, 5, 7 mock and 2, 4, 6, 8 isoxaben-treated. Size of DNA ladder fragments shown in bp	83
Figure 33 Rosette diameter measurement for <i>kor2</i> mutant plants that showed rosette growth complete stage. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average rosette diameter calculated from three biological replicates where the data was analysed using sixteen individual plants ± SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).....	84
Figure 34 Biomass measurement for four week-old <i>kor2</i> mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight of either fresh or dried materials and calculated from three biological replicates where the data was analysed using eight individual plants ± SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).	85
Figure 35 Biomass measurement for senescence <i>kor2</i> mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight calculated from three biological replicates where the data was analysed using eight individual plants ± SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).....	85
Figure 36 Results from germination assays for <i>kor2</i> mutants and Col-0 wild type. Sterile seeds were put on ½ MS plates and stratified at 4°C for forty-eight hours before grown in standard conditions as described. Seeds with radicals emerged were counted as germinated seeds after 1-4 days after plates were put under standard conditions. The values represent	

average percentage of germination calculated from three biological replicates where the data was analysed using fifty individual seeds $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).....	86
Figure 37 Roots of seven day-old <i>kor2</i> mutants and Col-0 wild type seedlings grown vertically on $\frac{1}{2}$ MS plates were captured. Photos were analysed using ImageJ software. The values represent the average of root length of seedlings calculated from three biological replicates where the analysis was performed using at least twenty individual seedlings $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test, $P \leq 0.05$).....	87
Figure 38 Percentage of root length decreased up on isoxaben treatment. Roots of seven day-old <i>kor2</i> mutants and Col-0 wild type seedlings grown vertically on $\frac{1}{2}$ MS plates for four days before transferred to fresh $\frac{1}{2}$ MS plates in presence or absence of 1, 2 or 3 nM isoxaben were analysed. Photos were captured as described and analysed using ImageJ software. Percentage of root length decreased were calculated from four biological replicates where the analysis was performed using at least ten individual seedlings $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type under the same treatment (student t-test, $n \geq 40$ per line and condition, $P \leq 0.05$).	88
Figure 39 Phloroglucinol stained seven day-old <i>kor2</i> and Col-0 seedling roots upon 12 or 24 hour-isoxaben treatment. A) The representative results of the phloroglucinol staining. Scale bar is 0.5 mm, at least twenty seedling roots were analysed for each genotype/treatment. B) Percentage of the phloroglucinol staining for seven day-old <i>kor2</i> and Col-0 seedling roots after 12 or 24 hour-isoxaben treatment $\pm$ SE. * indicates significant difference (student t-test, $P \leq 0.05$).....	89
Figure 40 Phloroglucinol staining for lignin deposition in Col-0 and <i>kor2</i> seedling roots upon 0, 1, 2 and 3 nM isoxaben treatment for three days. Scale bar is 0.5 mm, for each genotype/treatment at least twenty seedlings were analysed.	91
Figure 41 Percentage of phloroglucinol staining for lignin deposition in Col-0 and <i>kor2</i> seedling roots treated with 1, 2 and 3 nM isoxaben for three days. The values represent lignification area of total root area in percentage where the data was analysed using at least twenty individual seedlingss $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).	92
Figure 42 Cellulose measurement from stems of mature plants of Col-0 wild type, <i>kor2-1</i> , <i>kor2-2</i> and <i>kor2-3</i> mutants. The values represent average cellulose content in μ g glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using at least four individual plants . * indicates significant difference of mean of the data set compared with Col-0 (student t-test, $P < 0.05$). Error bars are based on standard deviation.	93
Figure 43 Uronic acids measurement results in mature stems of Col-0, <i>kor2-1</i> , <i>kor2-2</i> and <i>kor2-3</i> plants. The values represent average uronic acid content in μ g uronic acid per mg dry weight of plant materials which were calculated from three biological replicates where the	

data was analysed using at least four individual plants. * indicates significant difference of mean of the data set compared with Col-0 (student t-test, P<0.05). Error bars are based on standard deviation.....	93
Figure 44 Neutral cell wall sugar measurements in stems from Col-0, <i>kor2-1</i> , <i>kor2-2</i> and <i>kor2-3</i> plants. The values represent average values of seven different neutral cell wall sugar contents in µg sugar per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using at least four individual plants. * indicates significant different of mean of the data set compared with Col-0 (student t-test, P<0.05). Error bars are based on standard deviation.....	94
Figure 45 Microarray expression data for <i>At5g49720</i> (<i>KOR1</i>) (red), <i>At1g65610</i> (<i>KOR2</i>) (purple) and <i>At4g24260</i> (<i>KOR3</i>) (yellow) under mock treatment (dash line) and isoxaben treatment (bold line). Data derived from (Hamann et al., 2009).....	99
Figure 46 Results from PCR reactions amplifying fragments encoding the extracellular domain of XTH31 flanked by <i>Bam</i> HI and <i>Not</i> I restriction sites using Phusion® DNA polymerase. Size of DNA ladder fragments shown in kb	106
Figure 47 Colony PCR to isolate DH5α harbouring pCR2.1XTH31 (1-10) and DH5α with empty plasmid (11). Size of DNA ladder fragments shown in kb.....	106
Figure 48 <i>Bam</i> HI- <i>Not</i> I double digest for pCR2.1XTH31 construct. Size of DNA ladder fragments shown in kb.	107
Figure 49 Colony PCR for DH5α pET28XTH31 (1-12) and DH5α harbouring pET28 empty plasmid (13) are shown. Size of DNA ladder fragments shown in kb.....	107
Figure 50 <i>Bam</i> HI- <i>Not</i> I double digest for pET28XTH31 construct. Size of DNA ladder fragments shown in kb.	108
Figure 51 SDS-PAGE analysis for XTH31 expressed in <i>E.coli</i> strain BL21 (A), SoluBL21™ (B) and BL21 (DE3) (C). 4 different clones were tested; clone1 = lane 1 and 2, clone2 = lane 3 and 4, clone3 = lane 5 and 6 and clone4 = lane 7 and 8. Samples from 0 hour (lane 1, 3, 5 and 7) and 6 hours (2, 4, 6 and 8) of IPTG induction were shown. Putative XTH31 protein fragments are indicated with an arrow. Size of protein marker shown in kD	108
Figure 52 SDS-PAGE for XTH31 after purification for His-tagged proteins. Shown above are cell lysate (1), soluble proteins (2) Flow-through (3), 1 th Wash (4), 4 th Wash (5), 1 st Elution (6), 2 nd Elution (7), 3 rd Elution (8) and 4 th Elution (9). Target proteins were indicated with the arrow. Size of protein marker shown in kD.....	109
Figure 53 PCR-based genotyping to isolate plants homozygous for insertions into <i>XTH31</i> . Gene specific fragments and T-DNA fragments were amplified. For <i>xth31-1</i> , PCR reactions from ten independent mutant plants were shown on the upper panels (1-10) and Col-0 (11). For <i>xth31-2</i> , PCR reactions from twelve independent mutant plants were shown on the lower panels (1-12) and Col-0 (13). Size of DNA ladder fragments shown in kb.	110
Figure 54 Semi-quantitative RT-PCR for <i>XTH31</i> from Col-0, <i>xth31-1</i> and <i>xth31-2</i> . Gene specific reactions (right panel) and <i>UBIQUITIN 10</i> was used as a control (left panel). Col-0 mock and isoxaben-treated samples are shown in (1) and (2), <i>xth31-1</i> mock- and isoxaben-treated	

ones shown in (3) and (4) and <i>xth31</i> -2 mock- and isoxaben-treated samples shown in (5) and (6). Size of DNA ladder fragments shown in bp.....	111
Figure 55 Rosette diameter measurement for <i>xth31</i> mutant plants that showed rosette growth complete stage. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average rosette diameter calculated from three biological replicates where the data was analysed using sixteen individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).....	112
Figure 56 Biomass measurement for four week-old <i>xth31</i> mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight of either fresh or dried materials and calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).	112
Figure 57 Biomass measurement for senescence <i>xth31</i> mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).	113
Figure 58 Results from germination assays for <i>xth31</i> mutants and Col-0 wild type. Sterile seeds were put on $\frac{1}{2}$ MS plates and stratified at 4°C for forty-eight hours before grown in standard conditions as described. Seeds with radicals emerged were counted as germinated seeds after 1-3 days after plates were put under standard conditions. The values represent average percentage of germination calculated from three biological replicates where the data was analysed using fifty individual seeds $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).....	114
Figure 59 Root of seven day-old <i>xth31</i> mutants and Col-0 wild type seedlings grown vertically on $\frac{1}{2}$ MS plates were captured and the photos were analysed using ImageJ software. The values represent the average of root length of seedlings calculated from three biological replicates where the analysis was performed using at least forty individual seedlings $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test, $P \leq 0.05$).	114
Figure 60 Impact of different isoxaben concentrations on <i>xth31</i> and Col-0 seedling root growth. Percentage of root length decreased up on isoxaben treatment. Roots of <i>xth31</i> mutants and Col-0 wild type seedlings grown vertically on $\frac{1}{2}$ MS plates for four days before transferred to fresh $\frac{1}{2}$ MS plates in presence or absence of 1, 2 or 3 nm isoxaben were analysed. Photos were captured as described and analysed using ImageJ software. Percentage of root length decreased were calculated from four biological replicates where the analysis was performed using at least ten individual seedlings $\pm$ SE. * indicates significant difference of	

mean between data sets compared to Col-0 wild type under the same treatment (student t-test, $n \geq 40$ per line and condition, $P \leq 0.05$).....	115
Figure 61 Phloroglucinol stained seven day-old <i>xth31</i> and Col-0 seedling root tips after isoxaben treatment for 12 or 24 hours. A) is the representative results of the phloroglucinol staining. Scale bar is 0.5 mm, for each genotype/treatment at least twenty seedling roots were analysed. B) Percentage of the phloroglucinol staining for seven day-old <i>xth31</i> and Col-0 seedling root tips after isoxaben treatment for 12 or 24 hours.....	116
Figure 62 Lignin deposition in root tips of Col-0 and <i>xth31</i> seedlings after treatment with 0, 1, 2 and 3 nM isoxaben for four days. Scale bar is 0.5 mm.	117
Figure 63 Percentage of phloroglucinol staining for lignin deposition in Col-0 and <i>xth31</i> seedling roots treated with 0, 1, 2 and 3 nM isoxaben for three days. A) percentage of phloroglucinol stained for vascular bundle tissue. B) percentage of phloroglucinol stained for other areas along the roots. The values represent average stained area measured via ImageJ and calculated from three biological replicates where the data was analysed using at least twenty individual plants $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).	118
Figure 64 Cellulose measurements from stems of mature Col-0 and <i>xth31</i> plants. The values represent average cellulose content in μ g glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants. * indicates significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are based on SD values ($n = 3$).	119
Figure 65 Uronic acids measurement of material from stems of mature Col-0 and <i>xth31</i> plants. The values represent average uronic acid content in μ g uronic acid per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants. * indicates significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are based on SD values ($n = 3$).	120
Figure 66 Neutral cell wall sugar measurements in stems from mature Col-0 and <i>xth31</i> plants. The values represent average values of different seven neutral cell wall sugar contents in μ g sugar per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants. * indicates significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are based on SD values ($n = 3$).	120
Figure 67 Microarray expression data for <i>At3g44990</i> (<i>XTH31</i>) (orange) and <i>At2g36870</i> (<i>XTH32</i>) (red) under mock treatment (dash line) and isoxaben treatment (bold line). Data derived from Hamann et al. (2009).	124
Figure 68 Phylogenetic analysis of the <i>Arabidopsis</i> UGT super family, fourteen distinct groups are categorised based on amino acid similarity. UGT71C3 encoded by <i>At1g07260</i> resides in group E. (Ross et al., 2001)	129

Figure 69 The chemical structures of two hydroxycoumarins, esculetin and scopoletin, the substrates used for functional UGT assays and the three possible products deriving from glycosylation at different hydroxyl positions (Lim et al., 2003 ^a).....	130
Figure 70 Results from PCR reaction amplifying <i>UGT71C3</i> and introducing <i>Bam</i> HI and <i>Not</i> I restriction sites using Phusion® DNA polymerase. Size of DNA ladder fragments shown in kb.	131
Figure 71 Results from colony PCR to identify pCR2.1UGT71C3 positive colonies. Size of DNA ladder fragments shown in kb, expected fragment size: 1.5kb.....	132
Figure 72 Results of <i>Bam</i> HI- <i>Not</i> I double digest to generate the pCR2.1UGT71C3 construct. Size of DNA ladder fragments shown in kb.....	132
Figure 73 Results from a colony PCR to identify clone containing pET28UGT71C3 (1-10) are shown. Size of DNA ladder fragments shown in kb, expected fragment size: 1.5 kb.....	132
Figure 74 Results from a <i>Bam</i> HI and <i>Not</i> I double digest of pET28UGT71C3 (1-3) and pET28 empty plasmid (4) are shown. Size of DNA ladder fragments shown in kb, expected fragment size: 1.5 kb.	133
Figure 75 Immuno dot blot for UGT71C3 expressed under 0, 0.25, 0.5 and 1 mM IPTG induction for 0, 3, 6 and 18 hours.	133
Figure 76 SDS-PAGE analysis of three different <i>E.coli</i> strains BL21 (A), SoluBL21™ (B) and BL21 (DE3) (C). Four different clones were tested per strain; clone1 = lane 1 and 2, clone2 = lane 3 and 4, clone3 = lane 5 and 6 and clone4 = lane 7 and 8. Samples from 0 hour (lane 1, 3, 5 and 7) and 6 hours (2, 4, 6 and 8) of IPTG induction were shown. Protein markers shown in kD.....	134
Figure 77 SDS-PAGE for UGT71C3 after purification for His-tagged proteins. Shown above are cell lysate (1), Flow-through (2), 1 st Wash (3), 4 th Wash (4), 1 st Elution (5), 2 nd Elution (6), 3 rd Elution (7) and 4 th Elution (8). Target proteins are indicated with the arrow. Size of protein marker shown in kD.....	134
Figure 78 PCR-based screening to isolate homozygous <i>ugt71c3</i> plants. Gene specific fragments and T-DNA fragments were amplified from eleven independent mutant plants (2-12) and Col-0 (13). Size of DNA ladder fragments shown in kb.....	135
Figure 79 Semi-quantitative RT-PCR for <i>UGT71C3</i> from Col-0 and the insertion lines (left). <i>UBIQUITIN 10</i> was used as loading control (right). Col-0 mock- and isoxaben-treated shown in (1) and (2), <i>ugt71c3-1</i> mock and isoxaben treated shown in (3) and (4) and <i>ugt71c3-2</i> mock and isoxaben treated are shown in (5) and (6). Size of DNA ladder fragments shown in bp.	135
Figure 80 Rosette diameter measurement for <i>kor2</i> mutant plants that showed rosette growth complete stage. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average rosette diameter calculated from three biological replicates where the data was analysed using sixteen individual plants ± SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).....	137

- Figure 81** Biomass measurement for four week-old *ugt71-1*, *ugt71-2* and Col-0 wild type plants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight of either fresh or dried materials and calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference for the data set compare with wild type (student t-test, $P < 0.05$). 137
- Figure 82** Biomass measurement for senescence *ugt71c3* mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$). 138
- Figure 83** Results from germination assays for *ugt71* mutants and Col-0 wild type. Sterile seeds were put on $\frac{1}{2}$ MS plates and stratified at 4°C for forty-eight hours before grown in standard conditions as described. Seeds with radicals emerged were counted as germinated seeds after 1-3 days after plates were put under standard conditions. The values represent average percentage of germination calculated from three biological replicates where the data was analysed using fifty-two individual seeds $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$). 139
- Figure 84** Root of seven day-old *ugt71* mutants and Col-0 wild type seedlings grown vertically on $\frac{1}{2}$ MS plates were captured and the photos were analysed using ImageJ software. The values represent the average of root length of seedlings calculated from three biological replicates where the analysis was performed using at least twenty individual seedlings $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test, $P \leq 0.05$). 139
- Figure 85** Percentage of root length decreased up on isoxaben treatment. Roots of *ugt71* mutants and Col-0 wild type seedlings grown vertically on $\frac{1}{2}$ MS plates for four days before transferred to fresh $\frac{1}{2}$ MS plates in presence or absence of 1, 2 or 3 nM isoxaben were analysed. Photos were captured as described and analysed using ImageJ software. Percentage of root length decreased were calculated from four biological replicates where the analysis was performed using at least ten individual seedlings $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type under the same treatment (student t-test, $n \geq 40$ per line and condition, $P \leq 0.05$). 140
- Figure 86** Phloroglucinol stained seven day-old *ugt71c3* and Col-0 seedling root tips after isoxaben treatment for 12 or 24 hours. A) is the representative results of the phloroglucinol staining. Scale bar is 0.5 mm, for each genotype/treatment at least twenty seedling roots were analysed. B) Percentage of the phloroglucinol staining for seven day-old *ugt71c3* and Col-0 seedling root tips after isoxaben treatment for 12 or 24 hours. 141
- Figure 87** Phloroglucinol staining for lignin deposition from Col-0 and *ugt71c3* mutant seedlings treated with 0, 1, 2 and 3 nM isoxaben. Scale bar is 0.5 mm. 143

- Figure 88** Percentage of phloroglucinol staining for lignin deposition in Col-0, *ugt71c3-1* and *ugt71c3-2* seedling roots treated with 0, 1, 2 and 3 nM isoxaben for three days. A) percentage of phloroglucinol stained for vascular bundle tissue. B) percentage of phloroglucinol stained for other areas along the roots. The values represent average stained area measured via ImageJ and calculated from three biological replicates where the data was analysed using at least twenty individual plants $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$). 144
- Figure 89** Cellulose measurement from stems of mature Col-0 and *ugt71c3*. The values represent average cellulose content in μg glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants for each biological replicate. * indicates statistical significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are SD values. 145
- Figure 90** Uronic acids measurements in stems from mature Col-0 and *ugt71c3* plants. The values represent average cellulose content in μg glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants for each biological replicate. * indicates statistical significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are SD values. 146
- Figure 91** Neutral cell wall sugar measurement from stems of mature Col-0 and *ugt71c3* plants. The values represent average cellulose content in μg glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants for each biological replicate. * indicates statistical significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are SD values. 146
- Figure 92** Microarray derived expression data for *UGT71C1* (*At2g29750*) (red), *UGT71C2* (*At2g29740*) (yellow), *UGT71C3* (*At1g07260*) (blue), *UGT71C4* (*At1g07250*) (pink) and *UGT71C5* (*At1g07240*) (grey) under mock treatment (dash line) and isoxaben treatment (bold line). Data derived from Hamann et al. (2009). 149
- Figure 93** Heat map showing expression patterns of genes belonging to *UGT71* clad C; *At2g29750* (*UGT71C1*), *At2g29740* (*UGT71C2*), *At1g07260* (*UGT71C3*), *At1g07250* (*UGT71C4*) and *At1g07240* (*UGT71C5*) in different plant organs. Data derived from Genevestigator. 150
- Figure 94** PCR to amplify AOX1 region using genomic DNA templates from *Pichia* clones, pPICZ α A_UGT71C3 (A, 1), pPICZ α A_76B1 (A, 2), Empty plasmid (A, 3), pPICZ α A_PME (B, 1), GS115 harbouring empty plasmid (B, 2), pPICZ α A_PLL17 (C, 1), pPICZ α A_PLL26 (C, 2), pPICZ α A_KOR2 (C, 3), pPICZ α B_GH9C2 (C, 4) and GS115 (C, 5). 175

List of tables

Table 1 Nucleotide sugars and their cell wall polysaccharide components (adapted from Reiter, 2008).	26
Table 2 Glycoside hydrolase enzyme families in <i>Arabidopsis</i> involved in cell wall metabolism (adapted from Minic & Jouanin, 2006; Minic, 2008)	41
Table 3 Initial selected genes with their putative protein functions.....	42
Table 4 Genetics of <i>E.coli</i> strains used in the project.....	53
Table 5 Structure of the candidate genes showing untranslated regions (light blue box), exons (blue box) and introns (solid blue line). Transcription orientation of each gene is indicated by arrow-shaped ends. Red triangles represent putative T-DNA insertion sites for selected insertion lines. (illustrated based on data derived from arabidopsis.org)	59
Table 6 Query enzymes with predicted conserved domains based on amino acid sequences. Amino acid sequences are represented in purple, a specific hit for which the query sequences displayed highest identities is shown in red, the multi-domains are shown in grey other non-specific hits and domains for superfamilies are in pale red and light blue. Plant secondary product glycosyltransferase (PSPG) domains are in green. (adapted from blast.ncbi.nlm.nih.gov)	60
Table 7 Overview of family members for the different candidate genes that show similar expression patterns with respect to developmental stages and plant organs based on transcript levels.....	62
Table 8 Prediction of signal peptides for candidate enzymes.	63
Table 9 Growth stage-based analysis results for the wild type and <i>kor2</i> mutants.....	84
Table 10 Growth stage-based analysis results for the wild type and <i>kor2</i> mutants.....	111
Table 11 Growth stage-based analysis results for the wild type and <i>ugt71c3</i> mutants.....	136
Table 12 Summarised observed phenotypes from <i>kor2</i> , <i>xth31</i> and <i>ugt71c3</i> mutants	154
Table 13 Other primers used throughout this study.....	170
Table 14 Predicted protein domains, families and functional sites for the candidates from Prosite database (adapted from Sigrist et al., 2010).	171
Table 15 Overview of T-DNA insertion lines for phenotypic study	173
Table 16 Codon usage analysis using Rare Codon Analysis Tools from GenScript.....	174

Abbreviations

4CL	4-(hydroxy) cinnamoyl CoA ligase
AA	amino acid
ABA	abscisic acid
ABRC	the Arabidopsis Biological Resource Centre
ACE	ACGT-containing element
AFM	atomic force microscopy
AGP	arabinogalactan protein
AGPs	arabinogalactan proteins
BSA	bovine serum albumin
BZIP	basic region/leucine zipper
C3'H	<i>p</i> -coumaroylshikimate 3'-hydroxylase
C4H	cinnamate 4-hydroxylase
Ca ²⁺	calcium ion
CaCl ₂	calcium chloride
CAD	(hydroxy)cinnamyl alcohol dehydrogenase
cAMP	cyclic adenosine monophosphate
CAZy	Carbohydrate-Active enZymes database
CBI	cellulose biosynthesis inhibitor
CBM	carbohydrate binding module
Cch	high-affinity calcium channel protein
CcoAOMT	caffeoyl CoA <i>O</i> -methyl transferase
CCR	(hydroxy)cinnamoyl CoA reductase
cDNA	complementary deoxyribonucleic acid
Cel	cellulase
CESA	<i>CELLULOSE-SYNTHASE A</i>
<i>cev</i>	<i>constitutive expression of VSP</i>
cGMP	cyclic guanosine monophosphate
CH ₃ CO ₂ K	potassium acetate
Col-0	Arabidopsis Columbia ecotype
COMT	caffeic acid/ 5-hydroxyferulic acid <i>O</i> -methyltransferase
COO ⁻	carboxylate ion
Crz	calcineurin-responsive transcription factor
CWD	cell wall damage
CWI	cell wall integrity
D-GLcA	D-glucuronic acid
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphate
DW	dry weight
EDTA	ethylenediaminetetraacetic acid
EGASE	endo-(1,4)-β-D-glucanases
<i>ein</i>	<i>ethylene insensitive</i>
<i>eli</i>	<i>ectopic lignification</i>
<i>etr</i>	<i>ethylene receptor</i>

EXT	extensin
F5H	ferulic acid/ coniferaldehyde/ coniferyl alcohol 5-hydroxylase
FAE	ferulic acid ester
FSK	Glucan synthase of cerevisiae
FTIR	Fourier transform infrared spectroscopy
Fuc	fucose
G-units	guaiacyl units
GA	gibberellic acid
Gal	galactose
GalA	galacturonic acid
GAUT	homogalacturonan α -1,4-galacturonosyltransferase
GAX	glucuronoarabinoxylans
gDNA	genomic deoxyribonucleic acid
GDP-Man	guanosine diphosphate mannose
GEF	guanosine nucleotide exchange factor
GH	glycosyl hydrolase
GRP	Glycine-rich protein
GTs	glycosyltransferases
GUS	β -glucuronidase
H-units	hydroxyphenyl units
HCl	hydrogen chloride
HCT	hydroxycinnamoyl CoA shikimate:quinate hydroxycinnamoyl transferase
HG	homogalacturonan
His	histidine
HRGPs	hydroxyproline-rich glycoproteins
HRP	horseradish peroxidase
IPTG	Isopropyl β -D-1-thiogalactopyranoside
JA	jasmonic acid
kb	kilobase
KCl	potassium chloride
kD	kilodalton
<i>kor</i>	<i>korrigan</i>
kV	kilovolt
LAC	laccase
LB	Lysogeny broth
LCRs	low complexity regions
MAPK	mitogen-activated protein kinases
MCA	MID1 COMPLEMENTING ACTIVITY
MgSO ₄	magnesium sulphate
MGT	macroside glycosyltransferase
MID	MATING PHEROMONE-INDUCED DEATH
MnCl ₂	Manganese (II) chloride
MOPs	3-(N-morpholino) propanesulfonic acid
Mpk	MITOGEN-ACTIVATED PROTEIN
MS	Murashige and Skoog medium
MTL1	MID2-LIKE 1
<i>mur</i>	<i>murus</i>
MYB	MYELOBLASTOSIS
NaCl	sodium chloride

NASC	the Nottingham Arabidopsis Stock Centre
NMR	nuclear magnetic resonance
PAL	phenylalanine ammonia lyase
PCR	polymerase chain reaction
PER	peroxidase
PKC	PROTEIN KINASE C
PME	pectin methylesterase
PSPG	plant secondary product glycosyltransferase
PttCel	<i>Populus tremula</i> × <i>Populus tremuloides</i> cellulase
PVDF	polyvinylidene fluoride
qPCR	quantitative real time polymerase chain reaction
RbCl	rubidium chloride
RBOH	RESPIRATORY BURST OXIDASE HOMOLOGUE
RG-I	rhamnogalacturonan-I
RG-II	rhamnogalacturonan-II
RGXT	1,3- α -D-xylosyltransferases
Rha	rhamnose
Rho	RAS HOMOLOG
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
ROM	RAS HOMOLOG (RHO) MULTICOPY SUPPRESSOR
ROS	reactive oxygen species
rpm	revolutions per minute
RT-PCR	reverse transcription polymerase chain reaction
S-units	syringyl units
SD	standard deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIGNAL	Salk Institute Genomic Analysis Laboratory
SKN	SUPPRESSOR OF KRE NULL
SLN	SYNTHETIC LETHAL OF N-END RULE
SOB	super optimal broth
SOC	super optimal broth with catabolite repression
T-DNA	transfer deoxyribonucleic acid
TBST	Tris-Buffered Saline with Tween 20
THE	THESEUS
T _m	melting temperature
UBQ10	Ubiquitin10
UDP	uridine diphosphate
UGT	UDP-glycosyltransferase
UTR	untranslated region
UXE	<i>UDP-D-xylose 4-epimerase</i>
VSP1	VEGETATIVE STORAGE PROTEIN 1
WSC1	SYNTHETIC LETHAL WITH GAP 1
XEH	xyloglucan endohydrolase
XET	xyloglucan endo-transglycosylase
XG	xyloglucan
XTH	xyloglucan endotransglucosylase-hydrolase
XXT	Xyloglucan Xylosyltransferases
YPD	TYROSINE (Y) PHOSPHATASE DEPENDENT

Chapter 1 Introduction

1.1 General introduction

The plant cell wall is similar to an exoskeleton that protects the cell from biotic and abiotic stresses. The wall plays important roles in cell morphogenesis, structural support and cell-cell interaction (Cosgrove, 2005; Minic & Jouanin, 2006). Cellulose, the backbone polysaccharide of the plant cell walls, is the most abundant polysaccharide component in biomass and could provide most of the raw material to generate bioenergy so called lignocellulosic biofuels or second generation biofuels (Fairley 2011; Schiermeier et al., 2008; Somerville et al., 2010). In order to manipulate the wall to be used for generating biofuels and food production, it is important to understand the architecture and function of each component, and also the control mechanisms regulating cell wall biosynthesis and metabolism. Using the available genetic and biochemical tools, basic knowledge about the mechanisms that regulate metabolism of the wall has been gained. In parallel with the study of architecture of the wall, studies on enzymes regulating cell wall structure have been performed. For example, Taylor et al. (2008) have reviewed the available knowledge regarding heterologous expression of glycosyl hydrolases in plants. Interestingly, many attempts have been made to express bacterial or fungal enzymes in plants; however, the strategies used need optimisation to generate recombinant enzymes that at the same time have only limited negative effects on growth and development of the host plants. To optimise the strategies of the heterologous expression, many prospects need to be concerned including getting the correct folding and post-translational modification of the heterologous enzymes in order to be able to express highly specific active forms of those enzymes. It has been shown that alterations in polysaccharide components of the walls could lead to an enhancement of productivity of bioenergy plants for biofuel production. For instance, transgenic plants have been produced with increased cellulose and decreased lignin content of lignocellulosic biomass making them more available to processing (Fornalé et al., 2011; Patten et al., 2010; Sticklen, 2008).

Plant cell walls are dynamic and complex. The walls are formed mainly from polysaccharides which consist of diverse types of monosaccharides linked by different linkages like, for example, β -(1-4)-linkage or ester linkages (Geshi et al., 2010; Popper & Fry, 2008; Somerville, 2006). Therefore, polysaccharide biosynthesis and remodelling of the wall require a large number of different enzymes (Shipp et al., 2008). Although studies of cell wall structure are quite extensive, the regulatory mechanisms governing plant cell wall biosynthesis are still not well understood (Hamann, 2012; Peaucelle et al., 2012; Somerville, 2006). Many enzymes involved in cell wall metabolism are not well characterised (Gille et al., 2009; Lopez-Casado et al., 2008). For example, there are very limited data regarding the large enzyme families encoding plant glycosyl hydrolases (GHs), and much of the available data is derived from microbial and fungal studies (Bauer et al., 2006; Lopez-Casado et al., 2008). Even in *Arabidopsis*, knowledge regarding cell wall biosynthetic enzymes is limited. For example many cell wall-modifying enzymes have been identified only based on sequence similarity using available genome sequence without functional characterisation of their biological roles in cell wall metabolism (Sticklen, 2008).

1.2 Architecture of the plant cell wall

Plant cell walls are categorised into two types (type I and II) based on their chemical composition. Type I cell walls are found in dicotyledons and gymnosperms while type II walls are only found in graminaceous plants (Cosgrove, 2005). In type I walls xyloglucans and pectins form about 35 % of the wall mass (Minic & Jouanin, 2006), whereas in type II walls the arabinoxylan content is higher and there are lower levels of pectins and xyloglucans. Additionally, the type II wall contains mixed linked β -D-glucans and glucuronoarabinoxylans (GAX), which are cross-linked into the cell wall network by dimerisation of ester molecules through feruolyl groups (Carpita, 1996).

The walls are categorised into two forms by their functions, so called primary and secondary cell walls. The primary wall is deposited during cell elongation, thus, it can be found in every cell types. Whereas, the secondary wall is formed after elongation and expansion has ceased, to strengthen

the wall. Therefore, it can be found mostly in specialised cells such as fibres, tracheids and vessel elements in the xylem, which require reinforcement to prevent cell wall collapse during water transport (Zhong & Ye, 2007). Biochemical components that form most of the primary wall are cellulose, hemicelluloses, pectins and structural proteins (Figure 1) (Cosgrove, 2005).

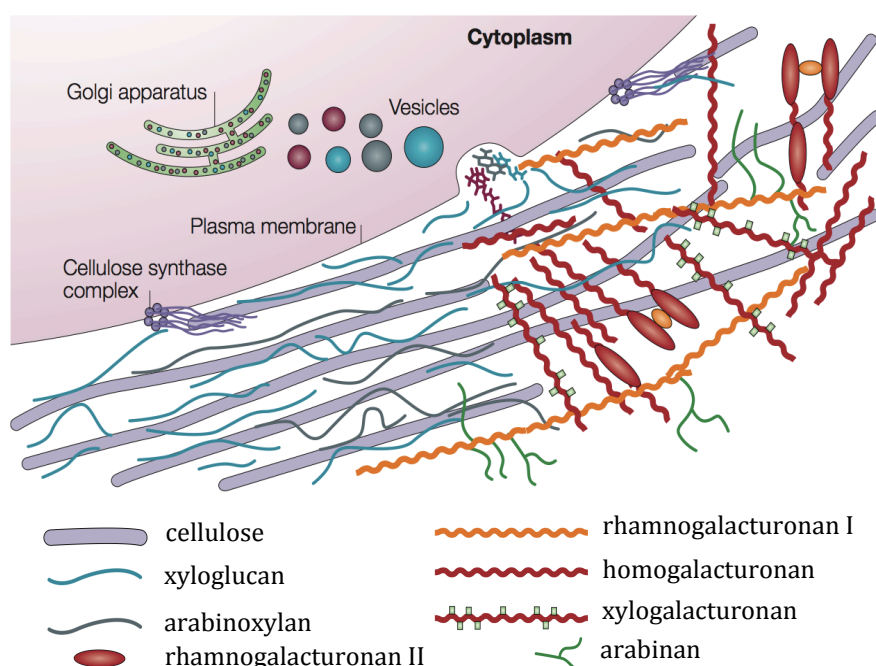


Figure 1 Simplified schematic structure of primary plant cell wall where cellulose and matrix components are showing. Cellulose microfibrils (purple rods) are synthesised by cellulose synthase complexes embedded in the plasma membrane. Matrix polysaccharides like hemicelluloses and pectins (consisting of glycans synthesised in the golgi apparatus) are transported to the cell surface by vesicles. The schematic diagram shows hemicelluloses, xyloglucan and arabinoxylan tethering the cellulose microfibrils together. Pectins are shown on the right. Notably, other components such as lignin and structural proteins are not shown (adapted from Cosgrove, 2005).

1.2.1 Nucleotide sugar inter-conversion enzyme family

During the dark reaction of photosynthesis, plants fix carbon dioxide and convert carbon molecules into glucose (Seifert, 2004). Most of the glucose molecules are incorporated into cell wall carbohydrates whereas the rest of the molecules are converted into glycosides, oligosaccharides, glycoproteins, glycolipids, sucrose and starch for different biological processes or simply storage (Reiter, 2008). Therefore, changes in the abundance of the nucleotide

sugars may affect structure, composition and integrity of the wall. Previous work has also shown that CWD causes changes in neutral cell wall sugars generated in these biosynthetic processes and that genes encoding the biosynthetic activities exhibit changes in transcript levels (Hamann et al., 2009). The nucleotide sugar precursors for cell wall polysaccharides are shown in Table 1. Most of the genes belonging to this enzyme family have been cloned and characterised (Reiter & Vanzin, 2001; Reiter, 2008; Seifert, 2004).

Table 1 Nucleotide sugars and their cell wall polysaccharide components (adapted from Reiter, 2008).

Nucleotide sugars	Polysaccharides
UDP- α -D-glucose	Cellulose, callose, mixed-linked glucans and xyloglucan
UDP- α -D-galactose	Rhamnogalacturonan-I, arabinogalactan proteins and xyloglucan
UDP- α -D-glucuronate	Glucurono(arabino)xylans and rhamnogalacturonan-II
UDP- α -D-galacturonate	Homogalacturonan, rhamnogalacturonan-I and -II
UDP- α -D-xylose	Xylans and xyloglucan
UDP- α -D-apiose	Rhamnogalacturonan-II
UDP- β -L-arabinose	Rhamnogalacturonan-I, arabinogalactan proteins and (glucurono)arabinoxylans
UDP- β -L-rhamnose	Rhamnogalacturonan-I and -II
GDP- α -D-glucose	Glucomannans
GDP- α -D-mannose	(Gluco)mannans
GDP- β -L-galactose	Rhamnogalacturonan-II
GDP- β -L-fucose	Rhamnogalacturonan-II and xyloglucan

As illustrated in the pathways in Figure 2, Fructose-6-Phosphate (D-Fructose-6-P) produced during photosynthesis is an essential precursor for further cell wall polysaccharide biosynthesis (Reiter, 2008). The D-Fructose-6-P molecules can be sequentially processed by phosphoisomerases and phosphomutases to produce glucose-1-phosphate and mannose-1-phosphate, which are converted to UDP-D-glucose (UDP-Glc) and GDP-D-mannose (GDP-Man) (Reiter and Vanzin, 2001; Reiter, 2008). By conversion of UDP-Glc and GDP-Man, monosaccharides and other nucleotide sugars are derived, while GDP-D-glucose is derived from D-glucose-1-phosphate (Reiter, 2008).

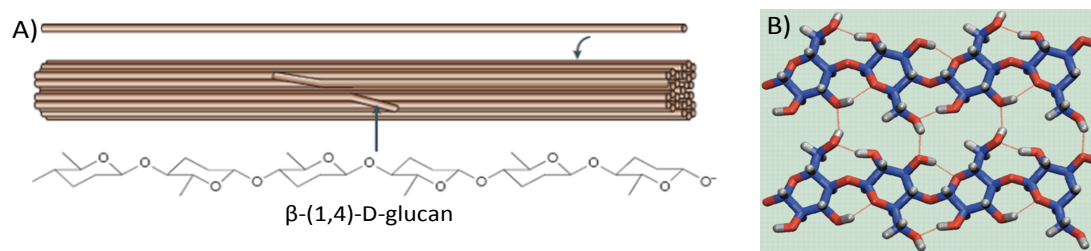


Figure 3 A) The structure of cellulose consists of β -(1,4) linked-D-glucan molecules. B) The intra- and inter-chain hydrogen bond network in cellulose (orange lines) (adapted from Cosgrove, 2005; Himmel et al., 2007).

Cellulose microfibrils are synthesised by the cellulose synthase complex (CESA), which is localised in the plasma membrane and forms hexameric rosettes (Figure 4). The *CESA* genes encode the main proteins forming the complex (Cosgrove, 2005). In *Arabidopsis*, the *CESA* gene family contains ten genes, which show different expression patterns. To form a functional cellulose synthase complex, three different CESA proteins are required and different sets of CESAs are involved in formation of the primary or secondary cell walls (Taylor et al., 2003). Results from genetic studies suggest that *CESA1*, *CESA3* and *CESA6* are required for primary cell wall biosynthesis, while *CESA4*, *CESA7* and *CESA8* are required for secondary cell wall biosynthesis (Atanassov et al., 2009; Carroll et al., 2012; Desprez et al., 2007).

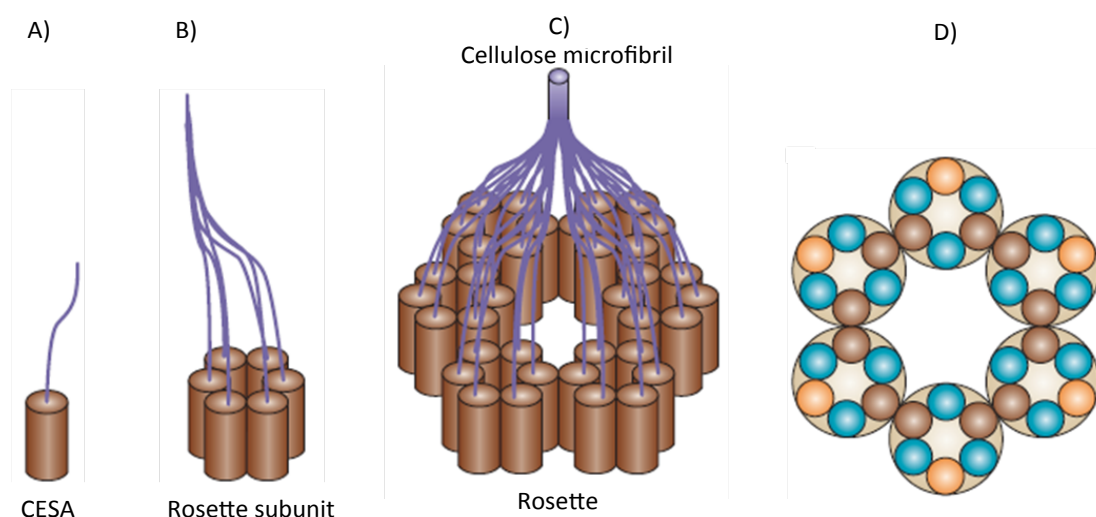


Figure 4 A schematic diagram of the CESA complex. (A) an individual CESA protein, (B) a rosette subunit containing six CESA proteins, (C) the rosette producing a cellulose microfibril and (D) a cross-view of the rosette consisting of six subunits with three different CESAs (blue, brown and orange) (adapted from Cosgrove, 2005).

1.2.3 Hemicelluloses

Hemicelluloses are β -(1,4) linked glucans that resemble cellulose based on their backbone structure (Cosgrove, 2005). Due to the cellulose-like conformation of the backbone, hemicelluloses can form hydrogen bonds with the cellulose microfibrils (Caffall & Mohnen, 2009). Xyloglucan and arabinoxylan are two of the most abundant hemicelluloses which differ only in the side branches (Scheller & Ulvskov, 2010). The Xyloglucan backbone has branches with xylose residues, which can be additionally decorated with galactose (Gal) and fucose (Fuc). The backbone of arabinoxylan is decorated with arabinose, glucuronic acid (Gal A) or ferulic acid ester (FAE) branches (Figure 5) (Cosgrove, 2005).

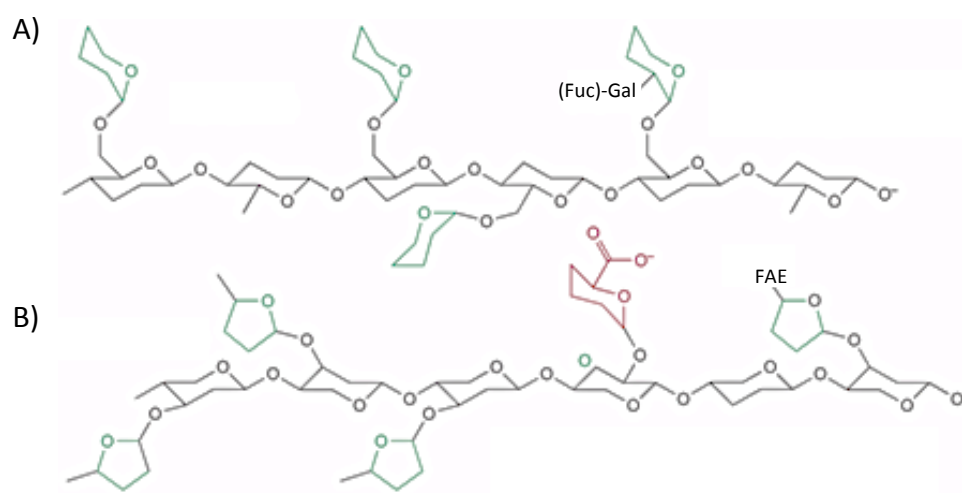


Figure 5 Hemicellulose structures, (A) xyloglucan and (B) arabinoxylan. Both types of hemicelluloses have β -(1,4)-D-glucan backbones but differ in their branches. Xyloglucan has xylose branches while arabinoxylan can have arabinose (green), glucuronic acid (red) and ferulic acid ester (FAE) branches (adapted from Cosgrove, 2005).

Xyloglucans are widely found across plant species, in which they are deposited, both in structural and storage forms (Zhou et al., 2007). Interaction of xyloglucan with cellulose via hydrogen bonds contributes to wall integrity. The binding capacity between cellulose microfibrils and xyloglucan cross-linking is determined by numbers of side chains of xyloglucan (Madson et al., 2003). Recently, Zhou et al. (2006) reviewed the possible use of xyloglucans and xyloglucan endo-transglycosylase (XET) as tools to alter cellulose properties and provide cellulosic fibre materials. They discussed the possibility of developing a

molecular toolbox consisting of *PttXET16A* from poplar, high molecular weight xyloglucans from tamarind and chemically modified xyloglucan oligosaccharides. Many applications for industrial usages such as paper or textile fibres conjugated with fluorophores or optical brightening reagents have also been proposed (Zhou et al., 2006). An example for the suitability of genetic analysis to study xyloglucan formation in *Arabidopsis* is the isolation of the *mur3* (*mur3*) mutant (Madson et al., 2003). Based on positional cloning, it was shown that the mutated gene encoded a xyloglucan galactosyltransferase, which specifically acts on the third xylose residue within the XXXG structure of xyloglucan (Madson et al., 2003).

1.2.4 Pectins

Pectins are mainly found in the plant primary cell wall and middle lamellae (Caffall and Mohnen, 2009). Pectic polysaccharides include rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II), homogalacturonan (HG), xylogalacturonan (XG), arabinan and arabinogalacturonan (Cosgrove, 2005). The structures of major pectins are shown in Figure 6. RG-II, HG and XG have linear backbones that consist of (XG), galacturonic acid while RG-I has an alternating backbone of ($\rightarrow\alpha$ -D-GalpA-1,2- α -L-Rhap-1,4 $\rightarrow$) and complex side chains can be found in both rhamnogalacturonans (Park & Cosgrove, 2012a).

Xylogalacturonan (XG) consists of a linear chain of GalA residues where some are methylesterified while the non-methylesterified GalA residues are decorated with xylose (Cosgrove, 2005).

RG-I contains a repeating ($\rightarrow$ 4)- α -D-GalA-(1,2)- α -L-Rha-(1 $\rightarrow$) backbone where the GalA residues may be acetylated (Harholt, Suttangkakul, & Vibe Scheller, 2010). The backbone is probably connected to other pectin domains, for example, arabinan, galactan and arabinogalactan (Caffall & Mohnen, 2009; Cosgrove, 2005).

RG-II is a complex structure that consists of a backbone composed of α -(1,4)-linked GalA residues, that exhibits some similarity to a HG backbone and possess oligosaccharide and disaccharide branches that are attached to the C-2

and C-3 of the backbone, respectively (Caffall & Mohnen, 2009; Harholt et al., 2010).

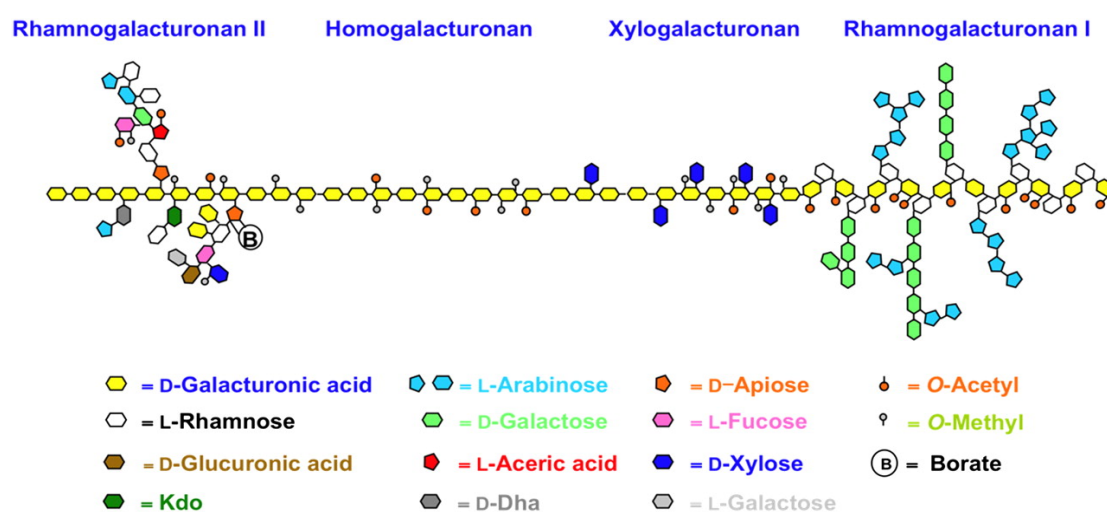


Figure 6 Schematic structure of pectin. Pectin consists of four different types of polysaccharides as their structures shown (from left to right), RhamnogalacturonanII (RG-II), Homogalacturonan (HG), Xylogalacturonan (XG) and RhamnogalacturonanI (RG-I). RGII is a complex domain, containing a galacturonic acid backbone (GalA, yellow hexagons), which can form dimers through borate esters. HG is a linear chain of GalA residues, where the carboxyl groups of some residues are esterified. Similar to HG, XG consists of the GalA backbone with esterification but differs in the xylose branches. RG-I consists of a backbone with the repeating units of GalA and rhamnose (adapted from Harholt et al., 2010).

In primary cell walls, RG-II is predominantly present as a dimer with the monomers cross-linked through borate ester bonds (Figure 7) (Cosgrove & Jarvis, 2012; Harholt et al., 2010). Detailed cell wall analysis of two L-fucose deficient *Arabidopsis* mutants (*mur1-1* and *mur1-2*) has been performed (O’Neil et al., 2001). The mutants showed dwarf phenotypes, which could be rescued by application of exogenous boric acid or L-fucose to provide boron for crosslinking (O’Neil et al., 2001).

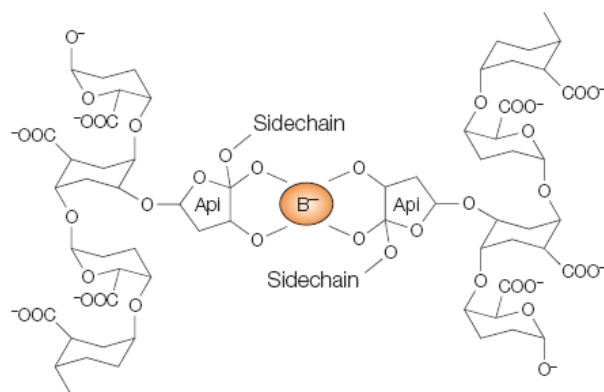


Figure 7 The structure of the RG-II boron di-ester in the walls (adapted from Cosgrove, 2005).

HG is a linear chain of α -(1,4)-linked GalA residues, where some carboxyl groups (COO^-) are methylesterified at the C-6 and O-acetylated at the O-2 or O-3 (Caffall & Mohnen, 2009; Cosgrove, 2005). HGs are synthesised in the Golgi apparatus in a methyl-esterified state and transported inside vesicles to the cell surface. Wall-bound pectin methyl esterases (PME) demethylate the chains at distinct positions within the wall (Caffall & Mohnen, 2009). In the presence of Ca^{2+} ions demethylated HG can form Ca^{2+} -bridges, thus, creating an ordered structure reminiscent of an egg-box (Figure 8) (Caffall & Mohnen, 2009).

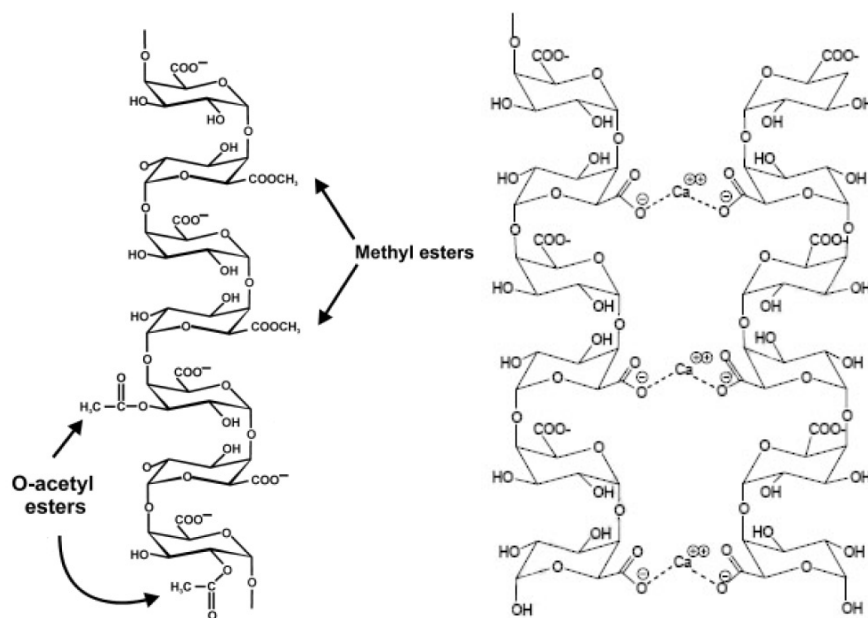


Figure 8 The structure of homogalacturonan (HG) as a linear polymer of α -(1,4)-linked galacturonic acid (GalA) residues. Methylesterification at the C-6 and O-acetylation at the C-2 or C-3 of the GalA residues are shown on the left. The egg-box model of calcium crosslinking in HG is shown on the right (adapted from Caffall & Mohnen, 2009).

1.2.5 Lignin

Plant cell walls are not only comprised of cellulose, hemicelluloses and pectins, but also lignins which are deposited in secondary walls in sclerenchyma like sclereids, fibres and tracheary elements and vessel elements of xylem (Zhong & Ye, 2007). Lignin deposition plays an important role in waterproofing and strengthening of cell walls. Accordingly, it is found in many tissues and cell types where water transport and/or high tension occur such as vascular tissue (Ralph et al., 2004). Recent studies also demonstrated that lignin is crucial for the generation of Casparian strips, the cellular junctions in roots that are important in many processes like selective nutrient uptake and protection of pathogens (Naseer et al., 2012). In addition, lignins are also involved in plant defence, responses to stresses and perturbations of the walls (Cano-Delgado et al., 2003).

The process of lignification requires expression of the genes encoding enzymes involved in monolignol synthesis which are mostly controlled by MYB transcription factors (Rogers & Campbell, 2004). The pathways of monolignol biosynthesis in tracheophytes are shown in Figure 9. The branch for S lignin (highlighted in red) is restricted to angiosperms and *Selaginella* which convergently evolved from angiosperms in the way that syringyl subunits are directly derived from the precursors of *p*-hydroxyphenyl lignin by using a unique bifunctional enzyme SmF5H (Weng et al., 2011).

In order to generate the complex mature lignin molecules, three different hydroxycinnamyl alcohol precursors (so called monolignols) are polymerised randomly via radical coupling (Holmgren, Norgren, Zhang, & Henriksson, 2009; Weng & Chapple, 2010). As shown in Figure 9, the molecules differ in the degree of aromatic ring methoxylation as para-coumaryl alcohol has no methoxyl group, coniferyl alcohol has one methoxyl group in position 3 and sinapyl alcohol possesses two methoxyl groups in positions 3 and 5 (Holmgren et al., 2009). After synthesis and modification in the cytoplasm, these monolignols are secreted and subsequently radicalized by extracellular peroxidases and laccases, which are deposited at the walls (Li et al., 2008; Ralph et al., 2004; Weng & Chapple, 2010). Through that process, three types of monolignols are produced,

H-units (hydroxyphenyl), G-units (guaiacyl) and S-units (syringyl) (Li et al., 2008; Vanholme et al., 2010; Weng & Chapple, 2010).

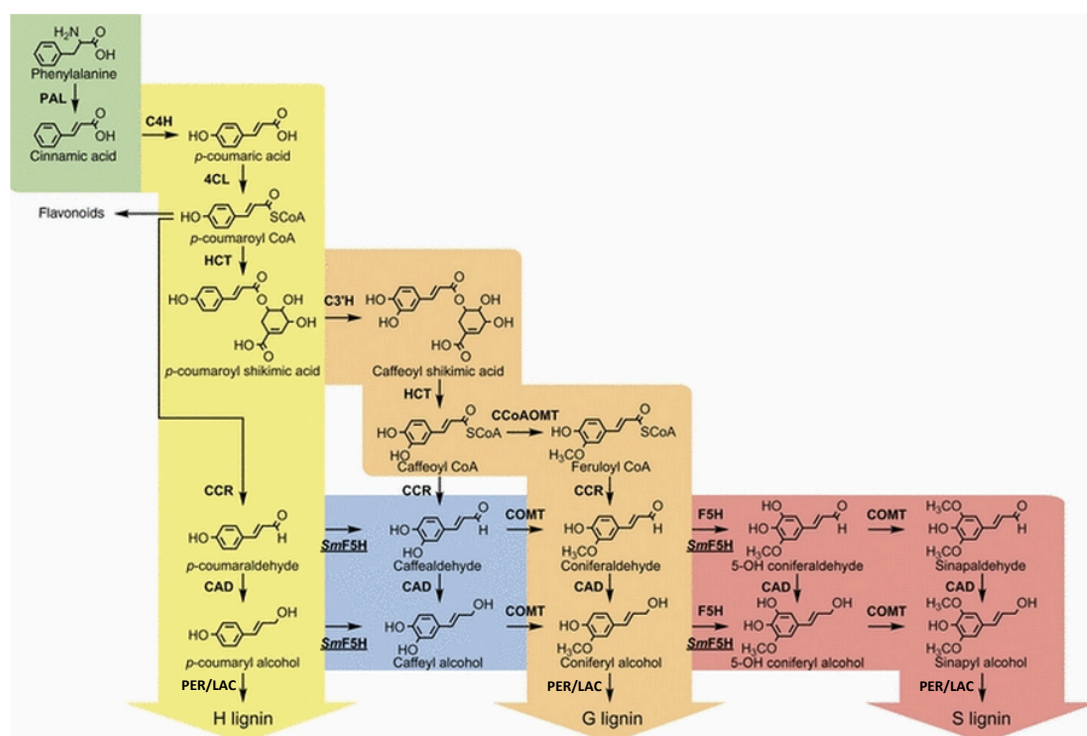


Figure 9 Schematic pathway of monolignol biosynthesis showing enzymes involved in the pathway are phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-(hydroxy) cinnamoyl CoA ligase (4CL), hydroxycinnamoyl CoA shikimate:quinic acid hydroxycinnamoyl transferase (HCT), *p*-coumaroylshikimate 3'-hydroxylase (C3'H), caffeoyl CoA *O*-methyl transferase (cCoAOMT), (hydroxy)cinnamoyl CoA reductase (CCR), ferulic acid/ coniferaldehyde/ coniferyl alcohol 5-hydroxylase (F5H); caffeic acid/ 5-hydroxyferulic acid *O*-methyltransferase (COMT), (hydroxy)cinnamyl alcohol dehydrogenase (CAD), peroxidase (PER) and laccase (LAC) (adapted from Weng & Chapple, 2010).

1.2.6 Minor components of the walls

In addition to cellulose and the matrix components, about 10% of the wall components are structural proteins, for instance, arabinogalactan proteins (AGPs), Gly-rich proteins (GRPs), Hyp-rich glycoproteins (HRGPs) like extensins (Jamet et al., 2008, 2009). Extensins are extracellular HRGPs which have hydrophobic tyrosine residues as potential cross-linking sites (Lamport et al., 2011). In *Arabidopsis*, there is a family of twenty extensins that are highly similar in sequences but their expression patterns are different (Cannon et al., 2008). An example for the function of one of *Arabidopsis* EXTENSIN has been characterised by Cannon et al. (2008). The *Arabidopsis* EXTENSIN3, AtEXT3, is expressed in cell plates and mature cell walls and appears to have an important role in cell plate

formation, since loss of functional alleles cause embryonic lethality with incomplete cell plates being detectable (Cannon et al., 2008).

1.3 Cell wall integrity maintenance in yeast and plants

The plant cell wall is not only a protective structure, providing strength to the cell; it is also involved in regulation of plant growth and development as well as interaction with the environment. During these processes composition and structure change and/or cell wall damage occurs which can lead to impairment of cell wall integrity and/or cell death. To maintain functional integrity of the wall during for example exposure to abiotic stress, mechanisms are likely to have evolved that are capable of detecting integrity impairment and initiating appropriate responses to maintain functional integrity (Humphrey et al., 2007). Interestingly, yeast cells are faced with similar challenges to plant cells when it comes to monitoring and maintaining the functional integrity of their cell walls during interaction with the environment and cell morphogenesis. Genetic studies in yeast have revealed that yeast cells possess a complex signalling network perceiving signals from the wall (Levin, 2011). Cell wall integrity signalling pathways in yeast include cell surface sensors, which are thought to detect changes in surface tension or cell wall damage (Levin, 2005; Rodicio & Heinisch, 2010). They are encoded by two different gene families (WSC; MID) (Heinisch & Dufrêne, 2010; Levin, 2005) and recruit guanosine nucleotide exchange factors (GEFs), Rom1 and Rom2, to the membrane and cause stimulation of nucleotide exchange on the G-protein, Rho1 (Levin, 2005). The effectors of Rho1 include PKC1, which translates the signal downstream to a mitogen-activated protein kinase cascade (MAPK). This cascade amplifies and transfers signals to target proteins or transcription factors, regulating gene expression (Levin, 2012). Targets of the signalling cascades can be proteins regulating organisation of the cytoskeleton, cell cycle progression and cell wall metabolism (Levin, 2005). The cell wall integrity-signalling cascade is not the only cascade regulating the activity of these target genes. The gene activities are also regulated by signalling cascades involved in mechano- (WSC/MID/MTL) and

osmo-perception (SHO1/SLN1) (Levin, 2005; Rodicio & Heinisch, 2010). Thus the combination of the three different signalling cascades provides the yeast cell with a detailed overview of the state of the cell wall and thus indirectly its environment. This enables the yeast cell to custom tailor the cellular responses to changes in the environmental changes as summarised in Figure 10.

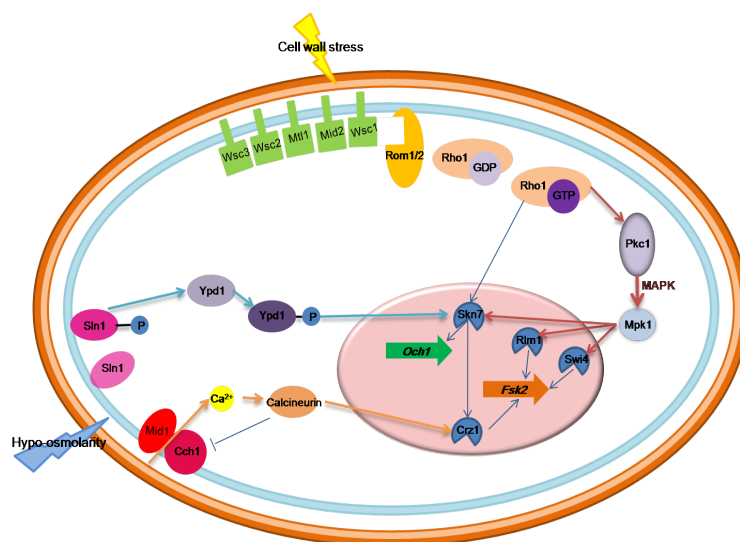


Figure 10 Schematic diagram for cell wall integrity pathway in yeast. Under cell wall stress, the signals are initiated at the plasma membrane. There are five putative mechanosensors which recruit Rom1/2 GEFs to the membrane and stimulate nucleotide exchange in Rho1 G-protein. Activated Rho1 will trigger the effectors in mitogen activated protein kinase cascade whose targets are transcriptional factors for genes encode cell wall modifying enzymes (OCH1=mannosyltransferase, FSK1=β-1,3-glucan synthase). Under hypo-osmolality condition, a part of the high osmolality pathway including the Sln1 sensor kinase is autophosphorylated. It could phosphorylate a histidine phosphotransfer protein, Ypd1 can phosphorylate Skn1 transcription factor. In calcium pathway, under stress conditions, Mpk1 can activate a Mid1/Cch1 calcium channel. This can lead to an active form of calcium-calmodulin dependent protein phosphase, calcineurin that will dephosphorylate Crz1 transcription factor into active form, allowing transcription of Fsk1 gene. (redrawn from Levin, 2005)

While the cell wall integrity signalling pathways in yeast are well characterized, the equivalent mechanisms in plants remain to be characterised. Genetic studies of mutants with defects in cell wall metabolism provide initial insights into the mode of action of the plant cell wall integrity maintenance mechanism. *Arabidopsis* mutants deficient in cellulose biosynthesis like *ectopic lignification1* (*eli1*) and *constitutive expression of VSP1* (*cev1*) both affecting the cellulose synthase gene, *CESA3* exhibited ectopic lignin deposition and activation of a stress-induced reporter (*VSP1*) (Cano-Delgado et al., 2000, 2003; Ellis et al., 2002). Phenotypic studies of these mutants detected shorter roots and hypocotyls in mutant seedlings that were stunted compared to wild type

controls. These changes were caused by the reduction of cellulose biosynthesis levels. Additionally, jasmonate and ethylene levels were elevated suggesting activation of traditional plant stress response pathways (Ellis et al., 2002). These phenotypes can be mimicked using the cellulose biosynthesis inhibitor, isoxaben (Cano-Delgado et al., 2003; Ellis et al., 2002). In accordance, Hamann et al. (2009) reported that isoxaben does not only cause changes in cellulose and neutral cell wall sugar contents, lignin deposition and the measured jasmonic acid, but also activates several defence responsive genes and affects expression of the *UDP-D-xylose 4-epimerase (UXE4)* gene, involved in sugar cell wall biosynthesis. The study also showed that these processes depend on the presence of hexose sugars, which do not have to be metabolisable (Hamann et al., 2009). Previous work has shown that manipulation of extracellular hexose concentrations through targeted expression of invertases can lead to activation of stress response pathways and lignin deposition (Roitsch, 2004). In parallel it has been shown that genes required for reactive oxygen species signalling RBOH(DF), mechano-perception (MCA1) and possibly ligand-binding (THE) are involved in the response to cell wall damage. In particular the mutant phenotypes observed with *THE* alleles suggest that the same mechanism detecting cell wall damage is active during development and in response to stress (Hematy et al., 2007; Denness et al., 2011). MCA1 was originally isolated because of its ability to rescue a yeast strain deficient for MID1 activity. Mid1 encodes a stretch-activated calcium channel protein that forms part of the cell wall integrity maintenance mechanism (Humphrey et al., 2007). This observation suggests that a certain degree of functional conservation may exist between yeast and plants (Denness & Hamann, 2012). The cell wall integrity maintenance mechanism in plants can be separated into an early and a late stage with reactive oxygen species, jasmonic acid and calcium based signalling processes mediating the cellular and cell wall responses (Denness et al., 2011). This also allows the separation into a perception or signalling phase and a later stage at which changes in cellular and cell wall metabolism seem to take place. The existence of this stage is supported by the observations that cell wall damage induces ectopic lignin and jasmonic acid production only after four to six hours and changes in

carbohydrate distribution start occurring after eight to ten hours (Denness et al., 2011; Wormit et al., 2011).

Additional insights into the mode of action of the cell wall integrity mechanism come from research into the control of plant growth by regulation of calcium based crosslinking of pectins and crosslinking between xyloglucans and cellulose microfibrils (Carpita, 2011; Peaucelle et al., 2012). It had been shown that an *Arabidopsis* double mutant (*xxt1/xxt2*), entirely lacking xyloglucans, exhibits subtle growth phenotype with all the wall properties comparable to wild type (Cavalier et al., 2008). Three dimensional nuclear magnetic resonance spectroscopy (3D-NMR) studies found, that only a small amount of xyloglucan was bound to cellulose microfibrils (Dick-Pérez et al., 2011).

In parallel with the microscopy studies of cell wall structure components, creep measurements in cucumber where cell wall movement of the growing cells was observed using extensometer assays were performed. The creep measurement results showed that creep is promoted by extracellular glucanases which are not specific to either xyloglucan or cellulose (Park & Cosgrove, 2012a). Therefore it could be concluded that xyloglucan is probably not the single major control element of cell growth and expansion. Interaction of xyloglucan and covalently bound pectin has been discussed extensively (Popper & Fry, 2008). However, it has been shown from acid promoted cell wall creep study that the expansin dependent creep in *xxt1/xxt2* double mutant plants is consistent with one was measured in wild type plants, while the creep of calcium -pectate or -xylan treated walls in the mutant was more pronounced than in wild type (Park & Cosgrove, 2012b), suggesting a pivotal role of xylans as load bearing in *xxt1/xxt2* xyloglucan deficient mutant. For extensin dependent cell growth and expansion, it had been shown by atomic force microscopy that amphiphilic extensins can self assemble *in vitro* (Cannon et al., 2008).

In addition, it was also found that the positive charge extensins can form a network with negative charge pectin, resulting in a scaffold provided for cell plate formation (Valentin et al., 2010). The same study also showed that positive charge and conformation of extensins are crucial for interaction between extensins and pectins (Valentin et al., 2010). Together with the extensins properties, dimethyl esterified pectins that possess more negative charge also

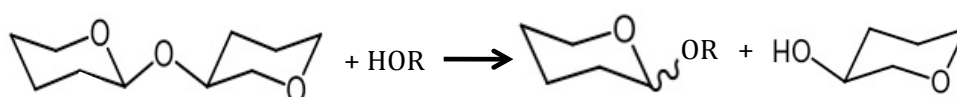
promote the interaction and thus promote compaction of the wall (Scheller & Ulvskov, 2010).

To summarise, the available data support the existence of a plant cell wall integrity maintenance mechanism. They provide some initial idea about the process, signalling components involved (i.e. receptors, messengers and regulators) and cellular processes affected. However, they also highlight that our understanding of the mode of action and the signalling mechanisms bringing about changes in wall structure and composition is very limited.

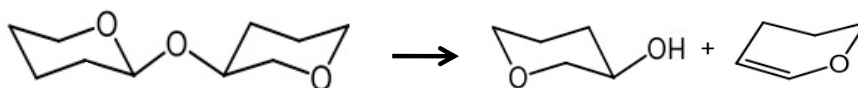
1.4 Enzymes modifying plant cell wall

Carbohydrate-active enzymes are categorised and described according to their catalytic and carbohydrate-binding modules (or functional domains) that they consist of. The central repository providing an overview of the resulting families is the Carbohydrate-Active Enzymes (CAZy) database (www.cazy.org). The cell wall polysaccharide modifying enzymes can be grouped into following four categories based on their putative catalytic activities (Henrissat et al., 2001).

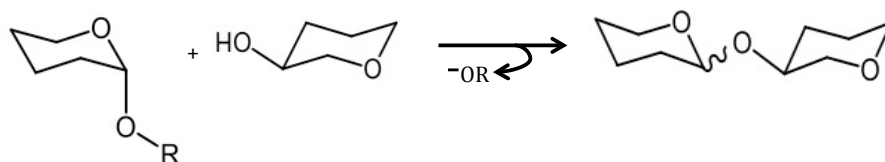
1. Glycoside hydrolases show hydrolytic activity when an activated donor substrate (HOR) is water or transglycosylation when HOR is a non-water nucleophile.



2. Polysaccharide lyases catalyse β -link elimination between glycosides.

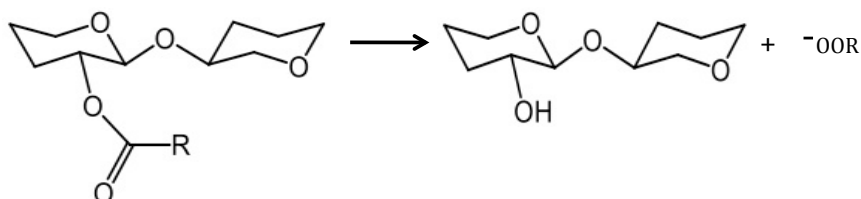


3. Glycosyltransferases catalyse glycosidic bond synthesis using activated sugar donors. The activating group (R) can be a nucleotide or lipid phosphate.

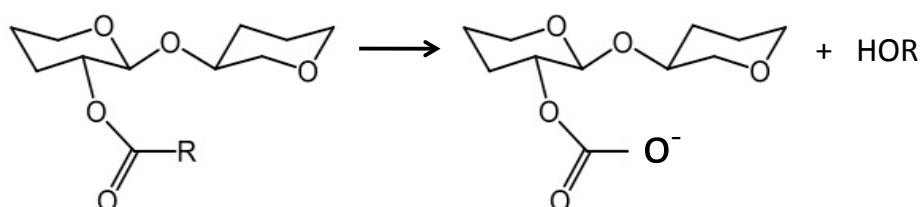


4. Carbohydrate esterases catalyse de-*O*-acetylation reaction.

The reaction of alcohol sugars



The reaction of acid sugars



1.5 Glycosyl hydrolases (GHs)

GHs represent one of the two major classes of carbohydrate-active enzymes (CAZY), the second class being Glycoside Transferases (GTs). Of the over seven hundred thirty open reading frames compiled in CAZY, three hundred seventy nine belong to GHs and are classified into twenty nine families (Henrissat et al., 2001). GHs are involved in the breakdown or hydrolysis of carbohydrates. As carbohydrates play important roles in various physiological processes, GHs have a multitude of functions, such as mobilization of storage reserves, plant defence, symbiosis, secondary metabolism, signalling, glycolipid metabolism and most important cell wall metabolism (Minic 2008). Glycoside hydrolase enzyme family in Arabidopsis with their putative functions and substrates are listed in Table 2.

Table 2 Glycoside hydrolase enzyme families in *Arabidopsis* involved in cell wall metabolism (adapted from Minic & Jouanin, 2006; Minic, 2008)

Family	Functions	Substrates
GH1	β -D-Glucosidase	Glucan/cellulose, xyloglucan
GH3	β -D-Xylosidase	Xylan, arabinan, arabinoxylan
	α -D-Arabinofuranosidase	
GH5	endo-1,4- β -Glucanase (cellulase)	Glucan/cellulose
	endo-1,4- β -Mannanase	Mannan
GH9	endo-1,4- β -Glucanase	Glucan/cellulose
GH10	1,4- β -Xylan endohydrolase	Xylan
GH16	Xyloglucan endotransglycosylase	Xyloglucan
GH17	endo-1,3- β -Glucanase	1,3- β -Glucan
	endo-1,4- β -Glucanase	(1,3)(1,4)- β -Glucan
GH27	α -D-Galactosidase	Galactomannan
GH28	endo-polygalacturonase	Homogalacturonan
	exo-polygalacturonase	
GH31	α -D-Xylosidase	Xyloglucan
	α -D-Glucosidase	
GH35	β -D-Galactosidase	Galactan
GH43	1,4- β -Xylan endrohydrolase	Xylan
GH51	α -L-Arabinofuranosidase	Arabinoxylan, xylan, arabinan

According to Minic & Jouanin (2006), thirteen of the twenty-nine GH families, containing around two hundred proteins in *Arabidopsis*, could be involved in reorganization of plant cell wall polysaccharides. So far for only very few GHs the physiological and biochemical functions have been experimentally demonstrated. The structures of plant cell walls change during growth and development as well as in response to biotic and abiotic stress. Members of the GH gene families are possibly playing important roles in these biological processes (Minic, 2008) The next two chapters will focus on two GH families, family 9 and 16, which have been implicated in cell wall polysaccharide modifications and are of particular interest.

1.6 Selection of candidate genes for functional analysis

From the previous study, a time course experiment using Affymetrix DNA microarrays was performed to characterise effects of isoxaben on *Arabidopsis* seedlings and to identify genes involved in cell wall remodelling and formation (Hamann et al., 2009). Isoxaben is a highly specific cellulose biosynthesis inhibitor that can be used to generate cell wall damage by preventing production of the load bearing cellulose microfibrils in elongating plant cells. The resulting

expression profiling data suggested that approximately three hundred fifty genes exhibited transcriptional changes from four hours after isoxaben treatment. Hundred twenty of them show high expression in tissues undergoing cell wall remodelling or formation (based on the Genevestigator database).

From the a hundred twenty genes that expression profiles suggest responses of the gene to isoxaben and being highly expressed in tissues undergoing cell wall formation, eight out of those genes are encoded putative enzymes involved in plant cell wall metabolism. List of the eight selected genes and the putative functions of encoded enzymes are shown in Table 3.

Table 3 Initial selected genes with their putative protein functions

ATG numbers	Putative encoded proteins
<i>At1g04680</i>	Probable pectate lyase 1 (PLY1)
<i>At1g07260</i>	UDP-glycosyltransferase 71C3 (UGT 71C3)
<i>At1g64390</i>	Endo-1,4- β glucanase 6 (GUN6)
<i>At1g65610</i>	Endo-1,4- β glucanase 7 (GUN7)/ Korrigan 2 (KOR2)
<i>At3g11340</i>	UDP-glycosyltransferase 76B1 (UGT 76B1)
<i>At3g44990</i>	Probable xyloglucan endotransglucosylase/hydrolase protein 31 (XTH31)
<i>At3g49220</i>	Probable pectinesterase/pectinesterase inhibitor 34 (PME34)
<i>At3g53190</i>	Probable pectate lyase 12 (PLY12)

These eight genes encoding cell wall modifying enzymes were selected and divided into two groups based on the expression data. The first group contains five genes showing down regulated expression in response to isoxaben (Figure 11, panel A), while the second group contains three genes exhibiting increased expression (Figure 11, panel B).

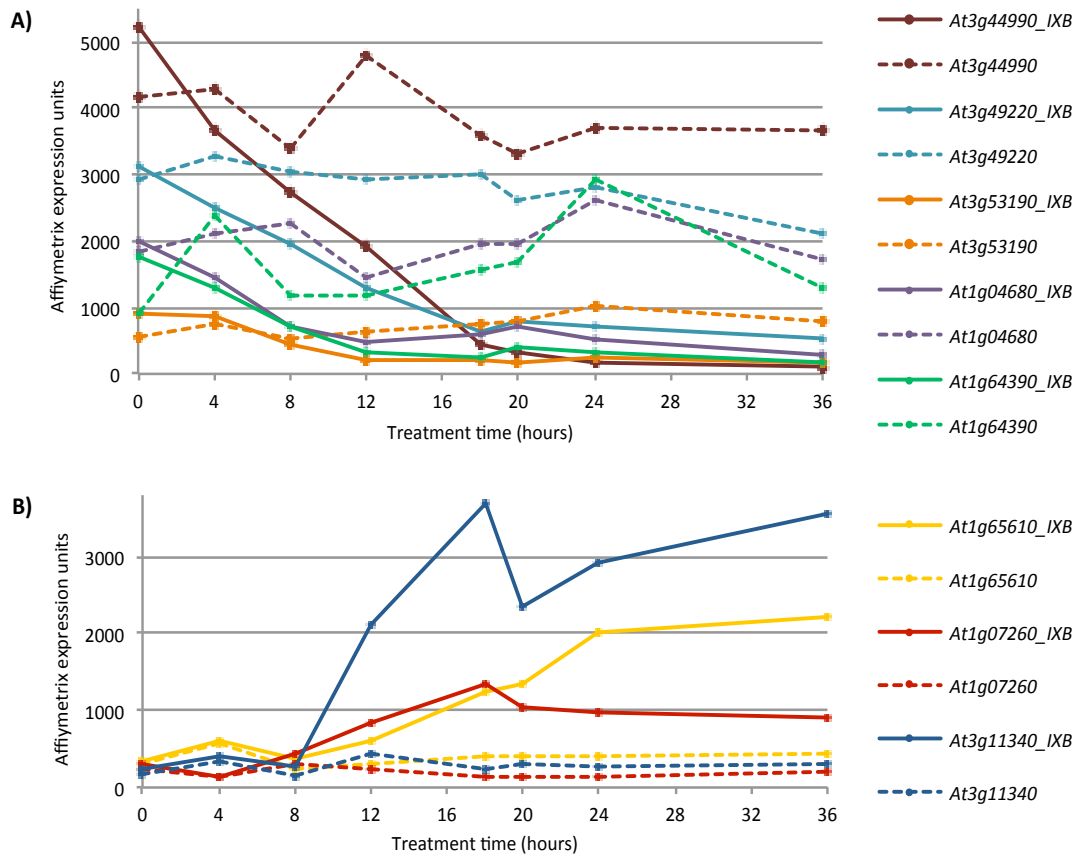
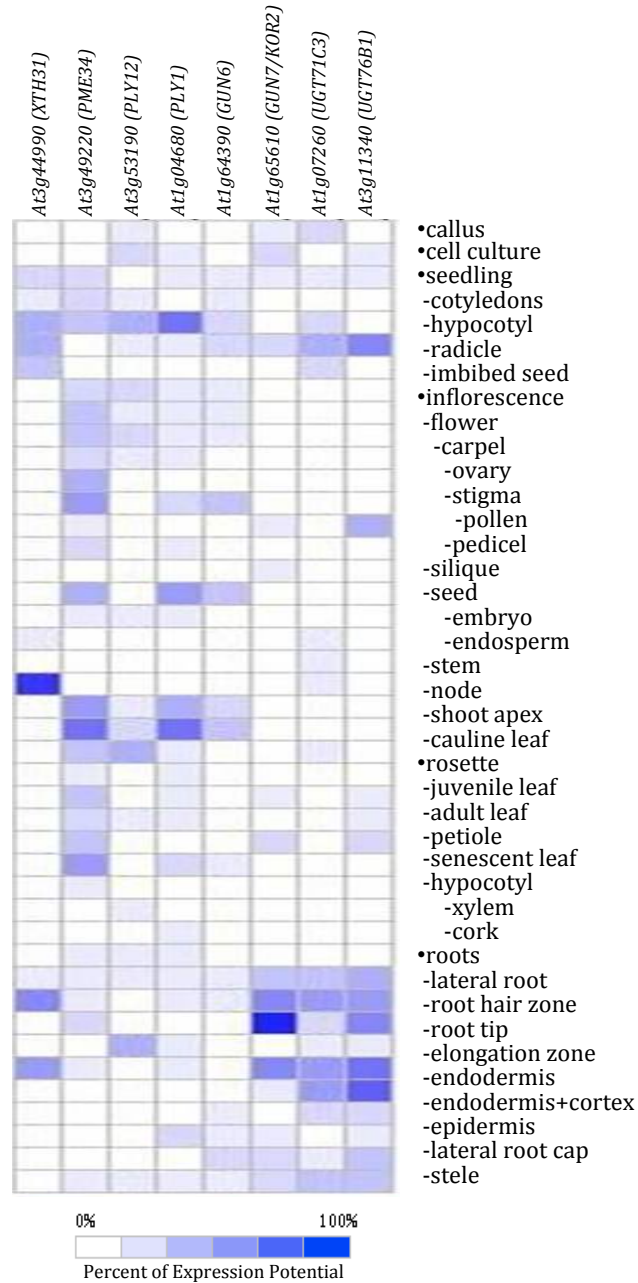


Figure 11 Candidate gene expression profiles from *Arabidopsis* seedlings isoxaben or mock treated (represented by bold and dash lines). The Y-axis represents the expression level, while the X-axis represents the treatment period (hours). Samples with isoxaben treatment are labelled as ‘_IXB’ (expression data derived from Hamann *et al.* 2009).

In order to determine possible biological functions of the selected genes and reduce their number at the same time, expression data for the selected genes was obtained from public expression databases. For this purpose expression profiles deposited in the Genevestigator database were used. Figure 12 shows an example for expression data derived from Genevestigator. Expression levels of the target genes are shown in a heat map pattern where the darkest blue colour indicates strongest expression of the genes whereas lighter colors indicate weaker expression.

A)



B)

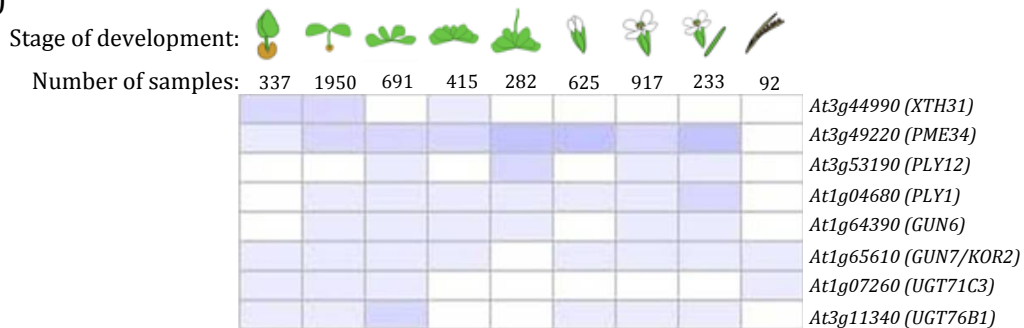


Figure 12 Expression profiles of the candidate genes derived from Genevestigator. (A) shows the expression profiles of the genes in different types of plant organs and tissues, while (B) shows the profiles from different plant developmental stages. Heat map represented percentage of expression potential for candidate genes were derived based on microarray expression profiles deposited on GeneVestigator database.

1.7 Aims of the project

This study aims to gain a better understanding of cell wall metabolism using molecular biological techniques and biochemical tools to characterise candidate enzyme activities. For the preliminary study, eight *Arabidopsis* genes that showed altered expression in response to the cellulose biosynthesis inhibitor, isoxaben, encode the candidate enzymes were selected. General bioinformatics analyses for gene structures and predicted domains of the putative proteins for the eight target genes are shown in Chapter 3 (see 3.1). To determine the biological function of the candidate genes, mutant plants generated through T-DNA insertion were employed as a genetic tool for the biological characterisation of target genes. At least two independent insertion lines were used and screened for homozygous plants (Appendix 3, 3.1). In order to determine biochemical functions of the encoded enzymes, heterologous expression using *Pichia pastoris* as the expression system were performed (Appendix 3, 3.2). Preliminary results are shown in Appendix 3. However, successful express heterologous proteins in active forms via *Pichia* could not be achieved. Due to time constrain, the scope of this study were adjusted. The selection to pursue throughout this study includes three genes: *At1g65610*, *At3g44990* and *At1g07260* that predicted functions based on amino acid sequences are involved in metabolisms of the main load bearing contents of the walls: cellulose, xyloglucan and probably lignin, respectively.

Therefore, the current study aims to characterize the activities and biological functions of three candidate genes (*At1g65610*, *At3g44990* and *At1g07260*) implicated by the previously described expression profiling experiments in cell wall metabolism (Figure 11 and 12). Interestingly, apart from the fact that the putative functions of these three target genes are involved in the metabolism of most abundant structures of the wall like lignin (*At1g07260*), cellulose (*At1g65610*) and xyloglucan (*At3g44990*), microarray expression profiles of these target genes show that under cell wall impairment via isoxaben treatment, the expression of *At1g07260* and *At1g65610* were up-regulated while *At3g44990* were down-regulated (Figure 13).

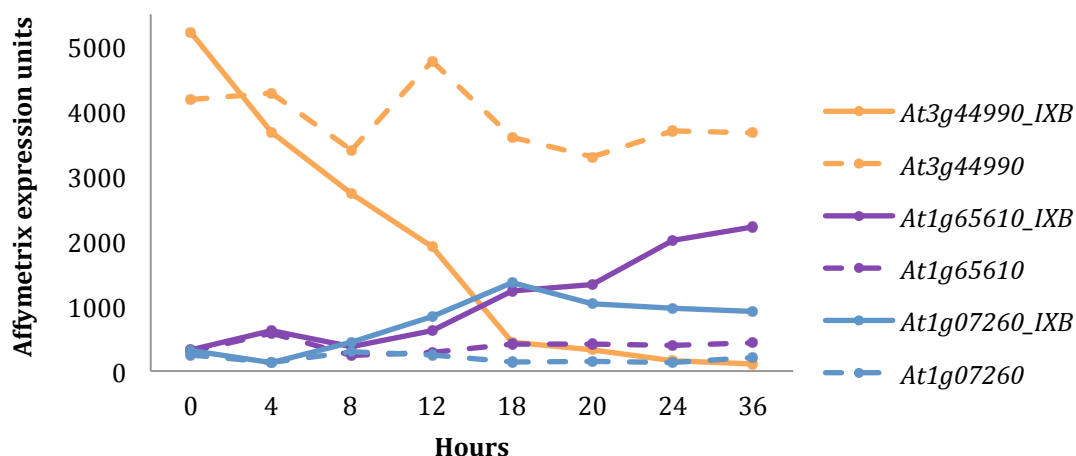


Figure 13 Candidate gene expression profiles of *Arabidopsis* seedlings in response to isoxaben and mock treatment (represented by bold and dash lines). The Y-axis represents the expression level, while the X-axis represents the period of treatment (hours). Samples with isoxaben treatment are labelled as ‘_IXB’ (expression data derived from Hamann, 2009).

To determine biological functions of the candidate genes, knock out plants generated via T-DNA insertion were phenotypically characterised. For the biochemical functions of the candidate enzymes, cloning and heterologous expression of the candidate genes were performed via *Escherichia coli* expression system. By monitoring and observing mutant phenotypes either from mutants of up-regulated or down-regulated under cell wall impairment, the results and knowledge gain from this study would be able to provide more evidence in the wall metabolism. More understanding support existing relevant pathways and the regulation mechanisms could be used either as molecular markers or manipulate target genes to select crops for human consume, livestock or lignocellulosic biofuel industries.

Chapter 2 Materials and methods

2.1 Bioinformatics

Expression and co-expression analysis of the candidate genes were performed using publically available expression data in the Genevestigator database (<https://www.genevestigator.com/gv/plant.jsp>). To analyse putative conserved domains of the candidates, amino acid sequences for the three candidate proteins were aligned against the non-redundant protein database (<http://blast.ncbi.nlm.nih.gov/>). Secondary-structures of the candidates were predicted using PredictProtein software (<http://www.predictprotein.org/>). For cloning purposes, amino acid sequences were analysed for any signal peptides or transmembrane domains using SignalP software via SignalP4.0 server (<http://www.cbs.dtu.dk/services/signalp/>).

2.2 Materials and methods for characterisation target genes

2.2.1 Plant materials and growth conditions

Seeds for *Arabidopsis thaliana* (Columbia ecotype; Col-0) and the T-DNA insertion lines were obtained from the Nottingham Arabidopsis Stock Centre (NASC, <http://arabidopsis.info/>). Plants were grown in a Fitotron growth room (Gallenkamp PLC, UK) under sixteen hour-day and eight hour-night cycles, fluorescent light irradiance 130-175 $\mu\text{E m}^{-2} \text{sec}^{-1}$ and 60 % relative humidity at 21-23°C. To prepare plant material for genetic screening, at least ten plants were screened by sowing two seeds for each insertion line in a 9.5 cm pot containing fine compost supplemented with sand, Levington®F2+S (Levington, UK).

2.2.2 Isolation of T-DNA insertion lines

Information on insertion lines and primers to identify homozygous mutant plants were obtained from the Salk Institute Genomic Analysis Laboratory (SIGNAL) (<http://signal.salk.edu>). For several insertion lines, alternative primers

were designed (Appendix1). The forward primer was designed based on the available T-DNA insertion sequences and reverse primers (gene specific primers) were generated against predicted regions of the insertions in the candidate genes.

2.2.3 Seed sterilization and media preparation for plant growth

For experiments requiring sterile conditions to grow seedlings, 50 mg of seeds were surface sterilized for five minutes with 70 % ethanol, followed by five minutes incubation in 50 % bleach and washing three times with sterile water. Sterile seeds were grown for six days in 250 ml flasks containing 125 ml ½ MS media supplemented with 1 % sucrose on a shaker at 135 rpm. After six days, fresh media was supplied to the cultures at the onset of isoxaben treatment. Seedlings were treated either with DMSO (mock treatment) or 600 nm isoxaben dissolved in DMSO for the time-course experiment.

To harvest samples for lignin deposition analysis, 50-100 mg of seedling material were collected and stored in 70 % ethanol to remove chlorophyll. Samples for RNA extraction and cell wall component analysis seedlings were rinsed well with MilliQ water before being flash frozen in liquid nitrogen and stored at -80°C.

2.2.4 Isolation of homozygous insertion lines

2.2.4.1 DNA extraction

The first fully expanded leaf of a four-week old plant was ground in 400 µl DNA extraction buffer (200 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA and 0.5% SDS) and centrifuged at 9,660 xg for two minutes. Subsequently, the supernatant was transferred into fresh tubes and 350 µl of isopropanol was added. The mixture was mixed by inverting the microcentrifuge tubes and incubated at room temperature for two minutes before centrifugation at 9,660 xg for ten minutes. After decanting the supernatant, the pellet was air-dried for three minutes and re-dissolved in 50 µl Tris buffer (pH 7.5).

2.2.4.2 RNA isolation

To confirm the disruption of transcript in mutant plants, total RNA was prepared from approximately 100 mg frozen seedlings using Trizol® reagent (Invitrogen) as described in the manufacturer's manual.

2.2.4.3 Polymerase chain reaction (PCR) - based genotyping

T-DNA PCR reactions were carried out using primers for T-DNA identification and gene specific primers (Appendix 1), and isolated DNA samples as described above. A 25 µl reaction contained 0.2 mM dNTPs, 1 µM of each primer, 1 unit Dreamtaq™ DNA Polymerase (Fermentas) in Dreamtaq™ buffer (20 mM MgCl₂) and approximately 0.5 µg DNA template. The following PCR cycling conditions were used: denaturation at 95°C for 30 seconds, annealing at a temperature 5°C below the melting point (T_m) of primers for thirty seconds, extension at 72°C for a minute per kilobase (kb) product (28-40 cycles) and a seven-minute final extension at 72°C.

2.2.4.4 Reverse transcriptase PCR (RT-PCR)

To confirm disruption of the target genes, RT-PCR was employed using One Step RT-PCR kit (Qiagen) as described by the manufacturer's protocol. RT-PCR reactions using *UBIQUITIN* primers (Appendix 1) was run as a positive loading control.

2.2.4.5 Agarose gel electrophoresis and DNA sequencing

PCR products were analysed by electrophoresis using 1 % agarose gel supplemented with 0.002 % SYBR® Safe DNA gel stain (Invitrogen). The gels were run at 80 V and imaged in an ultraviolet transilluminator light with a SYBR Sage filter (BioImaging Systems). In order to determine size and quantity of the PCR product, 0.5 µg of Gene Ruler™ 1 kb DNA ladder Plus (Fermentas) was used.

2.2.5 Phenotypic characterization of the mutants

2.2.5.1 Germination assay

To analyse the germination of each T-DNA insertion line compare to Col-0 wild type, surface sterile seeds were sown on $\frac{1}{2}$ MS basal media with 1 % sucrose plates. Wild type and mutant seeds were arranged in the same plate side by side as shown in Figure 14. After the arrangement on the plates, seeds were vernalized at 4°C for two days before putting in the Fitotron under the growth conditions described in 2.2.1. The seeds from which the radicle had emerged were defined as germinated seeds.

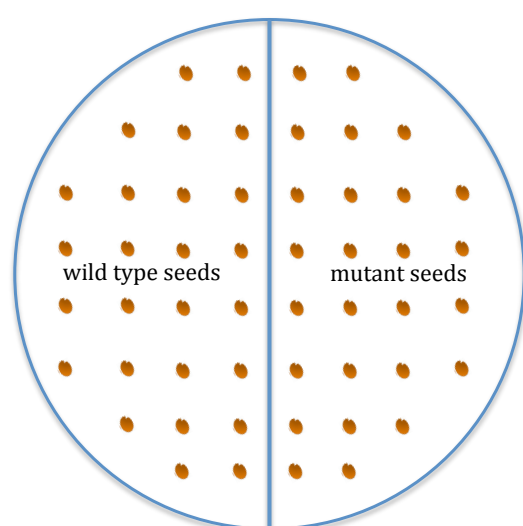


Figure 14 Schematic depiction for the arrangement of *Arabidopsis* seeds sown on a round plate of $\frac{1}{2}$ MS with 1% sucrose supplement for germination assay.

2.2.5.2 Growth stage based analysis

Growth stage based analysis was performed to characterize phenotypes of the mutant plants compared with Col-0 wild type plants as described (Boyes et al., 2001). Plants were grown under standard conditions as described in 2.2.1 and the pots randomly arranged as the example shown in Figure 15.



Figure 15 An example for the arrangement of plant pots grown for growth stage based analysis.

2.2.5.3 Root growth assay

To compare the root lengths of seedlings of the different T-DNA insertion lines with the Col-0 wild type, seedlings were grown on square plates of $\frac{1}{2}$ MS basal media supplemented with 1 % sucrose. Seeds were surface sterilised as described in 2.2.3 and arranged on the prepared plates (Figure 16). Then seeds were vernalized at 4°C for two days before grown vertically on the plates for four days. Pictures were captured using Leica MZ16F with the Leica image-processing program (Leica, Germany). The ImageJ (rsbweb.nih.gov/ij/) program was employed for the processing and measurement. Data analysis and statistics analysis was performed via Microsoft Excel (Microsoft).

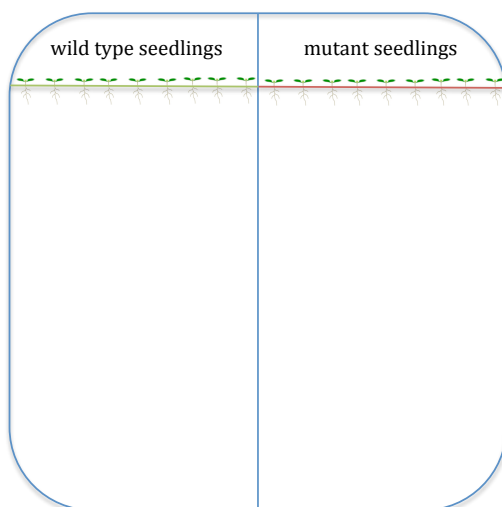


Figure 16 Schematic depiction for the arrangement of *Arabidopsis* seedlings grown on a square plate of $\frac{1}{2}$ MS with 1% sucrose supplement for root growth assay.

2.2.5.4 Phloroglucinol staining for lignin deposition analysis

To analyse ectopic lignin deposition in seedlings, saturated phloroglucinol stain was used (1 mg ml⁻¹ w/v in 20 % HCl). Before the staining, chlorophyll was removed from the seedlings using 70 % ethanol. The cleared seedlings were mounted with phloroglucinol onto a microscopic slide and incubated for three minutes. Imaging was performed using a digital camera (AxioCam-HR; Carl-Zeiss) connected to a light microscope (Axioskop2-plus; Carl-Zeiss) with image capture software (AxioVision-3.1.2.1, v5.07.03; Carl-Zeiss). To semi-quantitatively analyse stained area of lignin deposition from the photos, ImageJ (rsbweb.nih.gov/ij/) program was employed for the processing and measurement total area of the captured root and stained root area. Data analysis and statistics analysis was performed via Microsoft Excel (Microsoft). Values for percentage of staining were calculated as (stained area/total captured root area) × 100. Student t-test using Microsoft Excel was performed for significant differences between each mutant and wild type.

2.2.5.5 Cell wall analysis

To obtain fine powder of cell wall materials, main stem parts of eight week old mature plants were dried at 37°C for three days before processing with a tissue-lyser (TissueLyser-II, QIAGEN). For each assay, 1 mg powder was used. Four technical replicates were weighted out for each sample and three biological replicates were used for each genotype. Cellulose content was quantified as described by Updegraff (1969). Uronic acids content was quantified as described by Blumenkrantz & Asboe-Hansen (1973). Neutral cell wall sugars were quantified as described by Blakeney et al., (1983), with the modifications as described below. Hydrolysates from samples boiled in 400 µl of 1 M sulphuric acid for an hour were obtained and subsequently neutralised with 9 M ammonia. Sugar standards were used as controls (0.1 % w/v each: rhamnose, fucose, arabinose, xylose, mannose, galactose, glucose and inositol as an internal standard). Sugar acetates were extracted in 2 ml chloroform at 4°C overnight and analysed by gas chromatography mass spectrometry using MSD ChemStation software (B75; Agilent) using conditions with modifications as described below.

A Hewlett-Packard (HP) 6890 GC-5073 MSD was used for the analysis, with ionization achieved by electron impact at 70 eV. The capillary column used was a HP-5MS Agilent 19091S-433 (5 % phenyl methylpolysiloxane, 30 m × 0.25 mm ID, 0.25 µm thick film) with a capillary (30 m × 250 µm × 0.25 µm nominal). The injector was operated in the split mode. The operating conditions were an injection port temperature at 280°C; interface temperature, 280°C; column oven temperature hold at 160°C for ten minutes followed by 4°C min⁻¹ ramp to 204°C; helium carrier gas (flow rate of 1 ml min⁻¹); 1 µl injection volume. A student t-test analysis was used to compare grouped means of data sets between mutants and wild type (Microsoft Excel).

2.3 Molecular cloning for heterologous expression

2.3.1 *Escherichia coli* strains employed

The strains, genetics and other information for the *E. coli* used in this project are shown in Table 4.

Table 4 Genetics of *E. coli* strains used in the project

Strain	Genotypes	Sources
DH5α	F ⁻ endA ¹ glnV44 thi ⁻¹ recA ¹ relA ¹ gyrA96 deoR nupG Φ80d ^{lacZ} ΔM15 Δ(<i>lacZYA-argF</i>)U169, hsdR17(r _K ⁻ m _K ⁺), λ ⁻	Invitrogen
BL21	B F ⁻ <i>dcm ompT hsdS</i> (r _B ⁻ m _B ⁻) gal [malB ⁺] _{K-12} (λ ^S)	Stratagene
BL21 (DE3)	F ⁻ <i>ompT gal dcm lon hsdS</i> _B (r _B ⁻ m _B ⁻) λ(DE3 [<i>lacI lacUV5-T7 gene 1 ind1 sam7 nin5</i>])	Stratagene
SoluBL21™	F ⁻ <i>ompT hsdSB</i> (r _B ⁻ m _B ⁻) gal <i>dcm</i> (DE3)	Amsbio

2.3.2 Molecular cloning procedures

Unless stated otherwise, chemicals were purchased from Sigma-Aldrich (UK). cDNA clones were obtained from the *Arabidopsis* Biological Resource Centre (ABRC, the Ohio State University). For *At1g65610*, the cDNA clone was not available. Thus, reverse transcription of the coding region of the target gene was generated using the Omniscript RT Kit (Qiagen). *Arabidopsis* Col-0 total RNA prepared from seedlings were used as template.

2.3.3 *E. coli* transformation competent cells preparation

2.3.3.1 Electro-competent cells

Overnight cultures were grown in 5 ml of LB medium (1% w/v tryptone, 0.5 % w/v yeast extract, 0.5 % w/v NaCl) using a shaking incubator at 30°C 230 rpm to an OD₆₀₀ of 4-6. An active growing culture was started in 500 ml of SOB medium (2 % w/v tryptone, 0.5 % w/v yeast extract, 10 mM NaCl, 2.5 mM KCl) with the final cell density at OD₆₀₀ of 0.04 and grown until OD₆₀₀ of 0.3-0.5 was achieved. The culture was immediately cooled on ice. Cell were harvested at 4°C, 4,000 xg for fifteen minutes and washed for four times with 50 ml ice cold water before washing twice with 10 ml of ice cold 10 % glycerol. The cells were suspended in 5 ml of 10 % glycerol. Cell suspension was flash frozen in liquid nitrogen in 100 µl aliquots.

2.3.3.2 Chemically competent cells

Overnight cultures were grown in 5 ml of LB using a shaking incubator at 37°C 230 rpm to an OD₆₀₀ of 4-6. An active growing culture was prepared from 50 µl of the overnight cultures grown in 10 ml pre-warmed ψB medium (2 % w/v tryptone, 0.5% w/v yeast extract, 100 mM KCl, 34 mM MgSO₄) to until an A₅₅₀ of 0.3 was achieved. Then half of the active growing culture 5ml of the ψB cultures was inoculated into 200 ml pre-warmed ψB medium and grown until an A₅₅₀ of 0.3-0.4 was achieved. The culture was then cooled on ice. Cells were harvested by centrifugation at 4°C, 2,000 xg for eight minutes before re-suspended in 60 ml ice-cold TFβI (30 mM CH₃CO₂K, 50 mM MnCl₂, 100 mM RbCl, 10 mM CaCl₂, 15 % v/v glycerol pH 5.8). The suspension was centrifuged under the same conditions before re-suspended in 8 ml ice-cold TFβII (75 mM CaCl₂, 10 mM RbCl, 15 % v/v glycerol, 10 mM MOPs pH 7). Cell suspension was flash frozen in liquid nitrogen in 100 µl aliquots.

2.3.4 Transformation methods

2.3.4.1 Electroporation

Up to 600 ng DNA fragment or plasmid vectors were added to an aliquot of electro-competent cells (~50 µl) and mixed gently with a pipette. The mixture was transferred to an ice-cold electroporation cuvette (1 mm for bacteria, 2 mm for bacteria or yeast) and given a pulse at 18 kV, 250 µF, 200 Ω using a Gene Pulser®II (Bio-Rad). After the pulse, 950 µl of SOC medium (2 % w/v tryptone, 0.5 % w/v yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 20 mM glucose) were added immediately. The mixture was transferred to a 15 ml tube and cells were allowed to recover at 37°C for an hour with full aeration ensured through shaking at 230 rpm.

2.3.4.2 Heat-shock transformation

Chemically competent cells were prepared as described in 2.3.3.2. The competent cells were thawed on ice and gently mixed with 1-2 µl of either ligation reaction or plasmid DNA as appropriated. The mixture was incubated on ice for thirty minutes before performing heat-shock transformation. Heat-shock procedure is to shock cells at 42°C for forty-five seconds. The mixture was cooled on ice for five minutes immediately after the heat-shock. Then cells were recovered in 1 ml LB or SOC medium and shaken at 37°C for an hour with fully aeration. The recovered cells were plated on LB agar plates supplemented with appropriate antibiotics.

2.3.5 Protein expression and purification

2.3.5.1 Conditions for expression of the recombinant proteins

Overnight culture were grown in 20 ml LB and used for inoculation of 1 L of LB medium. The cultures were grown at 37°C with vigorous shaking until an OD₆₀₀ of 0.5-0.6 was achieved. Protein expression was subsequently induced by addition of IPTG (1 mM final concentration). The host cells were harvested by centrifugation at 4°C at 4000 xg for twenty min. The cell pellet was kept overnight at -20°C prior to purification.

2.3.5.2 SDS-PAGE

SDS-PAGE was performed as described by Laemmli (1970) with modifications. Either 10 % or 12 % separating gel with 4 % stacking gel were used (Bio-Rad mini gel size of 8.6 x 6.8 cm). To de-nature the proteins, samples were mixed with 2x Loading Buffer and heated at 95°C for three minutes before loading onto the prepared gel. The electrophoresis condition was 80V for fifteen minutes or until loading dye reached the junction between stacking and separating gels. Then the condition was changed to 120-200 V and continued until the loading dye reached the bottom of the gel.

2.3.5.3 Immunoblotting

After electrophoresis, the separated protein samples were transferred to PVDF membrane (GE Healthcare) using a Mini Trans-Blot® Electrophoretic Transfer Cell (Bio-Rad) as described in manufacturer's protocol with modifications. Equilibration of the protein gel and PVDF membrane is required before the transfer. The membrane was sized to 8 x 6 cm and soak in methanol for ten seconds. The membrane was washed with milliQ water for ten minutes. The protein gel, the pre-wet membrane, two gel-size pieces of GB005 Whatmann blotting paper and fibre pads were equilibrated in transfer buffer (25 mM Tris, 192 mM glycine) for at least ten minutes. Transfer cassettes were assembled as shown in and set in Figure 17 and the transfer cell was set as described in the manufacturer's protocol. The transfer conditions were 100 V, 350 mA (constant) for an hour using a Bio-Rad power supply. After transfer, the membrane was blocked for an hour at room temperature in TBST buffer (50 mM Tris-HCL pH 7.5, 150 mM NaCl, 0.1% v/v Tween™ 20) containing 1 % BSA. The membrane was washed twice for ten minutes each time with TBST buffer. The blocked membrane was incubated with a nickel (Ni²⁺) activated derivative of horseradish peroxidase (HRP), HisProbe-HRP (Thermo Scientific) TBST with at concentration 1:5,000 for an hour. The membrane was washed four times, for ten min each time, with TBST buffer.

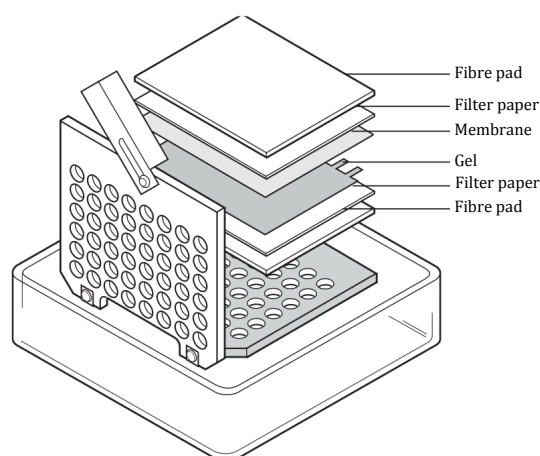


Figure 17 Schematic depiction for the assembly of the protein gel and the membrane for Western blotting (From Bio-Rad Mini Trans-Blot® Electrophoretic Transfer Cell manual).

To detect the signal, incubation with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) for five minutes and wrapped in a SaranWrap. The membrane was placed in a film cassette before the signal was detected using a Hyperfilm ECL (GE Healthcare). After exposure for 0.5, 1, 5 or 30 minutes (as required), the films were developed using a compact X4 automatic X-ray film processor (Xograph, UK). Films were scanned using an HP scanner to create digital images.

2.3.5.4 His-tagged protein purification using Ni-TED column

The cell pellet (3 g) was thawed on ice and lysed in 10 ml of Lysis Buffer. Then the lysate was treated with lysozyme (1 mg ml^{-1}) for thirty minutes. Cells were disrupted by sonication on ice for ten minutes. A cleared protein lysate was obtained after centrifugation at $30,000 \times g$ for thirty minutes and loaded onto a Ni-TED 2000 column (Macherey-Nagel). In order to remove unspecific bound proteins, 8 ml of Washing Buffer were applied to the column. Proteins of interest were eluted with 3 ml of Elution Buffer and separated on SDS-PAGE.

Chapter 3 Bioinformatics analysis

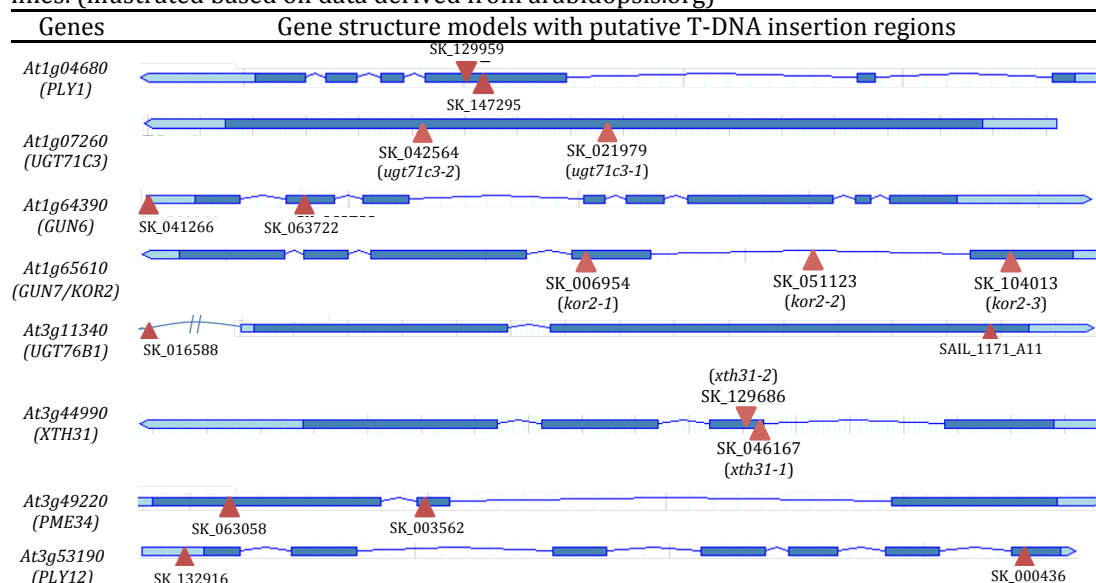
3.1 Functional analysis of candidate genes and proteins

This section describes bioinformatics analysis of gene and protein structure related to the criteria that was concerned while selected the strategies for biological and biochemical characterisation of the target genes. All eight-candidate genes, which had been selected at the start of the project, were analysed with the same methods and described here in order to perform the preliminary experiments before further selection of the final list candidates as described in the aims of the study (see 1.7). Before performing any experiments to characterise both biological and biochemical functions of the target genes, *in silico* analysis using bioinformatics and data-mining tools available were employed. The target genes were cross-referenced with data deposited on public database as described (see 2.1).

To determine the biological functions of the candidate genes, a reverse genetic approach using T-DNA insertion lines was employed. Predicted structures for the candidate genes were obtained from the TAIR database and are shown with an overview of selected insertion lines for the candidate genes with predicted insertion sites are shown in red triangles (Table 5). The selection criteria for these insertion lines were the location of the insertion in an exon and/or as close as possible to the transcriptional start codon.

As shown in Table 5, *At1g04680 (PLY1)* appears to be complex as there are six exons and five introns. The two selected *ply1* mutants show predicted T-DNA insertion positions both in the third exon. While the structure of *At1g07260 (UGT71C3)* is more simple as no intron presents. The both selected *ugt71c3* mutants show predicted T-DNA insertion positions. For *At1g64390 (GUN6)* that shows opposite orientation from those two, it has seven introns. The selected *gun6* mutants show predicted T-DNA insertion positions just after the start codon for the first line and in the second intron for the other one.

Table 5 Structure of the candidate genes showing untranslated regions (light blue box), exons (blue box) and introns (solid blue line). Transcription orientation of each gene is indicated by arrow-shaped ends. Red triangles represent putative T-DNA insertion sites for selected insertion lines. (illustrated based on data derived from arabidopsis.org)



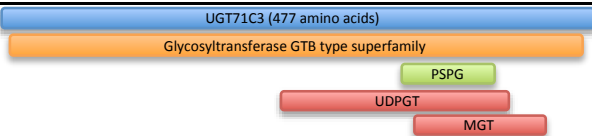
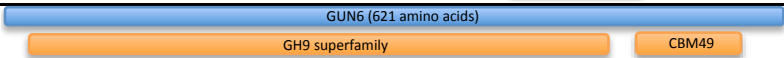
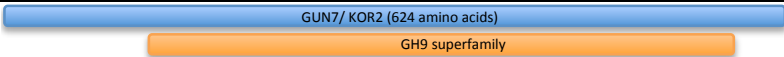
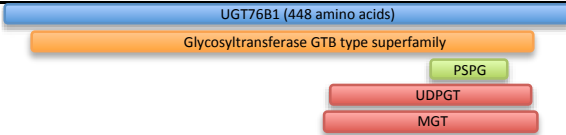
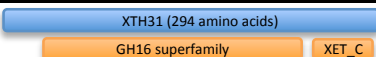
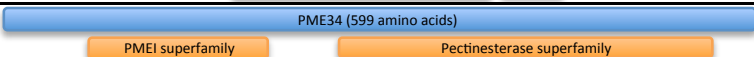
The *At1g65610* (*GUN7/ KORRIGAN2, KOR2*) appears to contain four introns with three selected *kor2* mutants showing predicted T-DNA insertion positions in the first exon, the first intron and the second exon. For *At3g11340* (*UGT76B1*), the gene structure appears to be rather simple as there is only a short intron presents. Two selected *ugt76b1* mutants show predicted T-DNA insertion sites in the untranslated region before the start codon for the first line and another site is nearly the end of the coding region. *At3g44990* (*XTH31*) appear to have four exons with three introns. Both two selected *xth31* mutants have putative T-DNA insertion positions in the second intron adjacent to each other. *At3g49220* (*PME34*) appears to have three exons and two introns, one short and another long one. The two selected *pme34* mutants show the putative T-DNA insertions in the first and second exons. *At3g53190* (*PLY12*) appears to have seven exons where two putative T-DNA insertion positions are in the first and the last exon for first and second *ply12* mutants, respectively. Further information about selected mutants related to putative protein structure and predicted conserved domains will be described below.

Amino acid sequences for the candidate proteins were aligned against the non-redundant protein database (<http://blast.ncbi.nlm.nih.gov>). Sequence analysis identified several conserved domains diagnostic for certain enzyme families (Table 6).

PLY1 appears to have pectate lyase C super family domains. Both selected *ply1* mutants (SK_129959 and SK_147295) have supposedly inserted before the predicted active sites or metal binding domains.

UGT71C3 possesses several conserved domains previously identified in proteins belonging to the UDP-glucosyl transferase (UGT) and macroside glycosyltransferase (MGT) families. In plant UGTs, the plant secondary product glycosyltransferase (PSPG) box ([FW]-x(2)-[QL]-x(2)-[LIVMYA]-[LIMV]-x(4,6)-[LVGAC]-[LVFYAHM]-[LIVMF]- [STAGCM]-[HNQ]-[STAGC]-G-x(2)-[STAG]-x(3)-[STAGL]-[LIVMFA]-x(4,5)-[PQR]- [LIVMTA]-x(3)-[PA]-x(2,3)-[DES]-[QEHNR]) is highly conserved (Gachon et al., 2005). The PSPG box is located at the C-terminus of the protein in close proximity to a nucleotide-sugar substrate recognition and binding site (Osmani et al., 2009). The available genetic and protein data suggest that the T-DNA insertion (SK_042564, *ugt71-2*) is probably located in the PSPG box. For SK_021979 (*ugt71-1*) the insertion site is before the PSPG domain.

Table 6 Query enzymes with predicted conserved domains based on amino acid sequences. Amino acid sequences are represented in purple, a specific hit for which the query sequences displayed highest identities is shown in red, the multi-domains are shown in grey other non-specific hits and domains for superfamilies are in pale red and light blue. Plant secondary product glycosyltransferase (PSPG) domains are in green. (adapted from blast.ncbi.nlm.nih.gov)

Enzymes	Conserved domains
PLY1	
UGT71C3	
GUN6	
GUN7/ KOR2	
UGT76B1	
XTH31	
PME34	
PLY12	

GUN6 contain glycosyl hydrolase family 9 (GH9) domains with a carbohydrate-binding module 49 (CBM 49) at the C-terminus. The CBM49 module is active during hydrolysis of crystalline cellulose. Two selected *gun6* mutants are SK_041266 and SK_063722 that show putative T-DNA insertion sites in the 5'UTR and the second exon before all predicted active sites.

KOR2 contains a domain characteristic for the glycosyl hydrolase family 9 (GH9) with KOR2 being a member of subclass A (ATGH9A2) whose members contain a trans-membrane domain at the N-terminus. The selected insertion lines for *KOR2* are SK_006954 (*kor2-1*) and SK_051123 (*kor2-2*) which have the putative insertions in the extracellular domain before all predicted active sites and SK_104013 (*kor2-3*) where the putative insertion is in the trans-membrane region.

UGT76B1 appears to have conserved domains including UDP-glucosyl transferase (UDPGT) and macroside glycosyltransferase family (MGT) with the PSPG box at the C-terminus of the protein in where nucleotide-sugar substrate recognition and binding sites are located. Two selected *ugt76b1* mutants are SAIL_1171_A11 and SK_016588. The former mutant shows putative T-DNA insertion site after PSPG box and the latter one shows putative insertion site in the promoter region.

The annotation of XTH31 suggests that the enzyme contains an O-glycosyl hydrolase domain, making it a member of glycosyl hydrolase 16 family (GH16) (blast.ncbi.nlm.nih.gov). Members of this family hydrolyse the glycosidic bond between two or more carbohydrates, or between a carbohydrate and a non-carbohydrate moiety. It also possesses an XET_C domain (xyloglucan endo-transglycosylase (XET)) at the C-terminus.

The PME34 enzyme appears to have a pectin methyl esterase inhibitor (PMEI) and a pectin esterase domain. For selected *pme34* mutants, putative insertion site for SK_063058 is in the PMEI domain, while the putative insertion site is located in the pectinesterase domain for SK_003562.

PLY12 appears to have pectate lyase C super family domains. The selected *ply12* mutants, SK_132916 shows putatively T-DNA insertion site in the 5'UTR and for another mutant SK_000436, the putative T-DNA insertion is located

within the last exon beyond all predicted active sites and the metal binding domains.

3.1.1 Addressing possible functional redundancies

The available information suggest that the insertions should lead to loss of function of the candidate genes. However, there is the possibility that functional redundancy could mask the mutant phenotypes especially for genes with a large number of family members. Therefore, *in silico* analysis was performed to compare expression profiles of other family members to the profiles of the target genes as a first step towards addressing possible functional redundancy problems.

To identify other family members that could possibly cause redundancy in the knockout lines to be isolated, gene expression profiles for all family members were compared with the expression profiles of the corresponding gene of interest. Microarray expression data deposited in the Genevestigator database (www.genevestigator.com) was employed for this purpose. Candidates showing similar expression patterns in terms of plant development and plant anatomy are listed in Table 7.

Table 7 Overview of family members for the different candidate genes that show similar expression patterns with respect to developmental stages and plant organs based on transcript levels.

Genes of interest	Similar expressed genes	Gene expression pattern	
		Developmental stages	Plant organs
<i>At3g44990</i>	<i>At4g28850, At5g48070</i>	Seedlings	whole plant
<i>At3g49220</i>	<i>At2g26440, At4g33220</i>	young rosettes	leaves
<i>At3g53190</i>	<i>At4g24780, At5g63180</i>	seedlings/ young rosettes	hypocotyl/ root tips
<i>At1g04680</i>	<i>At4g24780</i>	Seedlings	whole plant
<i>At1g64390</i>	<i>At2g32990, At1g75680</i>	Seedlings	whole plant
<i>At1g65610</i>	<i>At1g22880, At1g48930</i>	Seedlings	roots
<i>At1g07260</i>	<i>At5g05900, At5g53990</i>	young rosettes/ seedlings	roots
<i>At3g11340</i>	<i>At3g21770, At1g50580</i>	young rosettes/ seedlings	roots

3.1.2 Analysis of native signal peptides

The candidates were crossed referenced with the universal protein database, UniProt (<http://www.uniprot.org/>) to obtain all available information regarding predicted functions, active sites, signal domains and post-translational modifications. In addition, the proteins were analysed using SignalP 3.0 server (<http://www.cbs.dtu.dk/services/SignalP/>) to predict native signal peptides independently because it is crucial to remove the native signal domains before expressing the proteins fused to the yeast secretion factor provided in the pPICZ α expression vectors. Prediction results are presented in Table 8.

Table 8 Prediction of signal peptides for candidate enzymes.

Proteins	UniProt annotation		SignalP prediction	
	Signal peptide	Positions	Signal peptide	Positions
XTH31	Signal peptide	1-20	Signal peptide	1-20
PME34	Trans-membrane	46-66	Signal anchor	1-21
PLY1	Signal peptide	1-24	Signal peptide	1-24
PLY12	Signal peptide	1-20	Signal peptide	1-20
GUN7/KOR2	Trans-membrane	80-100	Non-secreted	-
GUN6	Signal peptide	1-22	Signal peptide	1-22
UGT71C3	N/A	-	Non-secreted	-
UGT76B1	N/A	-	Non-secreted	-

The analysis results suggest that UGT71C3 and UGT76B1 are non-secretory proteins, while PLY1, PLY12, GUN6 and XTH31 are secreted to the cell wall. For PME34, a trans-membrane region at the N-terminus was predicted, suggesting that it could be an anchor protein. Interestingly, there is a contradiction between the two databases for KOR2. From UniProt, it has been predicted that KOR2 has an eighty amino acid cytoplasmic domain with a single trans-membrane domain and an extracellular domain. However, those predicted domains and their positions within KOR2 amino acid sequences were not shown on the predicted results analysed by the signal region prediction tool on SignalP server.

3.2 Bioinformatics based analysis of *KORRIGAN2*

KOR2 exhibits typical glycosyl hydrolase family 9 (GH9) domains and is a member of the subclass A, whose members contain a trans-membrane domain at the N-terminus. The predictions for the secondary structure for KOR2 (623 amino acids) are shown in Figure 18.

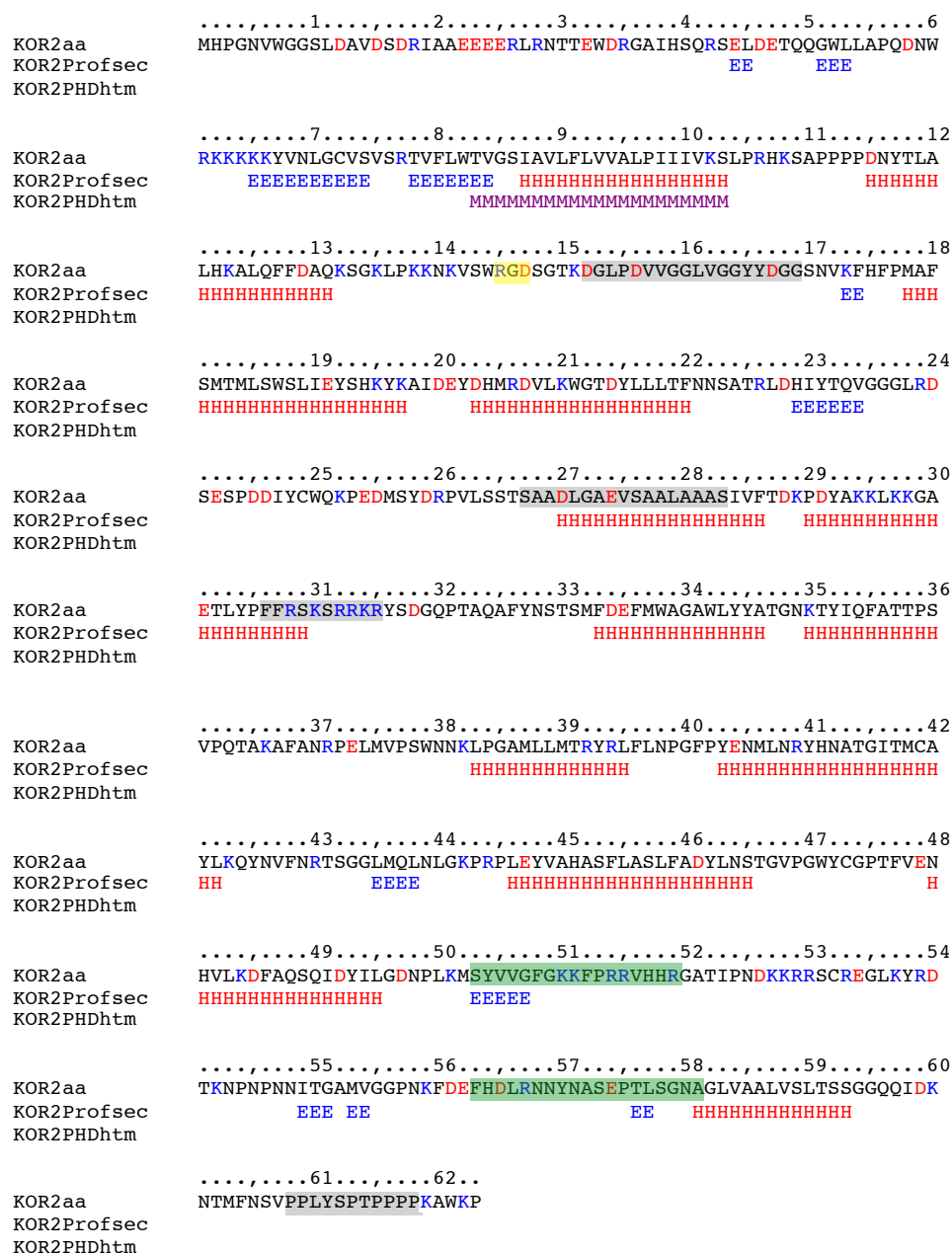


Figure 18 Predicted secondary structure of putative KOR2. KOR2aa = amino acid sequences, KOR2Profsec = predicted secondary structure for KOR2 with residues in α -helix regions (red H) and extended β -strand regions (blue E) colour coded. KOR2PHDhtm = profile network prediction Heidelberg where the helical trans-membrane region (purple M) is shown. Highlighted in yellow is a cell adhesion motif. Low complexity regions (LCRs) are highlighted in grey. Two glycosyl hydrolase family 9 active signature regions are highlighted in green (www.predictprotein.org; www.uniprot.org)

The secondary structure prediction results suggest that the protein has thirteen α -helix regions (34.67 % of amino acid residues being α -helix structure), while most parts of the structure (57.95 % of the sequences) were predicted as loops. Only 7.38 % being predicted as extended β -strand. The PHDhtm prediction identified a trans-membrane region between amino acid 84 and 101 (shown in purple M in Figure 18) with the C-terminus being cytoplasmic (1-83 AA) and the N-terminus forming an extra-cellular domain (102-623 AA). Other predicted sites affecting the activity of the protein are shown in Appendix 2 including ten N-glycosylation sites, ten N-myristoylation sites and more than ten possible phosphorylation sites. The low complexity regions where false positive annotation could occur are highlighted in grey. The Arginyl-glycyl-aspartic acid (RGD) motif represented a cell adhesion motif is highlighted in yellow (Mølhøj, Jørgensen, Ulvskov, & Borkhardt, 2001). Two glycosyl hydrolase family 9 active signature regions are highlighted in green (Figure 18). The selected insertion lines for *KOR2* were SK_006954, where the putative insertion is in the trans-membrane region, SK_051123 with the putative insertion located in the intron and SK_104013, which has predicted an insertion in the extracellular domain in front of all predicted active sites.

3.3 Bioinformatics based analysis of *XTH31*

The secondary structure of *XTH31* was analysed via the PredictProtein service (www.predictprotein.org) with general information for signature region derived from UniProt database (www.uniprot.org). Results were depicted as shown in Figure 19. The putative protein sequences sizes 293 amino acids with 8.53% of the sequences predicted as α -helix in two regions (Figure 19). The majority of the sequences were predicted as loop structures while other 32.76% are extended β -strands in twelfth separated regions. Only one membrane-helix was predicted (purple M,

Figure 19) at the N-terminus. Three low complexity regions are highlighted in grey. Predicted β -glucanase active region is highlighted in green. There are

two disulphide-bonds predicted at (238, 246 sites) and (280, 293 sites). From PredictProtein service, other predicted post-translational modification sites are shown in Appendix 2 including eight phosphorylation sites and two N-myristoylation sites. For protein localisation, it was predicted to be an extra-cellular protein involved in secretion pathway.

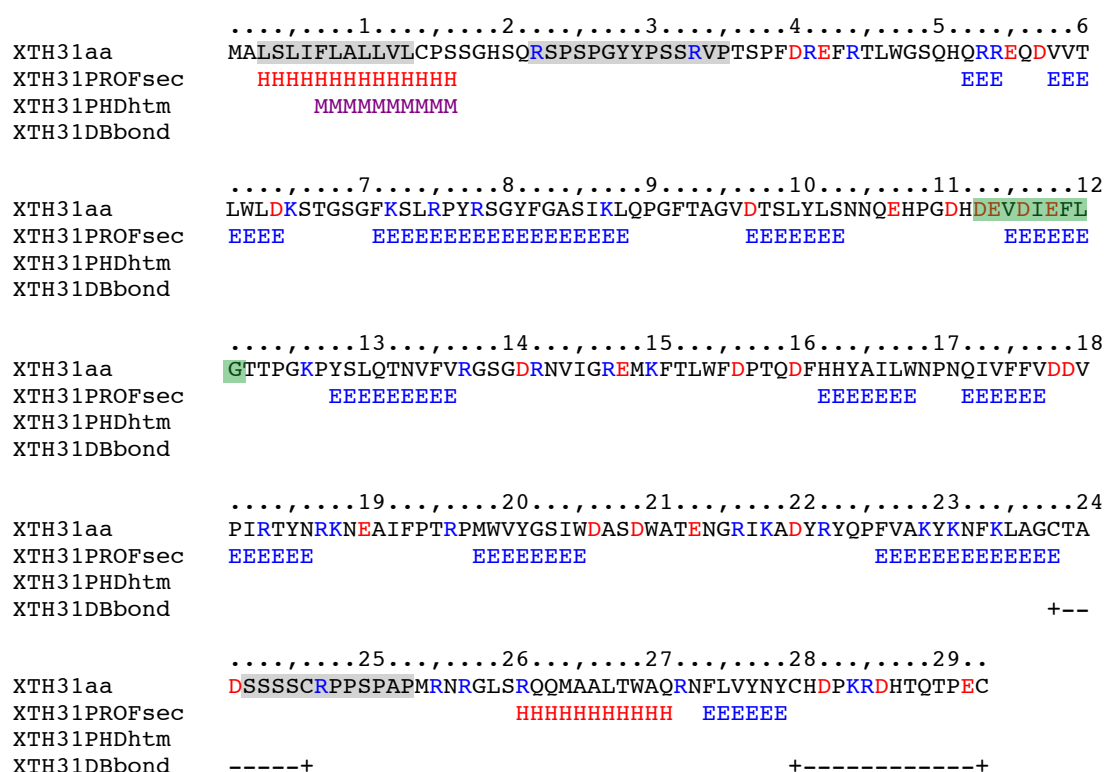


Figure 19 Predicted secondary structure of putative XTH31. XTH31aa = amino acid sequences where low complexity regions (LCRs) are highlighted in grey. XTH31Profsec = predicted secondary structure for XTH31 that the residues in α XTH31P regions (red H) and extended β -strand regions (blue E) are predicted. XTH31PHDhtm = profile network prediction Heidelberg where the helical trans-membrane region (purple M) is shown. XTHDBbond = predicted disulfide bonds. Predicted β -glucanase active region is highlighted in green (www.predictprotein.org; www.uniprot.org)

3.4 Bioinformatics based analysis of *UGT71C3*

The predicted secondary structure for UGT71C3 is shown in Figure 20. The protein is supposed to consist of 476 amino acids with 36.97 % forming an α -helix, 12.61 % an extended β -strand from twelve strands and 50.42 % loops. Profile network predictions suggest, two membrane domains exist (118 – 135 AA and 140 – 157 AA) the stretch between 1 – 117 AA forms the first outside

region, 136-139aa being located on the inside and 158 – 476 AA representing the second outside region.

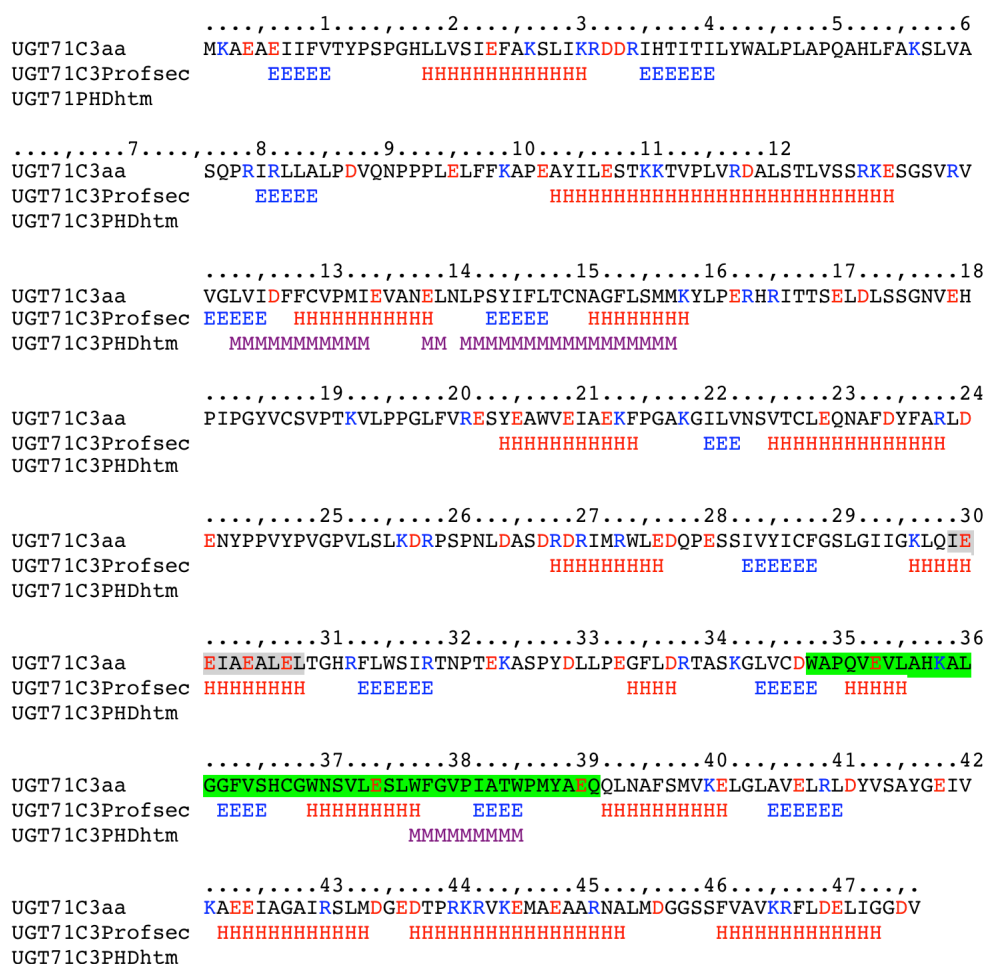


Figure 20 The predicted secondary structure of UGT71C3. UGT71C3 aa = amino acid sequences where low complexity regions (LCRs) are highlighted in grey. UGT71C3Profsec = predicted secondary structure for UGT71C3 with residues α -helix regions (red H) and extended β -strand regions (blue E) predicted. UGT71C3PHDhtm = profile network prediction Heidelberg where the helical trans-membrane region is highlighted with (purple M) is shown; and the UDP-glycosyltransferase signature region being highlighted in green (www.predictprotein.org; www.uniprot.org).

The data also suggests the existence of a low complexity region highlighted in grey (299 - 308 AA) and identifies the UDP-glycosyltransferase signature region highlighted in green (348 – 391 AA). While no disulfide bonds were predicted, there are more than thirteen phosphorylation sites and four N-myristoylation sites predicted (Appendix2).

In addition, it was shown for UGT71C3 that it possibly contains four predicted UDP-glucose binding site/regions (www.UniProt.org). Together with

the predicted secondary structure from PredictProtein service, all binding site/regions predicted are: Ser₂₉₀ (in 15th loop), Ala₃₄₉-Gln₃₅₁ (19th loop), His₃₆₆-Glu₃₇₄ (21st loop) and Tyr₃₈₈-Gln₃₉₁ (23rd loop).

3.5 Discussion and summary

3.5.1 Functional analysis of candidate genes and proteins

Bioinformatics analysis for eight candidate gene sequences, regarding to selection of the target genes showed at least two mutants on independent alleles were employed. The mutants that showed putative insertion sites the exon or close to start codon were targeted for the most possible disruption of the expression of the target genes. Functional analysis for predicted gene and protein structures of the target genes was performed and the alignment results showed that although gene structures for most of the candidate genes are complex, the protein coding sequences showed mostly conserved domains along the sequences (see Table 6). These indicate the high sequence similarity within enzymes family or super family for each target genes, highlighting the important of the enzymes from these families to plants as these domains are well conserved across plant species. In addition, this also suggests high possibility of functional redundancy by family members or other related genes, challenging the characterization of these target genes that any phenotypes could be masked by gene redundancy in the mutants. Therefore, expression profile of each target gene was compared with family members to observed possible family member that showed similar expression pattern in term of spatial and time for the expression. The results in Table 8 suggested there are some other family members that possibly have functional redundancy when the target genes are knocked out. However, to confirm this *in silico* analysis, double and triple or more sequential knock-out of these family members need to be performed. Subsequently, comparative analysis of single and double and other mutants' phenotypes need to be observed. In addition, further rescue of the knock-out mutants with the target gene will confirm if the gene is involved or causing the phenotypes observed.

In parallel with the functional analysis of gene and protein structures for selection of the T-DNA insertion mutants and some clues for the gene functions based on domain analysis results, amino acid sequences of the target genes were also analysed in order to guide strategy planning for molecular cloning of these genes. The alignment of amino acid of the candidate genes suggested that their structures are complex and post-translational modifications are required for proper folded-expressed-proteins. Therefore, *P. pastoris* was selected as an expression system as described (see 1.7 and Appendix 3.2). The sequences where signal peptide or trans-membrane regions were predicted, these regions were excluded from the cloning. However, the cloning strategies were revised after unsuccessful expression of the target enzymes in *Pichia* and *E. coli* system was employed as the results shown in other following chapters. Below, further bioinformatics analysis of three final selected genes will be discussed.

3.5.2 Bioinformatics based analysis of *KORRIGAN2*

From phylogenetic analysis, the closest related family members to KOR2 are KOR1 and KOR3 that display 55% and 54% amino acid identities (data not shown), suggesting that these three members could be key enzymes in different particular metabolic pathways. However, the amino acid sequence identity also suggested that these enzymes could divergently evolve from each other and have different functions. Thus, no likely functional redundancy on each other would occur. The selected insertion lines for KOR2 are SK_006954 (*kor2-1*) and SK_051123 (*kor2-2*) which have the putative insertions in the extracellular domain before all predicted active sites and SK_104013 (*kor2-3*) where the putative insertion is in the trans-membrane region. For *kor2-1* and *kor2-2*, in the case where truncated protein is translated, the enzyme would lose the active sites. For *kor2-3*, truncated protein would be a short sequence of peptides for trans-membrane region only.

It has been reported that *KOR2* is the only one member in *Arabidopsis* membrane bound endo- β -1,4-glucanase (EGase) family that has a cell adhesion motif (RGD; Figure 18) (Mølholm et al., 2001). The RGD motif is a cell adhesion motif found in many animal cells which can be recognised by a subset of integrins, cell-

surface receptors for cell adhesion molecules (D'Souza et al., 1991; Ruoslahti & Pierschbacher, 1987). Functions of integrins for animal cells are involved in the interaction both in cell to cell and cell to matrix communications, both in and out of two ways signal transmission which are important for cell adhesion, migration and interaction with the pathogens (Coppolino & Dedhar, 2000; Etzioni, 1999). In plants, immunolocalisation techniques employing polyclonal antibody against animal integrins have been carried out and the integrin-like proteins in *Clara*, *Arabidopsis* and maize have been identified via these techniques (Ciuffetti et al., 1997; Katembe et al., 1997; Knepper et al., 2011; Lü et al., 2007). Possible functions of plant integrin-like proteins has been studied in *Arabidopsis* and maize (Knepper et al., 2011; Lü et al., 2007; Lü et al., 2007¹). Knepper et al. (2011) studied an *Arabidopsis* integrin-like transmembrane protein, NON-RACE-SPECIFIC DISEASE RESISTANCE1 (NDR1) that the *ndr1-1* mutant plants showed alteration of the mRNA expression in response to pathogen effectors and other biotic and abiotic stresses, and alteration of membrane-cell adhesion in plasmolysed cells were observed. Thus, possible physiological roles of NDR1 in *Arabidopsis* has been suggested as a component in signal transduction triggering plant resistance mechanisms and its role in cell integrity maintenance via the membrane-wall adhesion (Knepper et al., 2011). Other roles of plant integrin-like proteins in osmotic stress induced ABA biosynthesis have been studied in *Arabidopsis* and maize using synthetic GRGDS pentapeptide containing RGD motif and found that the peptide contains RGD domain inhibited ABA biosynthesis while the GRGDS homolog without RGD domain showed no effect (Lü et al., 2007; Lü et al., 2007¹). The RGD domain has been found in a toxic protein so called Ptr ToxA from *Pyrenophora tritici-repentis*, a fungal pathogen to wheat (Ciuffetti et al., 1997; Meinhardt & Cheng, 2002). The study on the role of RGD domain suggested the role of this motif in host cell death probably via interaction of the RGD domain with integrin-like proteins of the host cells since the synthetic non-toxic RGD tripeptide could compete with the Ptr ToxA and reduced electrolyte leakage from the host cells, hence, reduce cell death (Meinhardt & Cheng, 2002). However, no evidences for others plant proteins containing RGD motif has been established to date. These available evidences for possible functions of RGD domain and its interaction with integrin-like proteins

suggest possible roles of KOR2 in membrane-wall adhesion, roles in signalling transduction in response to osmotic stress and pathogen probably via ABA induction. However, more supportive experimental data is needed to reveal functions of the KOR2 in plants.

Annotation data from UniProt show three functional domains of these are a cytoplasmic domain (positions 1-79), a trans-membrane domain (positions 80-100) and an extracellular domain (positions 101-623). Three predicted active sites are His-517, Asp-565 and Glu-574 with nine N-glycosylation sites (Asn-116, Asn-221, Asn-328, Asn-349, Asn-412, Asn-429, Asn-464, Asn-548 and Asn-571). In addition to the data derived from UniProt, the secondary structure protein prediction results for KOR2 (see Figure 18) suggest that the protein structure consists of 34.67 % α -helix, 57.95 % loops and Only 7.38 % extended β -strand. The PHDhtm prediction identified the C-terminus being cytoplasmic region, followed by a trans-membrane region where the first α -helix region is (shown in purple M in Figure 18) and the N-terminus forming an extra-cellular domain. Other predicted sites affecting the activity of the protein are shown in Appendix 2 including ten N-glycosylation sites, ten N-myristoylation sites and more than ten possible phosphorylation sites. These predicted post-translational modification sites indicate rather complex structure of KOR2, which means heterologous expression of this gene should be conducted in eukaryotic host. Ideally, over-expression of the target gene would assure proper folded active enzyme being expressed.

3.5.3 Bioinformatics based analysis of *XTH31*

The candidate gene *At3g44990*, encoding XTH31, is a member of subfamily GH16 in the CAZY database (Cantarel et al., 2009) of the 1,338 putative XTH plant proteins available in the database, 33 members derived from *Arabidopsis thaliana*. Plant XTHs are divided into three subgroups based on the phylogeny generated from amino acid sequence similarities. XTH31 is a member of subgroup III with six other XTHs (XTH 27-33). It has been suggested that the group III XTHs should be further divided into two subgroups, A and B, based on amino acid sequences and catalytic activities (Eklöf and Brumer, 2010).

Group III A, to which nasturtium XTH (TmNXG1, Mark et al., 2009) and Azuki bean XTH (VaXGH, Tobuchi et al., 2001) belong, has xyloglucan hydrolase (XEH) activity. Only AtXTH31 and AtXTH32 are designated into this group thus, XEH activity has been predicted for these two *Arabidopsis* enzymes (Baumann et al., 2007). Group III B, to which AtXTH27, SlXTH5 and HvXTH8 belong, possesses predominantly XET activity (Campbell and Bramm, 1999; Saladie et al., 2006 and Kaewthai et al., 2010). Amino acid identity analysis via BlastP suggests that *At2g36870* encoding XTH32 is the closest family member to XTH31 (65% amino acid identities). Although this subgroup is quite small with only seven family members, no structural information of the *Arabidopsis* members has been revealed so far. Based on detailed kinetic data, it has been suggested that the XEH activity is the result of a gain of function over the evolution of GH16 ancestral gene for the functions related to germination, fruit ripening and rapid cell expansion (Baumann et al., 2007). Hence, the most related structural data could only be gained from enzymes derived from other plants such as a nasturtium XTH, TmNXG1 and a hybrid aspen XTH (PttXET16-34). The first 3-D structures for XET and XEH active sites have been published for the Hybrid aspen protein PttXET16-34 (Johanssen et al., 2004 for XET activity and Baumann et al., 2007 for XEH activity).

Based on available data, the general structure for the GH16 family has been suggested to display the β -jellyroll fold that generated specificity for highly branched substrates (Eklöf & Brumer, 2010). The 3-D structure of PttXET16-34 and TmNXG1 have been studied and it has been indicated that the highly conserved Cysteine residues is located at the C-terminus, where two disulfide bonds are formed, is important for protein stabilisation (Johanssen et al., 2004 and Baumann et al., 2007). In addition, the N-glycan structure plays crucial roles in the folding and stabilisation of the proteins (Johanssen et al., 2004 and Baumann et al., 2007). Notably, for group III B members, it was found that the N-glycosylation site is shifted toward the C-terminus (Kallas et al., 2005). From available structural data of the characterized enzymes, it was found that the catalytic motif of XTH is (H/W/R)-(D/N)-**E**-(I/L/F/V)-**D**-(F/I/L/M)-**E**-(F/L)-(L/M)-G with catalytic residues in bold and the most frequent residues underlined (Eklöf and Brumer, 2010). Compared with the candidate gene XTH31,

whose catalytic domain is ¹¹²HDEVDIEFLG¹²¹, the catalytic amino acids could be predicted. The predicted catalytic domain of XTH31 is similar to the catalytic region of β -1,3-1,4-glucanase from *Bacillus macerans* (WDEIDIEFLG) (Borriss et al., 1990). It has been proposed that the first glutamate residue (E) in this region is critical for cleavage of 1,4- β -glycosyl linkages (Baumann et al., 2007). From UniProt database, it was predicted that Glu-114 is the catalytic nucleophile and Glu-118 is the proton donor, while the function of Asp-116 is still unknown but is was predicted to modulate the ionisation state of the catalytic nucleophile based on the hydrogen bonding pattern (Johanssen et al., 2004). Notably, the proton donor residue, Glu-118, is the site where putative insertions are predicted for SK_046167 (*xth31-1*) and SK_129686 (*xth31-2*). Therefore, there is a high possibility for these insertion lines that the expression/function of the target gene might be affected.

3.5.4 Bioinformatics based analysis of *UGT71C3*

Arabidopsis UGTs family is a large and diverse family, having roles in many metabolic pathways in plants due to their glycosylation activities over wide range of aglycons including nucleic acids, proteins, sugar, lipid, plant secondary metabolite, xenobiotics and other various molecules (Caputi et al., 2012; Paquette et al., 2003; Ross et al., 2001; Yonekura-Sakakibara & Hanada, 2011). Most of plant UGTs contain Plant Secondary Product Glycosyl Transferase conserved domain (PSPG box; [FW]-x(2)-[QL]-x(2)-[LIVMYA]-[LIMV]-x(4,6)-[LVGAC]-[LVFYAHM]-[LIVMF]- [STAGCM]-[HNQ]-[STAGC]-G-x(2)-[STAG]-x(3)-[STAGL]-[LIVMFA]-x(4,5)-[PQR]- [LIVMTA]-x(3)-[PA]-x(2,3)-[DES]-[QEHNR]) that is believed to be a signature for the enzymes mediating glycosyltransferase of plant secondary metabolites (Gachon et al., 2005; Ross et al., 2001). However, some UGTs from *Arabidopsis* have been reported to catalyse glycosylations of phytohormones and monolignols (Caputi et al., 2012; Yonekura-Sakakibara & Hanada, 2011). For examples, UGT71B6 is involved in regulation of ABA metabolism (Priest et al., 2005). For the screening of *Arabidopsis* glycosyltransferase family 1 members over two enantiomers of ABA ($\pm$ ABA), it has been discovered that eight UGTs exhibited glycosylation activity over ABA

(E.-K. Lim et al., 2005). However, only UGT71B6 that showed enantioselective activity over (+)-ABA, the natural form of ABA (Oritani & Kiyota, 2003). Another *Arabidopsis* UGT that found to be involved in regulation of BA metabolism is UGT73C6 (Husar et al., 2011). It was found that the over expression *UGT73C6* mutants showed stunt phenotypes and high level of BL-23-*O*-glucoside (Husar et al., 2011). The observed phenotypes from over expression *UGT736* suggested that the gene is involved in glycosylation of BL, however, functions of glycosylated BL are still need to be investigated. Although not much data is available for UGT71C3, the sequence analysis detected a triterpene UGT from *Medicago* (MtUGT71G1) was identified, showing 71 % amino acid identities (data not shown). MtUGT71G1 protein has already been characterised and crystal structure models are available (www.uniprot.org/uniprot/Q5IFH7) (Shao et al., 2005). Although MtUGT71G1 showed higher glycosylated efficiency to quercetin than triterpene, other genetics and metabolomics evidences suggested the role of MtUGT71G1 in terpenoid biosynthesis (Achnine et al., 2005). This may allow prediction of 3D structures and possibly the function of UGT71C3. However, the alignment of UGT71C3 against plant UGTs, suggests that the carbohydrate binding domains (C-terminal), where PSPG box is located, are highly conserved but the N-terminal domains for receptor recognition shows high variation among the different enzymes (BlastP, data not shown). This suggests a variety of aglycone receptors for UGTs.

Like other members of *UGT71C* clade, gene structure of *At1g07260* (*UGT71C3*) is simple and possesses only one exon region without any introns. Two knockout lines that show T-DNA insertions in the exon were selected from the NASC database (SK_021979 or *ugt71c3-1* and SK_042564 or *ugt71c3-2*). For *ugt71c3-1* it was showed that putative T-DNA insertion is before the PSPG box while the putative insertion position for *ugt71c3-2* is in the middle of the box. Thus, any truncated proteins expressed as the products should not be able to recognise and bind to the nucleotide-glucoside donors.

Chapter 4 Functional characterization of KORRIGAN2

4.1 Glycosyl hydrolase family 9

Glycosyl hydrolase enzymes are known for such an important group of enzymes whose functions were previously well characterised in microorganisms (Hématy et al., 2009; Henrissat et al., 2001). Available first completed genome of a plant model, *Arabidopsis*, allowed predictions of plant orthologues and suggested more abundant carbohydrate-active enzymes in plants than other organisms (Henrissat et al., 2001; Liepman et al., 2010). In the carbohydrate active enzymes database (CAZy; www.cazy.org), the predicted enzymes have been categorised based on amino acid sequence similarity. Several enzymes have been functionally characterised, therefore, allowing predictions regarding the activities of other enzymes that exhibit high degrees of identity. For the plant glycosyl transferase family 1 (GT1) three thousand fifty three members are currently listed in the CAZy database with hundred seven members of UGTs that have been derived from *Arabidopsis* (<http://www.p450.kvl.dk/gst.shtml>). There are five hundred thirty entries for the glycosyl hydrolase family 9 with twenty-five entries being assigned to *Arabidopsis*. The Glycosyl hydrolase family 9 (GH9) or cellulase family 9 (Cel9) has been categorised into three subclasses based on their amino acid sequence similarity (Figure 21) (Urbanowicz et al., 2007^b).

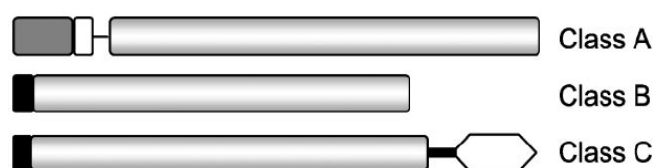


Figure 21 Structural variations among 3 subclasses of the plant GH9 family. The schematic illustrates the modular structure: cytoplasmic domain (dark grey), transmembrane domain (white), signal sequence (black), GH9 catalytic domain (light grey), linker region (thick black line), carbohydrate binding module (CBM) (hexagon). Structural subclasses are represented by GH9 members from tomato; SlGH9A1 (Subclass A, U78526), SlGH9B1 (Subclass B, U13054) and SlGH9C1 (Subclass C, AAD08699) (adapted from Urbanowicz et al., 2007^b).

At the N-terminal end, adjacent to the catalytic domain, subclass A has cytoplasmic and transmembrane domains while subclasses B and C have signal sequences (Urbanowicz et al., 2007^a). The subclass C can be distinguished from

subclass B by the carbohydrate-binding module (CBM) at the C-terminus, which is required for degradation of crystalline cellulose (Urbanowicz et al., 2007^b). It has been shown that the endo-1,4- β -glucanase KORRIGAN exhibits functional conservation between gymnosperms and angiosperms and is required for proper cell wall formation in gymnosperms (Maloney et al., 2012).

Even though little is known about activity and functions of plant GHs compared with the microbial enzyme GHs data available, a model for the tree subclasses of plant GH 9 family has been purposed (Figure 22). This could be inferred to predicted function and activities of the members from each subclass (Lopez-Casado et al., 2008). For example, GH 9 subclass A, that available experimental data suggested the involvement of the members with cellulose biosynthesis (Molhoj et al., 2002; Somerville, 2006), the locations of the enzymes are adjacent to cellulose synthase complexes (Lopez-Casado et al., 2008). However, more information for functions, specific activities and structures of the members still required for elucidating the roles of these enzymes in cell wall metabolism.

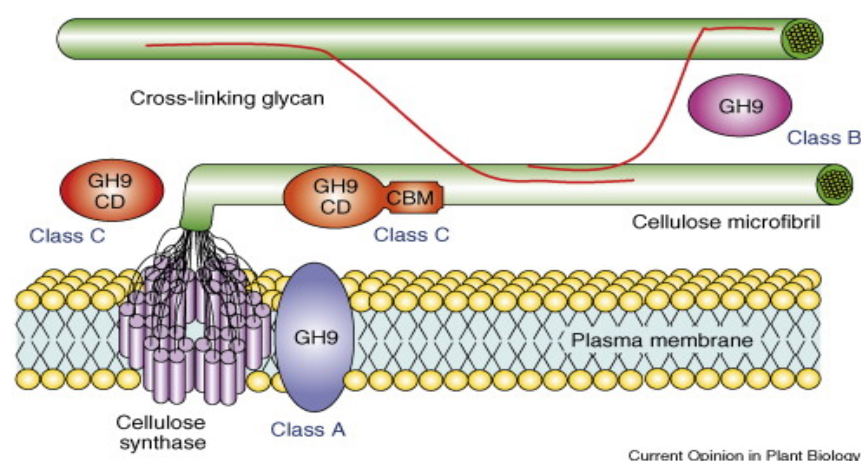


Figure 22 Schematic model showing plant GH9 enzymes that are membrane bound (Class A), associated with the cell wall through a CBM (Class C) or more freely soluble (Class B) (adapted from Lopez-Casado et al., 2008).

4.2 *Arabidopsis* membrane bound EGASEs

There are twenty-five *Arabidopsis* genes encoding GH9 family members with only three genes (*KOR1*, *KOR2* and *KOR3*) encoding membrane-bound proteins of the endo-(1-4)- β -D-glucanase (EGASE) family (Molhoj et al., 2001^a). It

was reported that mutation of KOR1 and the aspen homolog PttCel9A1 cause a low crystalline cellulose phenotype and affect secondary wall formation (Takahashi et al., 2009). KOR1 was implicated in cellulose biosynthesis (Lane et al., 2001; Sato et al., 2001; Szyjanowicz et al., 2004), cell elongation (Sato et al., 2001), cytokinesis (Lane et al., 2001) and loss of function causing compositional changes in pectic polysaccharides (His et al., 2001).

One of the candidate genes characterised in this study was *At1g65610*, which encodes KORRIGAN2 (KOR2, AtGH9A2). KOR2 is a member of GH9 subclass A. Analysis of *KOR2* gene expression using a GUS promoter reporter construct suggested that KOR2 may be involved in strengthening the cell wall at the junctions between leaf petioles and stems and between flower petioles and receptacles (Molhoj et al., 2001^b).

4.3 Heterologous expression of KORRIGAN2

In order to characterize the biochemical properties and functions of the KOR2 from *Arabidopsis*, heterologous expression of the KOR2 protein using other organisms was performed. Bearing in mind, the bioinformatics based analysis for KOR2 resulted in the predicted post-translational modification sites often N-glycosylation sites, ten N-moyristoylation sites and more than ten possible phosphorylation sites with a predicted transmembrane region. Therefore, eukaryotic hosts were considered and *Pichia pastoris* that had been showed the success in expression of other EGASEs were selected (Bauer et al., 2006). *P. pastoris* strain GS115 was used as a host with the expression constructs derived from pPICZ for non-secreted and pPICZ α for secreted expression constructs. However, all attempts at achieving expression were unsuccessful (see Appendix 3). Therefore, an *E.coli* based expression strategy was initiated. Two versions of the cDNA for KOR2 either containing the whole coding region or only the extracellular domain had been cloned into the expression vector for *Pichia*.

Since these constructs had been sequenced and confirmed to be error free they were used as the template to amplify fragments for expression in *E.coli* using TOPO[®] TA Cloning[®] (Invitrogen) (Figure 23).

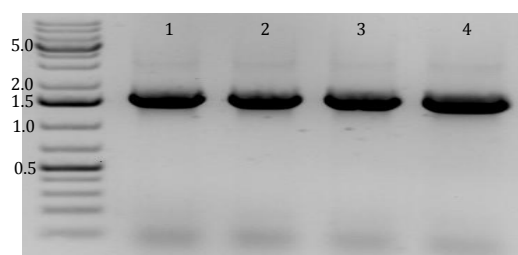


Figure 23 PCR to amplify extracellular domain for KOR2 flanking with *Bam*HI and *Not*I restriction sites using Phusion® DNA polymerase. DNA ladder shown as kb.

The insert fragments were then extracted from agarose gel and cloned into pCR2.1 TOPO vectors as described in manufacture's protocol. Figure 24 shows colony PCR from 9 pCR2.1-KOR2 clones.

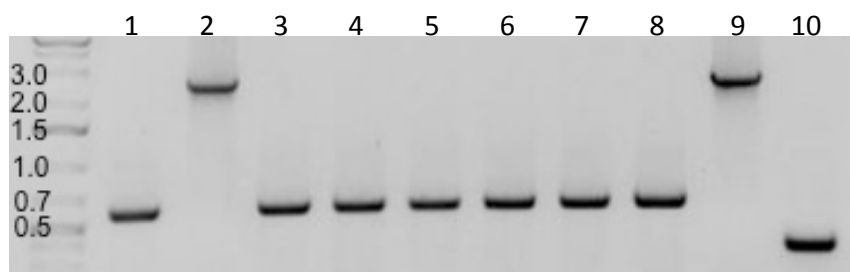


Figure 24 Colony PCR to isolate DH5α harbouring pCR2.1KOR2 (1-9) and DH5α with empty plasmid (10). Size of DNA ladder fragments shown in kb.

Colony PCR results shown in Figure 24, show that clones 2 and clone 9 are positive for integration of the expression construct. Cells from both clones were grown overnight for plasmid preparation and restriction digest using *Bam*HI and *Not*I. The results for restriction digest are showing below (Figure 25).

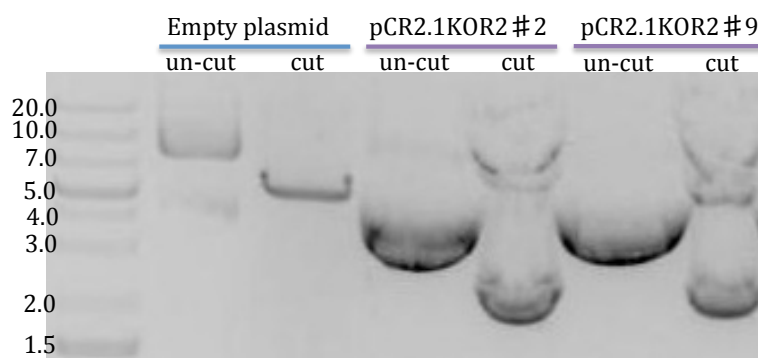


Figure 25 *Bam*HI-*Not*I double digest for pCR2.1KOR2 construct. Size of DNA ladder fragments shown in kb

The insert fragment (1.5 kb) from clone 9 (pCR2.1KOR2#9) was extracted from the gel, digested with *Bam*HI and *Not*I and ligated into pET28b⁺ vector. The ligation reactions of pET28KOR2 were transformed into DH5 α cells using heat shock, plated out and selected using kanamycin. Colony PCR reactions were performed using forward and reverse T7 primers, which should amplify 0.6 kb fragments from the empty plasmid and 2.1 kb fragments containing the expression construct. Results showed that almost all tested clones are positive for the expression construct (exceptions clone number 2 and 7) (Figure 26).

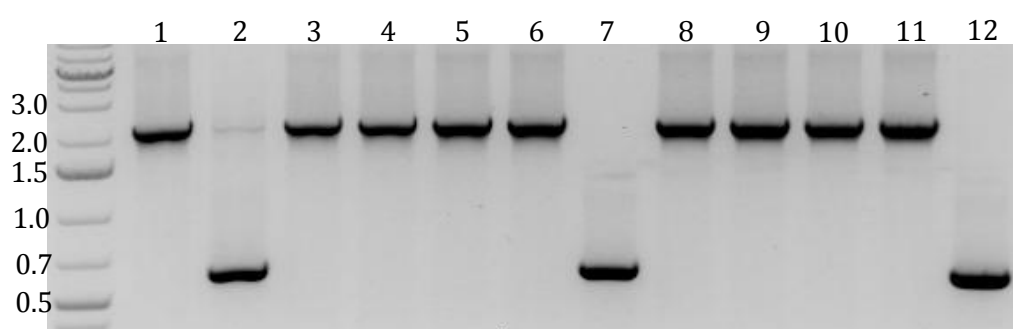


Figure 26 Colony PCR for DH5 α pET28KOR2 (1-11) and DH5 α harbouring pET28 empty plasmid (12) are shown. Size of DNA ladder fragments shown in kb

Three positive clones were selected and grown up for plasmid preparation as described before. Plasmid DNA was digested with *Bam*HI and *Not*I to confirm that correctly sized insert fragments are still present (Figure 27). The results showed that all three clones contain the expression construct. Clone pET28KOR2#9 was selected for sequencing. The sequenced pET28KOR2 construct was then used for further experiments.

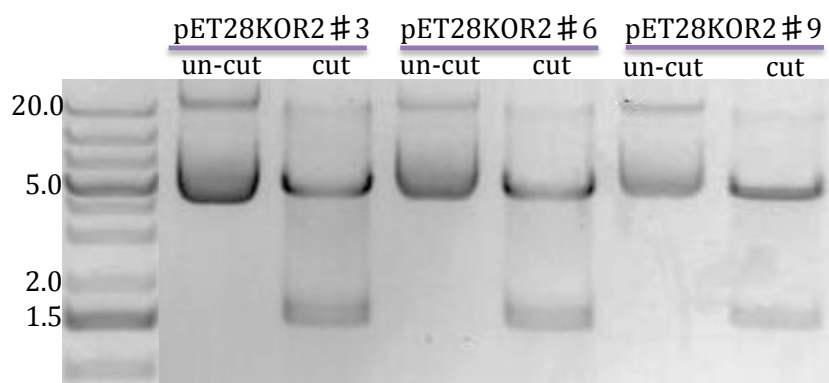


Figure 27 *Bam*HI-*Not*I double digest for pET28KOR2 construct. Size of DNA ladder fragments shown in kb

The expression construct for KOR2 was heat shock transformed into three different *E.coli* strains: BL21, SoluBL21™ and BL21 (DE3). After antibiotic selection and confirmation by colony PCR, positive strains were used for expression studies. Figure 28 shows SDS-PAGE analysis of protein preparations for four different clones from three different strains with the predicted size for KOR2 being 60 kD. However, the results from SDS-PAGE showed no dominant protein bands for the expected size.

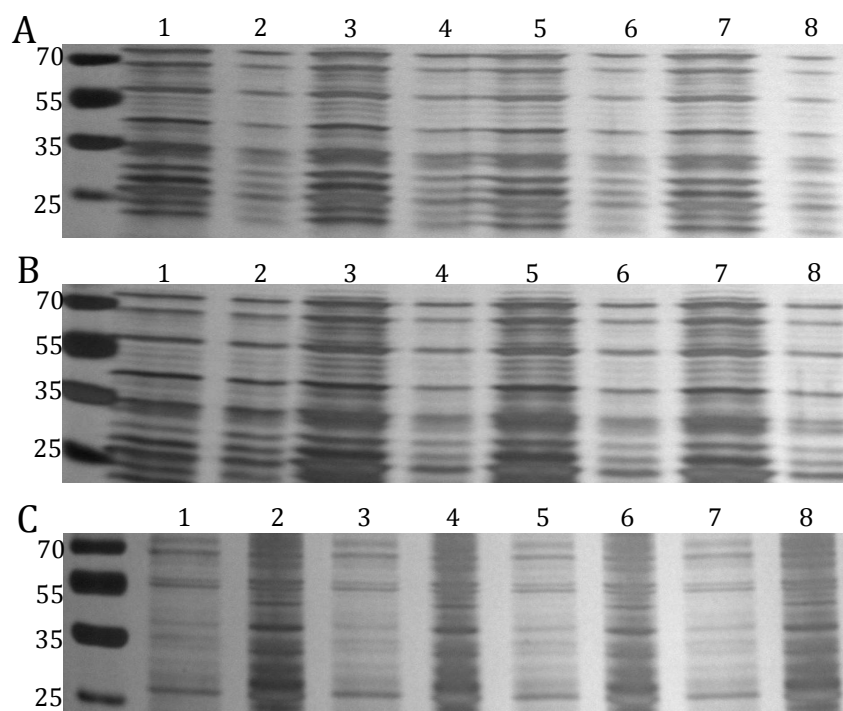


Figure 28 SDS-PAGE analysis for KOR2 expressed in *E.coli* strain BL21 (A), SoluBL21™ (B) and BL21 (DE3) (C). Four different clones were tested; clone1 = lane 1 and 2, clone2 = lane 3 and 4, clone3 = lane 5 and 6 and clone4 = lane 7 and 8. Samples from 0 hour (lane 1, 3, 5 and 7) and 6 hours (2, 4, 6 and 8) of IPTG induction are shown. Size of protein marker shown in kD

In parallel with the SDS-PAGE analysis for heterologous expression KOR2, immune-dot blot to detect target protein with histidine tags was performed (Figure 29). Although the signals obtained were faint, it was shown that target proteins could be induced with the strongest signals from culture induced with 0.5 mM IPTG for six hours and 1 mM IPTG for eighteen hours. However, there were also some signals from non-induced culture, indicating leaking of T7 expression that could cause the expression of the target gene that was cloned under T7 promoter.

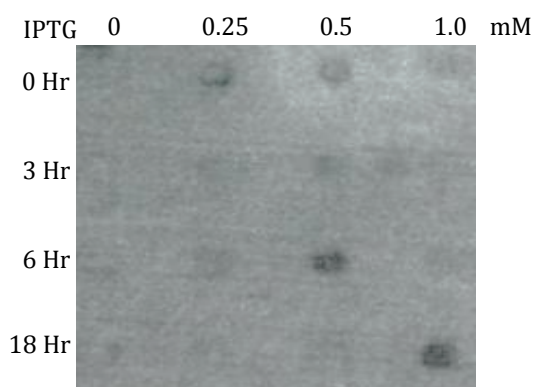


Figure 29 Immuno dot blot for KOR2 expressed under 0, 0.25, 0.5 and 1 mM IPTG induction for 0, 3, 6 and 18 hours.

Purification of the target protein was performed using Protino® Ni-TED Packed Columns. Proteins from each step of purification were isolated and analysed using SDS-PAGE (Figure 30).

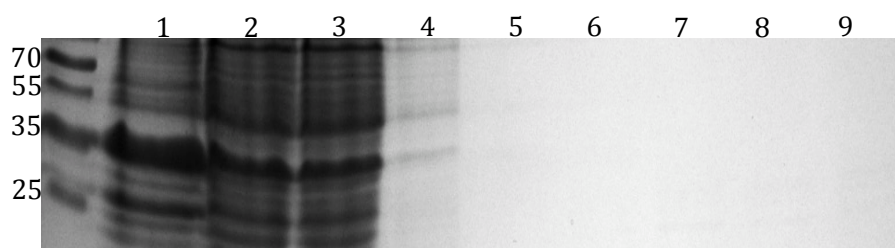


Figure 30 SDS-PAGE for KOR2 after purification for His-tagged proteins. Shown above are cell lysate (1), soluble proteins (2), Flow-through (3), 1th Wash (4), 2nd Wash (5), 1st Elution (6), 2nd Elution (7), 3rd Elution (8) and 4th Elution (9). Size of protein marker shown in kD

Although some signals were detected from crude culture by immuno dot blot, SDS-PAGE results show no purified proteins obtained for KOR2 (also no signal from Western blot, data not shown). This could be a result of leaking expression of KOR2 toxic for the host; therefore, no successful expression clones could be obtained. To achieve heterologous expression of the target gene, cloning strategies need to be optimised for appropriate hosts and expression conditions.

4.4 Isolation of homozygous T-DNA insertion plants

Reverse genetics tools were employed to determine the biological functions of the candidate genes. T-DNA insertion lines for *KOR2* were obtained from the stock centre and homozygous plants were isolated using a PCR-based strategy with primers specific for *KOR2* to amplify genomic fragments and a

combination of a primer specific for *KOR2* and a primer specific for the T-DNA insert to detect the T-DNA insertion in *KOR2*. Figure 31 shows exemplary results with plants homozygous for T-DNA insertions in *KOR2* showing amplification of fragments with gene and T-DNA specific primers but not with the gene specific primers alone.

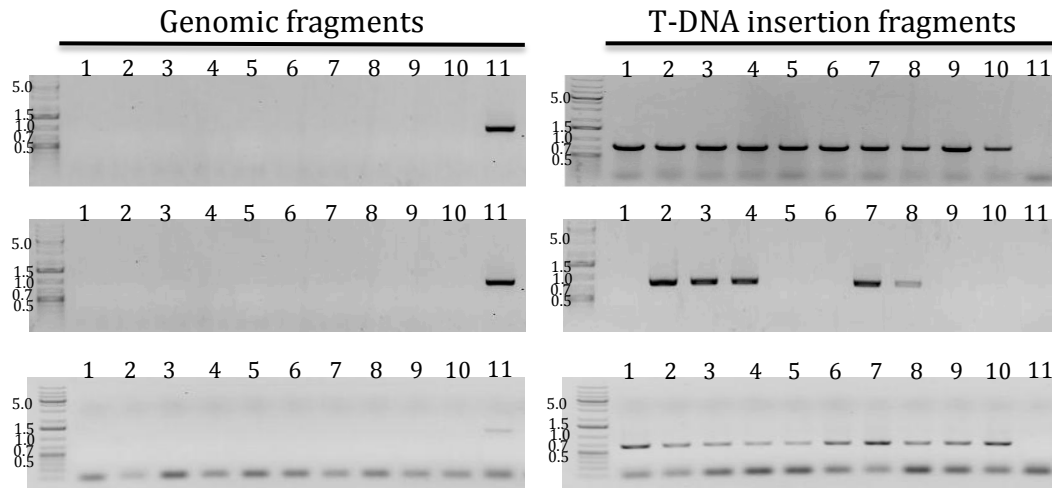


Figure 31 PCR-based screening to isolate homozygous *kor2* mutant plants. Gene specific fragments and T-DNA fragments were amplified from 10 independent mutant plants (1-10) for each T-DNA insertion line and Col-0 (11). Size of DNA ladder fragments shown in kb

Fragments amplified with gene and T-DNA specific primers for the three insertion lines were sent for sequencing to confirm putative insertion positions in *KOR2* (Table 5).

4.5 Analysis of gene expression in the *kor2* alleles

Once homozygous plants for each insertion line had been isolated, seedlings from the different lines and the Col-0 wild type control were treated with isoxaben for 24 hours and total RNA was isolated for semi-quantitative RT-PCR. Primers specific for the *KOR2* coding region and for *UBQ10* (loading control) were used (Figure 32). *UBQ10* expression was detected in all samples (Figure 32, right panel). *KOR2* expression was detected in the Col-0 controls (Figure 32, left panel, samples 1 and 2). No *KOR2* expression was observed in any of the samples from the *kor2* seedlings confirming that the insertions lead to loss of *KOR2* expression (Figure 32, left panel, samples 3 to 8).

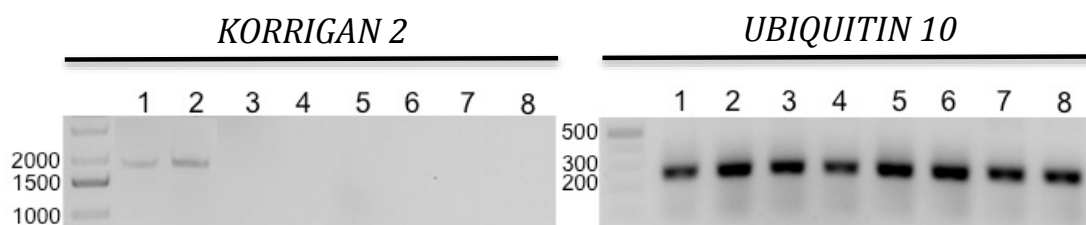


Figure 32 Semi-quantitative RT-PCR results for *KORRIGAN 2* expression in Col-0 and *kor2* seedlings. *UBIQUITIN 10* was used as a control. Col-0 shown in (1, 2), *kor2-1* in (3, 4), *kor2-2* (5, 6) and *kor2-3* (7, 8) with seedlings in 1, 3, 5, 7 mock and 2, 4, 6, 8 isoxaben-treated. Size of DNA ladder fragments shown in bp

4.6 Phenotypic characterization of *kor2* plants

After homozygous plants for *kor2* mutants were isolated and loss of *KOR2* expression confirmed with RT-PCR, plants were grown on soil under standard conditions as described along with the Col-0 wild type. Phenotypes of the plants from various developmental stages throughout the life cycle of plants were determined as described in materials and methods. The results for the time required at each growth stage of Col-0 wild type and *kor2* mutants are shown in Table 9. No obvious differences were observed between *kor2* and Col-0 plants.

Table 9 Growth stage-based analysis results for the wild type and *kor2* mutants.

Developmental stages	Col-0	<i>kor2-1</i>	<i>kor2-2</i>	<i>kor2-3</i>
Leaf development	4-5 days	4-5 days	4-5 days	4-5 days
2 rosette leaves > 1 mm in length	6 days	6-7 days	6-8 days	6-7 days
3 rosette leaves > 1 mm in length	7-9 days	7-8 days	8-9 days	8 days
4 rosette leaves > 1 mm in length	8-10 days	8-9 days	10 days	9 days
5 rosette leaves > 1 mm in length	11-12 days	10 days	11 days	10 days
6 rosette leaves > 1 mm in length	11-13 days	11 days	11-12 days	11 days
7 rosette leaves > 1 mm in length	13-14 days	12 days	13 days	12 days
8 rosette leaves > 1 mm in length	13-15 days	13-14 days	14-15 days	13-14 days
9 rosette leaves > 1 mm in length	14-15 days	15 days	15 days	15 days
10 rosette leaves > 1 mm in length	14-16 days	16 days	16 days	16 days
11 rosette leaves > 1 mm in length	17-18 days	17 days	17 days	17 days
12 rosette leaves > 1 mm in length	17-19 days	18 days	18 days	18 days
13 rosette leaves > 1 mm in length	18-19 days	19-20 days	19 days	19 days
14 rosette leaves > 1 mm in length	20-21 days	21 days	19-22 days	19-21 days
Rosette growth complete	21-22 days	20 days	20-22 days	20-21 days
Flower production	19-20 days	19-21 days	19-22 days	19-21 days
Senescence complete	60-65 days	58-63 days	60-65 days	60-63 days

Once all the plants were at the stage where the rosette growth completed, rosette diameter was measured. Results showed no significant differences among the wild type and three *kor2* mutants (Figure 33).

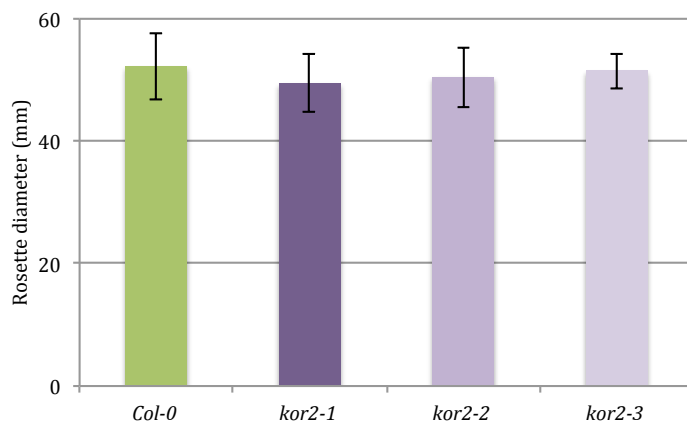


Figure 33 Rosette diameter measurement for *kor2* mutant plants that showed rosette growth complete stage. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average rosette diameter calculated from three biological replicates where the data was analysed using sixteen individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

At four weeks after sowing when the rosette growth is completed, biomass measurement was performed. The results showed that the mutant plants have significantly less biomass in both fresh and dry weight (Figure 34). Fresh weight of the *kor2* mutants is at least 20 % less than wild type while the dry weight is about 10 % less than wild type.

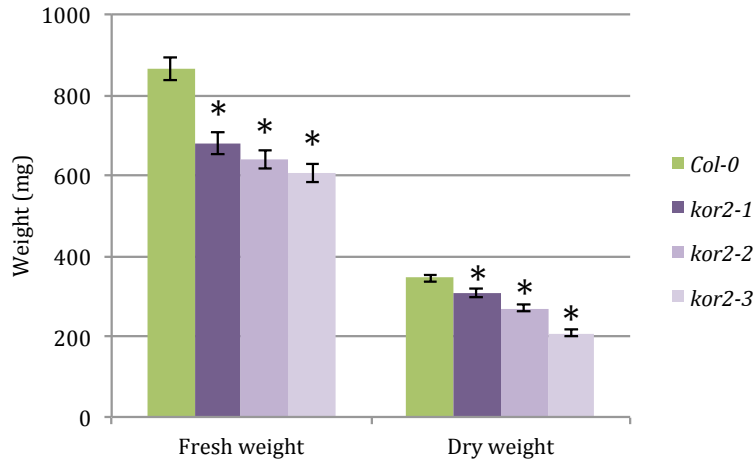


Figure 34 Biomass measurement for four week-old *kor2* mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight of either fresh or dried materials and calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

In addition to the measurement of biomass for plants at the growth stage where the rosette growth completed, the biomass for senescence plants were measure and the dry weights are shown (Figure 35).

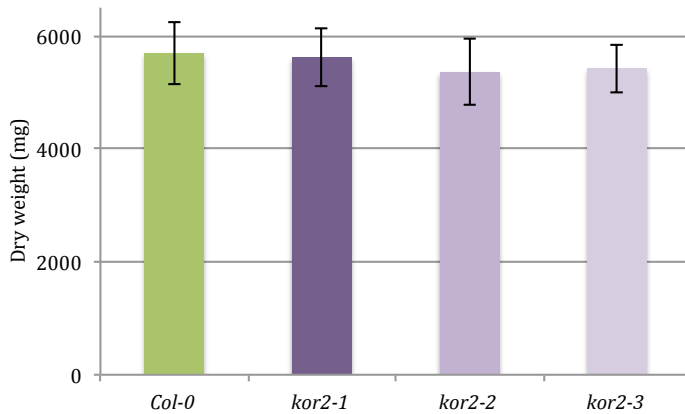


Figure 35 Biomass measurement for senescence *kor2* mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

Although significantly differences were observed for the mutant plants at the rosette-completed stage, no significantly differences was observed for senescence plants.

4.6.1 Characterisation of germination of mutant and wild type plants

The expression data for *KOR2* in the Genevestigator database suggests that the gene is expressed in seedling roots. Therefore, different phenotypic assays were performed to characterize growth and response to stress of the *kor2* radicle. Initially germination assays were performed under with *kor2* and Col-0 seedlings under standard growth conditions. Results are shown in Figure 36.

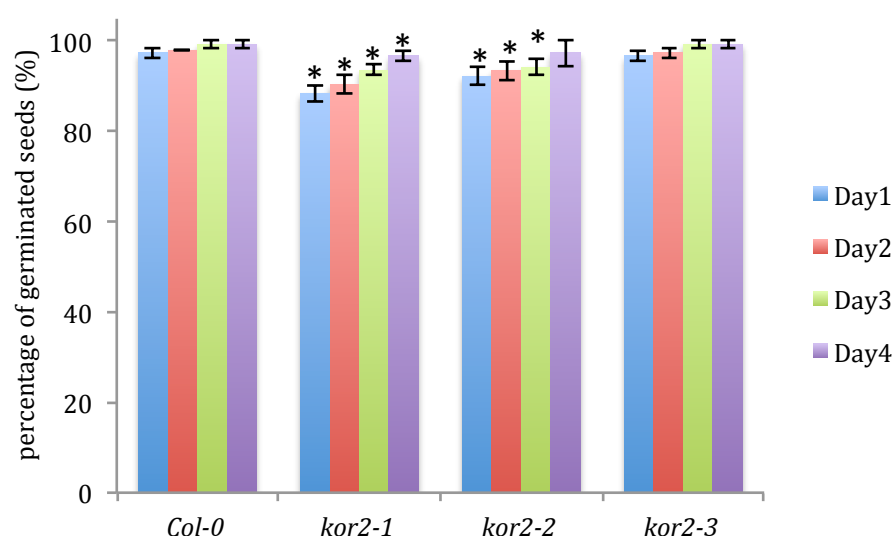


Figure 36 Results from germination assays for *kor2* mutants and Col-0 wild type. Sterile seeds were put on ½ MS plates and stratified at 4°C for forty-eight hours before grown in standard conditions as described. Seeds with radicals emerged were counted as germinated seeds after 1-4 days after plates were put under standard conditions. The values represent average percentage of germination calculated from three biological replicates where the data was analysed using fifty individual seeds $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

As shown in Figure 36, no significant differences were observed between *kor2-3* and Col-0. However, slightly differences were detectable between the germination rates of *kor2-2* and Col-0 on day 1 to day 3. For *kor2-1*, slightly but significantly differences were also shown between the germination rates of the mutants and Col-0 wild type from day 1 to day 4. These results indicated two loss of function mutants for *KOR2*, *kor2-1* and *kor2-2*, showed late germination phenotypes as less numbers of the emergence of radicles were observed. However, no germination phenotype was observed for *kor2-3*, therefore, root phenotypes in seedlings were studied as the results shown in 4.6.2.

4.6.2 Root growth of *kor2* mutants and wild type seedlings

As stated before that expression profiles from the Genevestigator database for *KOR2* suggests the gene is dominantly expressed in roots of seedlings. To follow up on the germination assays, root growth assays were performed for seedlings grown on plates (see methods in 2.2.5.3). The results shown in Figure 37, seedlings for all three alleles examined exhibit significantly longer roots than Col-0.

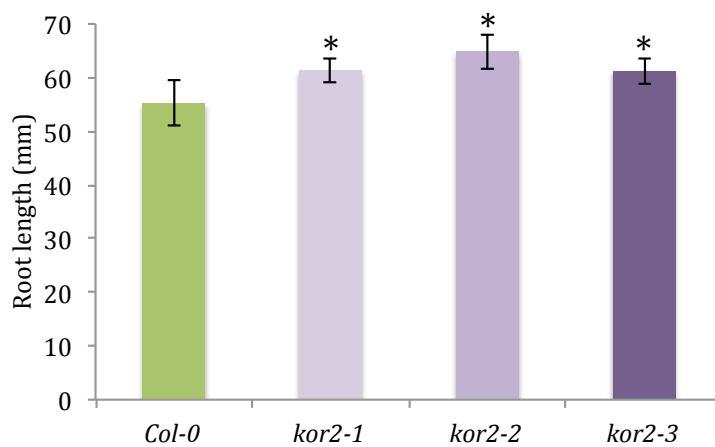


Figure 37 Roots of seven day-old *kor2* mutants and Col-0 wild type seedlings grown vertically on $\frac{1}{2}$ MS plates were captured. Photos were analysed using ImageJ software. The values represent the average of root length of seedlings calculated from three biological replicates where the analysis was performed using at least twenty individual seedlings $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test, $P \leq 0.05$).

As shown in Figure 37 from seedlings grown under standard growth conditions, results for root growth of seven-day-old seedlings of Col-0, *kor2-1*, *kor2-2* and *kor2-3* indicated slightly but significantly longer root phenotypes in mutant seedlings from all alleles compared with Col-0. These results showed here despite the fact that loss-of-function of *KOR2* mutants appeared to have negative effects and cause delay of radicle emergence during seed germination in *kor2-1* and *kor2-2* (Figure 36), this does not appear to restrict root elongation in seedlings but enhance primary root growth in mutant seedlings (Figure 37). Next experiment was designed to observe whether mutant seedlings would show different root phenotypes in response to a cellulose synthesis inhibitor, isoxaben.

To determine if loss of KOR2 activity affects the response of seedlings to isoxaben, root growth were measured from *kor2* and Col-0 seedlings grown on plates with ½ MS medium for four days. They were then transferred to plates with 0, 1, 2 and 3 nM isoxaben, their growth was measured every 24 hours over a period of three days. Figure 38 shows the differences in percentage of the root length grown on various concentration of isoxaben compared to the seedlings of the same age grown on the basal medium plates without isoxaben. The data suggested that *kor2* seedling roots are hypersensitive to isoxaben treatment.

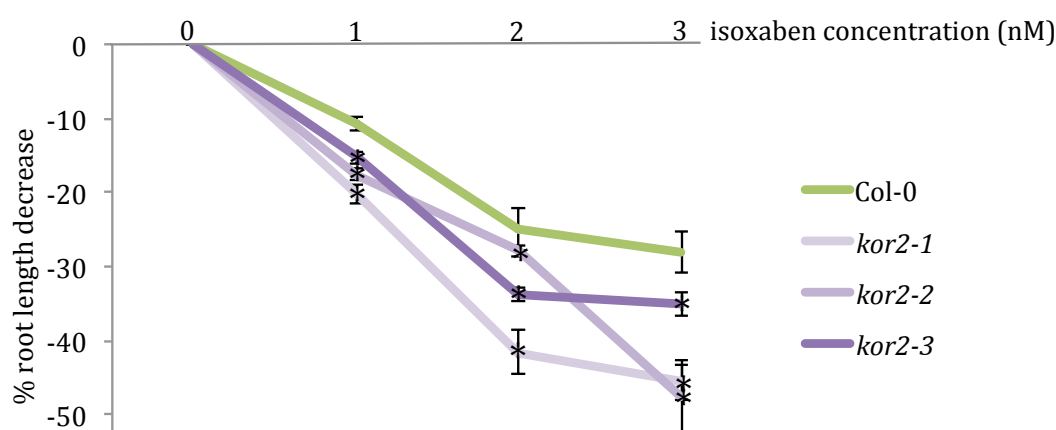


Figure 38 Percentage of root length decreased up on isoxaben treatment. Roots of seven day-old *kor2* mutants and Col-0 wild type seedlings grown vertically on ½ MS plates for four days before transferred to fresh ½ MS plates in presence or absence of 1, 2 or 3 nM isoxaben were analysed. Photos were captured as described and analysed using ImageJ software. Percentage of root length decreased were calculated from four biological replicates where the analysis was performed using at least ten individual seedlings $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type under the same treatment (student t-test, $n \geq 40$ per line and condition, $P \leq 0.05$).

To summarise germination and root growth phenotypes, the results suggested *KOR2* appeared to play roles in radicle emergence during seed germination even though one allele (*kor2-3*) did not showed significant delayed germinated phenotypes as in other two alleles (*kor2-1* and *kor2-2*). However, the gene was also appeared to have different regulations during root elongation of seedlings as the results showed that all mutants had longer roots than Col-0. These suggested the roles of *KOR2* in radicle or root growth with different regulatory functions during germination and root elongation. Moreover, results for root growth of seedlings grown under isoxaben treatment showed that loss-of-function mutants for *KOR2* exhibited hypersensitive responded phenotypes to isoxaben. These suggested prominent roles of the gene in the wall reinforcement during cell wall damaged by isoxaben.

4.6.3 Lignin deposition in seedlings upon exposure to isoxaben

It has been reported in previous study that isoxaben treatment causes cell wall damage and induces ectopic lignin deposition in the elongation zone of primary seedling roots (Hamann et al., 2009). Consequently, lignin staining in the roots of the *kor2* and Col-0 seedlings upon isoxaben treatment was performed. Seedlings of *kor2* and Col-0 were grown in liquid culture for six days before being treated with either DMSO (for mock treatment) or 600 nM isoxaben for 12 and 24 hours. The seedlings were then stained with phloroglucinol to detect ectopic lignin deposition. As shown in Figure 39 panel A, the elongation zone of *kor2* seedling roots appeared to exhibit more lignin deposition than the Col-0 controls (Figure 39 panel A, compare areas of red staining between *kor2* and Col-0 seedlings).

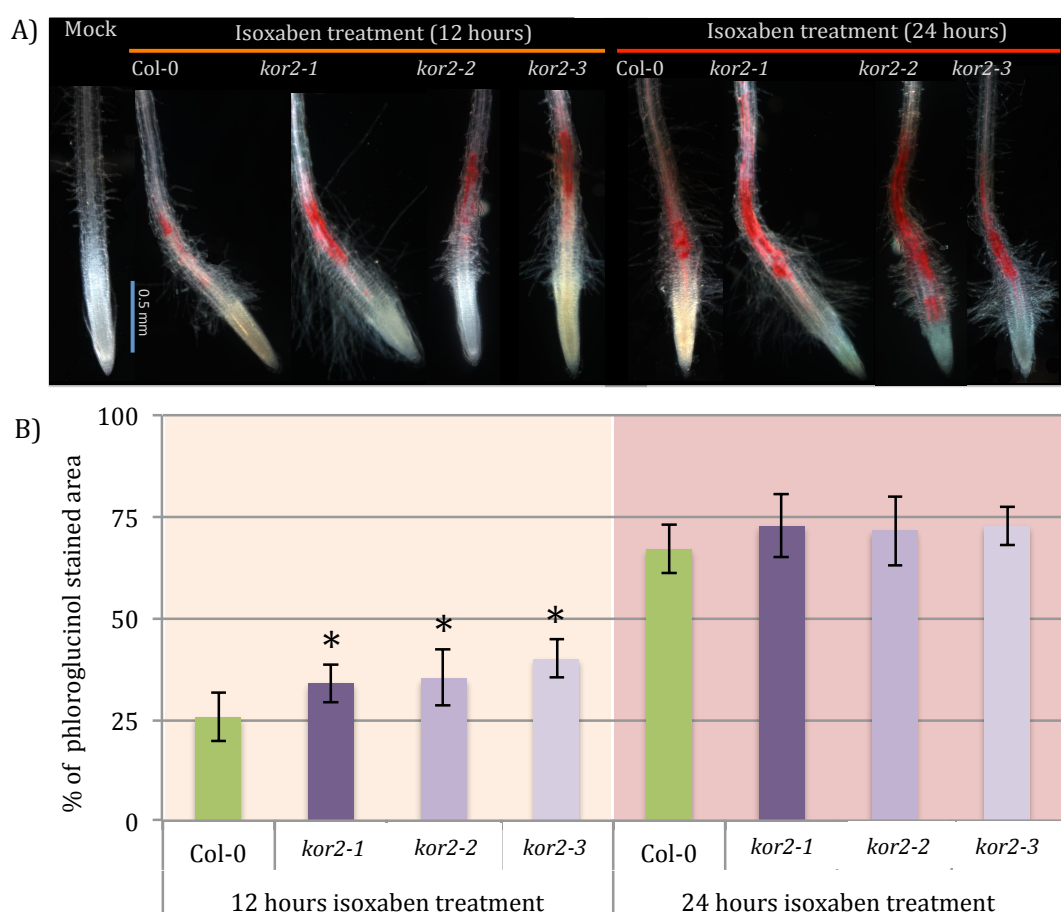


Figure 39 Phloroglucinol stained seven day-old *kor2* and Col-0 seedling roots upon 12 or 24 hour-isoxaben treatment. A) The representative results of the phloroglucinol staining. Scale bar is 0.5 mm, at least twenty seedling roots were analysed for each genotype/treatment. B) Percentage of the phloroglucinol staining for seven day-old *kor2* and Col-0 seedling roots after 12 or 24 hour-isoxaben treatment $\pm$ SE. * indicates significant difference (student t-test, $P \leq 0.05$).

As shown in Figure 39, panel A, the stained tissues for lignification can be observed in red and pink area and appeared in the elongation zone of roots. Although the picture had been captured immediately after staining, it was tricky to determine whether the light pink stained in some areas indicated less lignification or diffusion of the stained lignin from the deep red stained areas. Due to variations observed within the same line, at least twenty seedlings were stained. To compare the stained area of the roots quantitatively, the measurement was performed as described (2.2.5.4), percentage of stained area per root area appeared on the images was calculated and the results are shown in Figure 39, panel B. The results from 12 hour-treatment showed that all three alleles for *kor2* mutants appeared to show significantly more lignification in the elongation zone of the roots than in wild type. While no significant difference was observed for the seedlings treated for 24 hours (Figure 39, panel B).

In addition to the experiment on lignin staining from seedlings grown in liquid culture conditions, similar experiment had been performed also on seedlings grown on plates complemented with 1, 2 and 3 nM isoxaben. Phloroglucinol was used to detect lignin deposition in roots of the *kor2* and Col-0 seedlings after growth on ½ MS plates supplemented with different isoxaben concentrations for three days (Figure 40). Results from Figure 40 suggested different staining patterns of lignin deposition and root morphology in *kor2* seedlings compared to the Col-0 control after isoxaben treatment. It appeared in most roots of the mutants exhibited more swollen cells with smaller quiescent area than wild type. However, some variations among three *kor2* alleles were observed (compare *kor 2-1* with *kor2-2* and *kor2-3*).

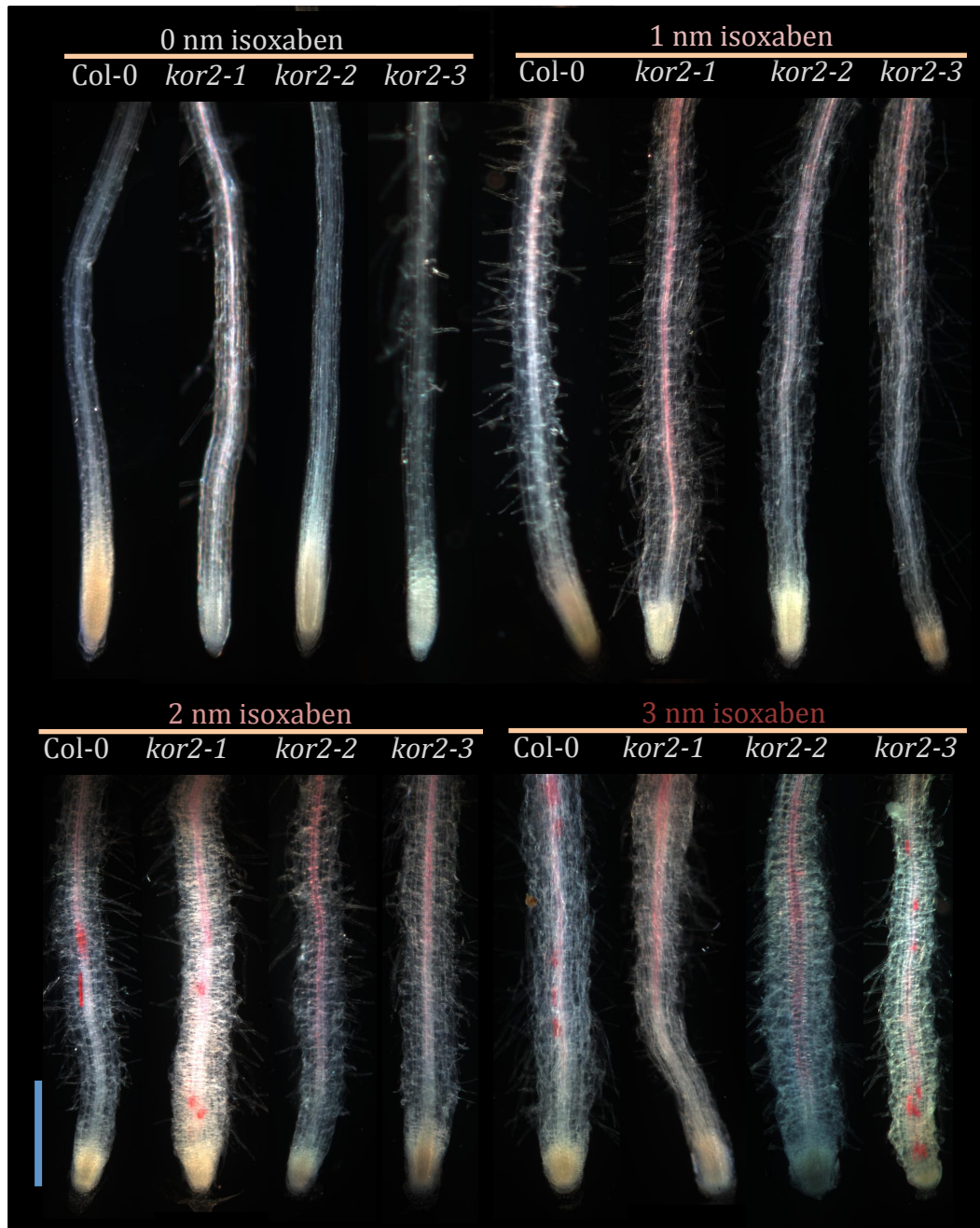


Figure 40 Phloroglucinol staining for lignin deposition in Col-0 and *kor2* seedling roots upon 0, 1, 2 and 3 nM isoxaben treatment for three days. Scale bar is 0.5 mm, for each genotype/treatment at least twenty seedlings were analysed.

In order to quantitatively compare lignin depositions on *kor2* mutants with the Col-0 wild type, lignin depositions were quantified from captured photos using ImageJ as described (2.2.5.4). Figure 41 shows the calculation of staining area measurement as percentage value staining of the total area of roots. Since no lignification was observed in samples grown upon 0 nM isoxaben treatment, the values for lignification on those are omitted from the figure.

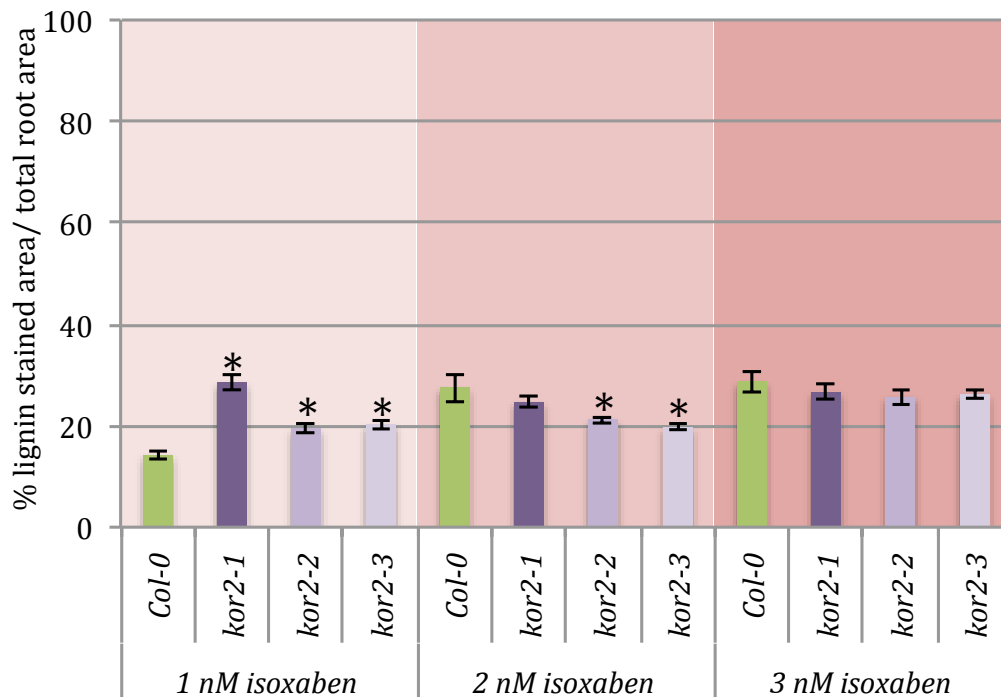


Figure 41 Percentage of phloroglucinol staining for lignin deposition in Col-0 and *kor2* seedling roots treated with 1, 2 and 3 nM isoxaben for three days. The values represent lignification area of total root area in percentage where the data was analysed using at least twenty individual seedlings $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

From the results shown in Figure 41, although the variations between different alleles for *kor2* mutants were observed, it was shown that for plants treated with 1-2 nM isoxaben, percentage values of lignin staining in this area for all three *kor2* mutants were different from Col-0 wild type. The result for 1 nM isoxaben treated seedlings exhibited more significant lignification was observed from all three mutants than that from Col-0. Variations were observed among three alleles where *kor2-1* showed more pronounced lignification than other two alleles. For 2 nM isoxaben treated seedlings, it was shown that the calculated percentage values for detected lignifications for *kor2-2* and *kor2-3* were nearly constant for both 1 nM and 2 nM isoxaben treated conditions. While it was increased in Col-0 but was decreased in *kor2-1*. However, no significant differences were observed on the staining values for the seedlings treated with 3 nM isoxaben. Percentage values of lignification for the mutants appeared to be more than Col-0 at 1nM isoxaben treatment then became lower at 2 nM isoxaben. That can be explained by variations of the observed root morphologies in Figure 40 where swollen and damaged cells were showed more in mutants than Col-0.

4.6.4 Cell wall in *kor2* and wild type plants

In order to determine the impact of loss of *KOR2* gene activity on cell wall composition, cell wall analysis of stems from mature plants was performed. The analysis consisted of quantifications of cellulose, uronic acids (indicative of pectic polysaccharides) and neutral cell wall sugars. Figure 42 summarizes the results from the quantification of crystalline cellulose. No significant differences were observed in any of the *kor2* alleles compared to the Col-0 wild type.

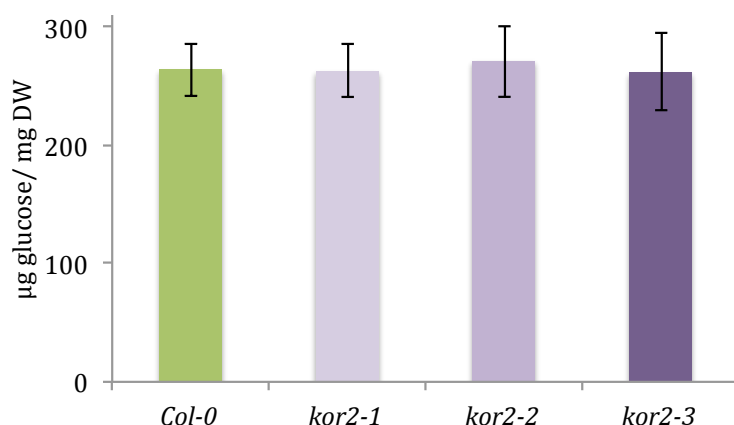


Figure 42 Cellulose measurement from stems of mature plants of Col-0 wild type, *kor2-1*, *kor2-2* and *kor2-3* mutants. The values represent average cellulose content in $\mu\text{g glucose per mg dry weight}$ of plant materials which were calculated from three biological replicates where the data was analysed using at least four individual plants. * indicates significant difference of mean of the data set compared with Col-0 (student t-test, $P < 0.05$). Error bars are based on standard deviation.

Figure 43 shows the results from the uronic acids measurements. All *kor2* alleles exhibit significantly lower amounts of uronic acids compared to Col-0 control plants.

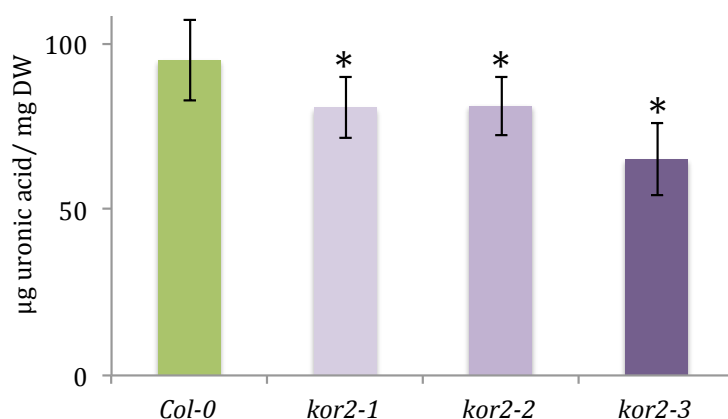


Figure 43 Uronic acids measurement results in mature stems of Col-0, *kor2-1*, *kor2-2* and *kor2-3* plants. The values represent average uronic acid content in $\mu\text{g uronic acid per mg dry weight}$ of plant materials which were calculated from three biological replicates where the data was analysed using at least four individual plants. * indicates significant difference of mean of the data set compared with Col-0 (student t-test, $P < 0.05$). Error bars are based on standard deviation.

Figure 44 shows the levels of neutral cell wall sugars in stems of mature plants. The results show that in all three *kor2* alleles levels of neutral cell wall sugars are elevated.

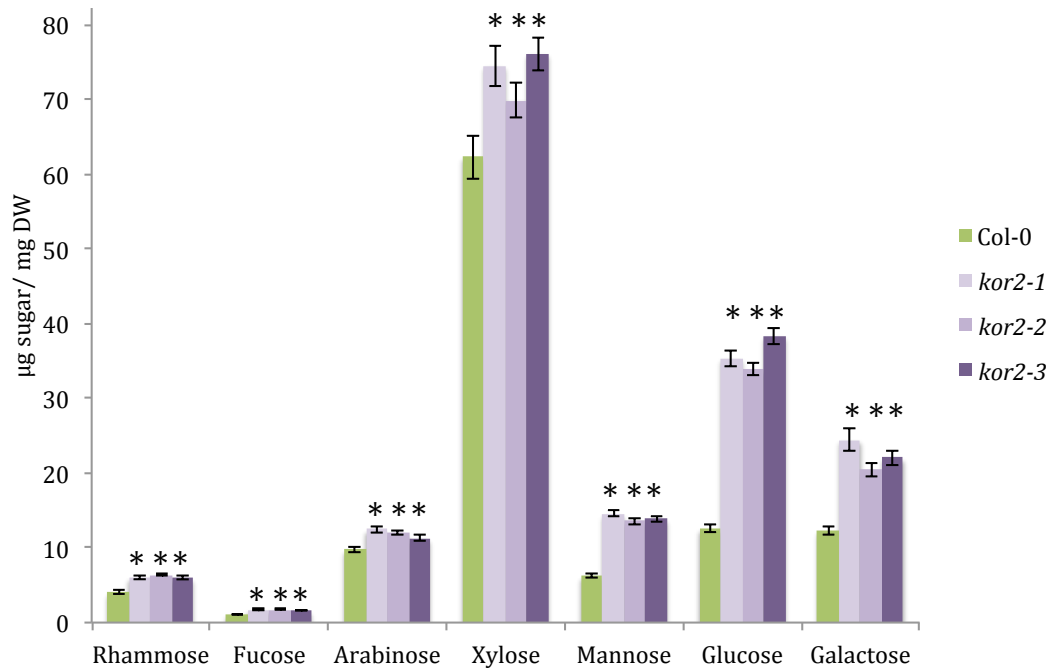


Figure 44 Neutral call wall sugar measurements in stems from Col-0, *kor2-1*, *kor2-2* and *kor2-3* plants. The values represent average values of seven different neutral cell wall sugar contents in µg sugar per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using at least four individual plants. * indicates significant different of mean of the data set compared with Col-0 (student t-test, $P < 0.05$). Error bars are based on standard deviation.

4.7 Discussion and summary

4.7.1 Heterologous expression of KORRIGAN2

The initial attempts to heterologously express KOR2 were performed in *Pichia* since previous attempts to express GH9 family members from *Aspergillus* sp. in *Pichia* had been successful (Bauer et al., 2006). In this study, integrity of the sequence of the expression construct and successful transcription in *Pichia* has been detected. However, this approach has not resulted in protein expression and there are different possible reasons why this could be the case. A particular type of posttranslational modification of the expression system that could cause unsuccessful secretion of the target protein or the native proteolytic activity of the host could prevent successful expression of KOR2 (Huang et al., 2011; Sinha et al., 2005). Because of these difficulties the expression system was changed and constructs for protein production in *E.coli* were generated and transformed into different *E.coli* strains.

Based on the available information regarding signal sequences and trans-membrane regions, a cloning strategy for expression of KOR2 in *E.coli* was developed. It had been decided to generate two constructs, one containing the predicted extracellular domain/ soluble part of the target protein and the other one containing the complete protein. Three different expression strains available in the lab were employed. BL21 strain provides the tightest control of uninduced expression; however, the induction is not as efficient as other strains used (Novagen competent cells user manual). For BL21 (DE3), the strain has highest level of expression; however, leaky expression of T7 polymerase could happen and that would cause uninduced expression of the target protein (Novagen competent cells user manual). Recently, SoluBL21™ has been launched. The strain provides increased expression of a soluble form of the target protein (Amsbio user manual). The pET28 vector was used for expression since a 6xHis tag forms part of the vector providing the means for easy purification of the expressed protein. In parallel the IPTG-inducible T7 promoter is a very high expression promoter allowing production of large amounts of target protein upon induction. However, no purified protein could be obtained. Different

possible causes could be responsible for this situation. One possibility is that the gene is not correctly transcribed, which could be checked by RT-PCR or northern blot with a probe for *KOR2*. For heterologous protein expression in *E. coli* expression system, one limitation for successful expression of proper folded target protein is insufficient chaperone available in the host cells (Martínez-Alonso et al., 2010). Alternative strains may also be used since the different strains have different predicted codon usage bias (<http://www.kazusa.or.jp/>) and different genetics, which could affect the expression, modification or stability of the target protein. An alternative would be the BL21 Codonplus-RIL strain, which enhances expression of eukaryotic proteins containing codons rarely used in *E. coli*. Other alternative eukaryote expression systems, which have been used successfully in the past, could be used. For instance, Vanzin et al. (2002) successfully expressed full-length MUR2 protein both wild type and mutant versions in African Green Monkey Kidney Cells (COS cells) and obtained active form of the target protein from wild type version. Employing the same expression system, COS cells, a group of CsIs from *Arabidopsis* and rice were also successfully expressed (Liepman et al., 2005). In addition, a homogalacturonan α -1,4-galacturonosyltransferase from *Arabidopsis* (AtGAUT1) was successfully heterologous expressed in the human embryonic kidney cell line HEK293 (Sterling et al., 2006). Another success in heterologous expression of two golgi localised (1,3)- α -D-xylosyltransferases, RGXT1 and RGXT2 was in baculovirus transfected insect cells, Sf9 and Hive five cells from where active forms of target proteins were obtained (Egelund et al., 2006). For plant expression system, transgenic *Nicotiana benthamiana* were used to express a xylogalacturonan xylosyltransferase from *Arabidopsis* (Jensen et al., 2008).

4.7.2 Isolation of homozygous T-DNA insertion lines for

KORRIGAN2

The criterion for selection of the T-DNA insertion lines to generate loss of function alleles was the predicted position in the gene in which the T-DNA fragments are inserted. Positions in exons or close to the beginning of the coding

region were selected as the most favourable sites. The three insertion lines selected for *KOR2* had putative insertion sites as shown in Table 5 and the sites were confirmed in this project by sequencing of the amplified fragments. For SK_006954 (*kor2-1*), the insertion site is in the 2nd exon, which may result in a 230 amino acids long truncated protein. This would encompass the cytoplasmic domain, trans-membrane region and about 130 amino acids of the extracellular domain, including two N-glycosylation sites (one in the trans-membrane region and another in the extracellular domain). For SK_051123 (*kor2-2*), the insertion position is in the first intron region before a predicted ACE-binding factor site (softberry.com), which may be a BZIP recognition site for transcriptional regulation of the gene (Singh, 2002; Wilkie et al., 2003). This could result in a 130 amino acids long fragment to be translated, which contains a single N-glycosylation site, the cytoplasmic domain, the trans-membrane region and about thirty amino acids of the extracellular domain. However, for both *kor2-1* and *kor2-2* mutants, if any truncated proteins would be expressed they would not have any predicted active sites essential for the enzyme activity (active sites; 517, 565 and 574 amino acid positions based on UniProt database). For SK_104013 (*kor2-3*), the insertion position is in the first exon, possibly generating a 79 amino acids long truncated protein. This fragment contains only the intracellular domain. To summarise, the different alleles with their different insertion positions may generate the tools to understand how different domains of the protein contribute to the activity of the complete protein.

PCR-based genotyping used to isolate homozygous plants for the insertion lines is a simple and efficient method, which allows a large number of samples to be analysed in a short period of time. As shown in the results section 4.4, homozygous plants for all selected insertion lines have been successfully isolated. Three independent lines were used for this gene since no backcrossing with Col-0 wild type could be done in this study due to time limitations. One drawback of using T-DNA insertion lines as a genetic tool is that there are many T-DNA insertion fragments, which could be inserted somewhere else in the genome and may cause unspecific mutant phenotypes. If the same mutant phenotypes are observed in all three lines the probability that the effect observed is unspecific is negligible. An alternative tool to overcome such a

problem would be RNAi knockout line where knocked out of the gene activity is more specific. In parallel, a construct to complement *KOR2* expression could be introduced into the mutant background to conclusively proof that the phenotypes observed are caused by loss of *KOR2* activity.

4.7.3 Analysis of *KOR2* gene expression in the insertion lines

After isolation of plants homozygous for the insertions it is crucial to determine whether expression of the gene is affected by insertion of the T-DNA fragment. In this study, semi-quantitative RT-PCR was used to investigate expression of *KOR2*. Similar amounts of total RNA were used as templates the gene specific primers were designed to amplify the whole coding region, allowing detection of whole gene expression and also distinguishing any products from contaminating gDNA templates. Based on the results in Figure 32, there was no detectable gDNA contamination. The RT-PCR results from the reactions using target gene specific primers suggested that the sample for 24 hours isoxaben treatment has higher expression of *KOR2* than the mock treatment one. However the results with the *UBQ10* primers, which served as loading control suggested that more RNA template might have been used for the isoxaben-treated sample. Therefore, this is not an ideal technique to distinguish differences in expression of target genes between samples. To be able to compare expression levels in a quantitative way, quantitative RT-PCR (qPCR) should be used in combination with suitable reference gene(s), whose expression level(s) are constant (Udvardi, Czechowski, & Scheible, 2008). This approach would provide quantitative data, allowing to detect differences in expression level between different treatments, genotypes and to confirm microarray derived expression data. Even though the results from semi-quantitative RT-PCR should not be used to compare the expression level of the target gene, the data showed clearly that *KOR2* expression is prevented in all three-insertion lines analysed. Therefore, these lines are suitable tools to determine the function of *KOR2* in cell wall metabolism.

4.7.4 Phenotypic characterisation of the mutant plants

In the past, functional studies of members of large gene families have frequently encountered the problem of functional redundancy. This may not be the case for *KOR2* in the response to isoxaben treatment, since other family members that show similar expression profiles on Genevestigator database do not appear to have same response to isoxaben treatment based on the available microarray data (Figure 45). Although it seemed like functional redundancy would not be an issue for seedlings grown upon isoxaben treatment, still that could be the case for plants grown under standard growth conditions since no obvious phenotypes were observed for *kor2* plants. Another explanation could be because the gene may only be essentially required under certain stress conditions, which would be in accordance with microarray derived expression data. Therefore, seedlings exposed to different treatments were investigated.

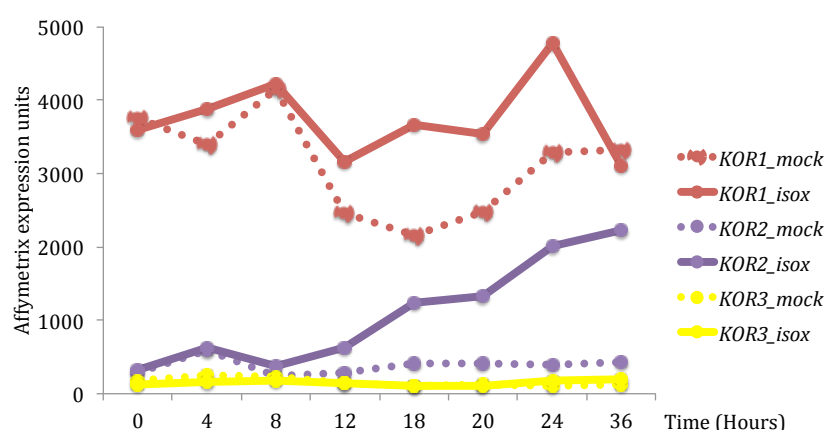


Figure 45 Microarray expression data for *At5g49720* (*KOR1*) (red), *At1g65610* (*KOR2*) (purple) and *At4g24260* (*KOR3*) (yellow) under mock treatment (dash line) and isoxaben treatment (bold line). Data derived from (Hamann et al., 2009).

Phenotypic analysis results showed no obvious phenotypes observed from growth stage-based analysis of *kor2* mutants compared with Col-0 wild type. In accordance with the growth stage-based analysis results, the mutants that their rosettes completely grew also showed no significant difference sizes compare with Col-0 wild type. However, biomass measurement in term of fresh weight and dry weight measurement from the plants at the same developmental stage showed lower biomass in the mutants than in wild type. No significant difference for biomass from senescence plants was observed. These results

suggested that disruption of *KOR2* involved in plant growth at early growth stage. Although the changes were not obvious by measuring plant growth (rosette diameter) but the changes are significant enough to be detected by measuring biomass. More in depth study to examine whether, in which tissues or cell types the phenotypes shown will provide clearer insight in the function of *KOR2* in plants at rosette growth complete stage and how or why the effects could not be detected for senescence plants. Possible hypothesis could be that *KOR2* plays quite an important role in early stage development in the way that disruption of the gene may disrupt development or differentiation of some cell types or cause thinner walls that resulting in less biomass observed. The experiments to investigate cell types and anatomical analysis need to be performed to prove the hypothesis.

Publically available expression data suggested that *KOR2* is expressed in radicles and roots of seedlings; therefore, germination assays and root growth assays were performed. The data presented in section 1.1.1 suggests that *kor2-1* and *kor2-2* mutants exhibit a late germination phenotype at day 1 and 2 after vernalisation but not on day 3. It is possible that the phenotype observed for *kor2-1* and *kor2-2* are caused by unknown insertions in the background. Alternatively, and more likely to be the case, is that the *kor2-3* allele did not exhibit the phenotype because of other genetic events occurring in the background of this line. To test whether the late germination phenotype is the result of loss of *KOR2* activity, complementation of *KOR2* gene activity in the *kor2-1* and *kor2-2* alleles with a wild type copy is the appropriate test. Also RNAi constructs, which could be cleaner (less genetic background effects) than T-DNA insertion, could be used. Using different methods to disrupt gene expression together with complementation would provide more evidences whether *KOR2* is required to regulate germination. However, more experiments could be carried out to investigate whether knock out and/or over expression of *KOR2* affects the response to phytohormones known to be involved in germination regulators, such as ABA and GA. Germination assays under supplementation of those hormones could be performed. In addition, crossing *kor2* mutants with mutants that are insensitive to the hormones could be used to identify at which stages of the transduction the target gene is involved.

For the root length phenotype, seedlings from all three insertion-lines show significantly longer roots than the wild type. However, the phenotype observed in *kor2-3* is again less pronounced than in the other two lines. These results suggest that *KOR2* limits root growth. In contrast, it has been reported that loss of *KOR1* causes shorter roots and shorter etiolated hypocotyls (Nicol et al., 1998). However, to determine the specific cause of the phenotype observed additional experiments are required. These would involve quantifying cell division/ elongation rate, cell size as well as cell counting. It is also possible that the target gene may involve in the control mechanism of radicle emergence during the germination and also could be a negative control for root growth; however, more evidence needed to be investigated. Interestingly, similar phenotypes (late/ lower germination rate with slightly longer primary roots) were observed in ethylene insensitive mutants, *etr1-1* and *ein2-5* (Wang, et al., 2007; Wang, 2007). It was found that *etr1-1*, which is a mutant for ethylene receptor, showed 10 % lower germination than wild type (Bleecker & Kende, 2000). The mutant also showed longer roots with hypersensitive response to salt and osmotic stresses (50-200 mM NaCl and 200-500 mM Mannitol) (Wang, et al., 2007; Wang, 2007). For *ein2-5* mutant, which is a mutant for a central component in ethylene transduction pathway, similar phenotypes to *etr1-1* were observed; additionally, it was found that the mutant is hypersensitive to ABA (Wang et al., 2007). Therefore, *KOR2* might be involved in ethylene/ ABA signalling pathway or knockout of the gene might affect these pathways. However, the experimental evidences provided in this study are not enough to draw any conclusion. Experiments to quantify ethylene and/or ABA, or to investigate hypersensitive phenotypes of the knockout and/or overexpression lines could be performed. In addition, phenotypes of double or triple mutants of *kor2* and the mutants for receptor or other components of the signalling pathways could provide more evidences regarding whether *KOR2* is involved in ethylene or ABA signalling pathway.

Short and long term experiment with Col-0 and *kor2* seedlings showed that the mutants exhibit hypersensitive responses to isoxaben both regarding root growth and lignin production. The root measurement (Figure 38) showed that all the *kor2* seedlings exhibit a dose dependent decrease in root length that

is more pronounced than wild type. Phloroglucinol staining results suggest that in mutant seedling roots more lignin is deposited upon isoxaben treatment than in wild type. In accordance to the observed staining results, quantitative analysis for the staining also suggest significant difference observed. These lignin-staining results suggested that disruption of *KOR2* lead to isoxaben-hypersensitive phenotypes. Mutant phenotypes in responses to cell wall impairment using isoxaben suggested the role of *KOR2* in cell wall integrity maintenance as if disruption of *KOR2* results in more damage of cell walls, thus, more lignin deposition to strengthen the walls. However, the quantification was performed by measuring the stained area from 2D photos of the roots, which could miss some part in other planes or other levels. Thus better imaging techniques or 3D scanning could provide more accurate and better results. Or other possible methods to quantify the lignin phenotype properly would be immune labelling, chemical assays (Klason lignin) or NMR-based analysis (Chabannes et al., 2001; Marita et al., 1999).

Although no striking phenotypes were observed in plants grown under standard growth conditions, it was possible that cell wall phenotypes exist that can only be observed through detailed investigation. For instance, *xxt1 xxt2* plants, which are totally deficient in xyloglucan have only subtle growth phenotypes (Park & Cosgrove, 2012b). Consequently cell wall analysis of stems from mature plants grown under standard conditions has been performed. Crystalline cellulose, uronic acids and neutral cell wall sugar levels were determined. Regarding cellulose, no significant differences were observed between *kor2* and Col-0 plants. Characterisation of the closely related *KOR1* mutant showed that disruption of that gene caused only small changes of cellulose content and did not change microfibril structure or crystallinity of cellulose but did affect other matrix structure (His et al., 2001; Lane et al., 2001; Sato et al., 2001; Szyjanowicz et al., 2004; Takahashi et al., 2009). The results presented here suggest that *KOR2* seems not to be required for cellulose biosynthesis or crystallisation of the cellulose microfibrils. However, other methods should be employed to confirm these results and measure the amounts of both crystalline and amorphous cellulose. Suitable methods would be for

example x-ray diffraction, AFM, FTIR and NMR-based analysis (Dick-Pérez et al., 2011; Notley et al., 2004; Park et al., 2010; Stephens et al., 2008).

The initial cellulose measurement results were complemented by uronic acids measurement results (representing pectic polysaccharides) and neutral cell wall sugars measurement results. The results showed that all mutant alleles have lower amounts of uronic acids (Figure 43) whereas the levels of all neutral cell wall sugars are higher than in wild type (Figure 44). However, it had been previously reported for *kor1* that pectin levels are increased to 162 % in the mutant background (Sato et al., 2001). Antibody based analysis of *kor1* cell walls suggests the walls contain more HG and RG-I molecules (His et al., 2001). Similarly enhanced amounts of uronic acids have also been reported for *Arabidopsis* seedlings treated with the cellulose biosynthesis inhibitors DCB and isoxaben (Hamann et al., 2009; Sato et al., 2001). The differences in the results between *kor1* and *kor2* could be due to the plant materials used for the measurements were derived from the different developmental stages (*kor1* seedlings versus *kor2* mature stems). In order to compare the phenotypes of the two genes, cell wall analysis of cell wall compositions could be done from plant materials at the same age or same developmental stage. To summarise, these results suggest that *KOR2* is probably not essentially required for cellulose biosynthesis but more likely to be involved in pectin metabolism since all the knockout mutants showed uronic acids phenotypes.

Chapter 5 Functional characterization of *XTH31*

5.1 Plant xyloglucan endotransglycosylase/ hydrolases (XTHs)

Xyloglucan endotransglucosylase/hydrolase (XTH) enzymes can mediate structural modifications of cell walls during cell expansion by splitting and reconnecting xyloglucan (XG) chains (Baumann et al., 2007; Eklöf & Brumer, 2010). Like all the matrix polysaccharides, XGs are synthesised in the Golgi and delivered via vesicles transported through cell membrane (Cosgrove, 2000; Cosgrove, 2005; Scheller & Ulvskov, 2010). For XGs, the precursor molecules are transported to the cell surface before the molecules are modified by glycosyl hydrolases and incorporated by endo-tranglycosylases (Scheller & Ulvskov, 2010).

The evolution of XTH has been studied in species from *Selaginella* to maize, and it was found that the enzyme activity is high in the root elongation zone (Vissenberg et al., 2003). In addition it was shown that the activities of enzymes are conserved from *Selaginella* and among vascular plants (Van Sandt et al., 2006). It has also been suggested that XEH activity of the enzymes had evolved from XET activity (Baumann et al., 2007). The functions of the enzymes have been suggested to be involved in germination, pollination, fruit growth and ripening (Eklöf & Brumer, 2010; Miedes & Lorences, 2009; Vissenberg et al., 2003).

The XTH protein family is a subgroup of the glycoside hydrolase family 16 (GH16; Cantarel *et al.*, 2008). Computer based analysis identified an eighteen amino acid long hydrophobic core at the N-terminus. This motif could direct the protein through the plasma membrane to the cell wall (Aubert & Herzog, 1996).

5.2 *Arabidopsis* XTHs

Based on CAZy database, there are thirty-three genes encoding GH16 family members predicted in *Arabidopsis* genome. The candidate gene *At3G44990* encodes AtXTH31 (AtXTR8) (Rose *et al.*, 2002). The cDNA of *AtXTH31* was isolated from an *Arabidopsis* germinating seed cDNA library and it was shown that the transcript was expressed twenty-four hours after imbibition and reached the highest level after seventy-two hours (Aubert & Herzog, 1996). Spatial expression of *AtXTH31*, observed in the experiment performed by Becnel *et al.* (2006), using a promoter reporter construct, *XTH31::GUS*, showed strong expression of the reporter in roots. Similarly, northern blot analysis of *XTH31* (*AtXTR8*) expression showed that the transcripts predominantly accumulate in roots (Aubert & Herzog, 1996). The spatial expression pattern of *AtXTH31*, which was observed in both germinating seeds and vegetative organs of three to four week old plants, suggested that the product might be involved in cell expansion rather than seed-reserve mobilisation (Aubert & Herzog, 1996). In the same study, the effect of gibberellic acid (GA) on *XTH31* expression was determined using GA-deficient mutants. Results from northern blot analysis showed no transcript of *XTH31* in severely GA-deficient *ga1-3* mutant plants grown without exogenous GA supplement, indicating that GA is required to turn on *AtXTH31* expression during germination (Aubert & Herzog, 1996).

5.3 Heterologous expression of XTH31

In order to characterize the biochemical properties and functions of the candidate protein heterologous expression of XTH31 using other organisms was initiated. Bearing in mind the predicted post-translational modification sites, it was decided to use *P. pastoris* strain GS115 as a host with the expression constructs based on pPICZ for the non-secreted and pPICZ α for the secreted version. However, the attempt to express the constructs in *Pichia* was not successful therefore, it was decided to attempt to generate expression constructs for expression in *E. coli*. Since the cDNA for XTH31 had been cloned into the expression vector in *Pichia* and this construct had been shown to be error free they were used as the template to amplify insert fragment for expression of the

extracellular domain (soluble part) of *XTH31* in *E. coli* using the TOPO® TA Cloning® (Invitrogen) (Figure 46).

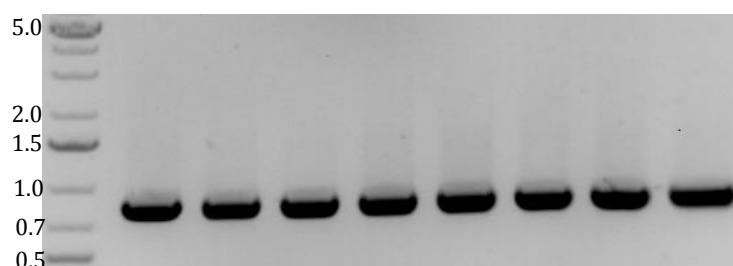


Figure 46 Results from PCR reactions amplifying fragments encoding the extracellular domain of *XTH31* flanked by *Bam*HI and *Not*I restriction sites using Phusion® DNA polymerase. Size of DNA ladder fragments shown in kb

The insert fragments were then extracted from agarose gel and cloned into pCR2.1 TOPO vectors as described before transformation and selection of the positive clones. Figure 47 shows colony PCR from ten pCR2.1-*XTH31* clones.

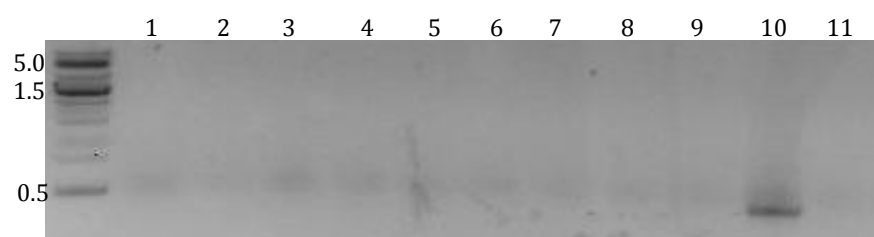


Figure 47 Colony PCR to isolate DH5α harbouring pCR2.1*XTH31* (1-10) and DH5α with empty plasmid (11). Size of DNA ladder fragments shown in kb

Figure 47 shows that out of ten clones tested by colony PCR only colony 10 seems to contain the insert. Therefore, an overnight culture was inoculated from this colony; plasmid DNA isolated and a restriction digest using *Bam*HI and *Not*I performed. The results for restriction digest are presented below in Figure 48. A fragment with the expected size of 0.85 kb was detected.

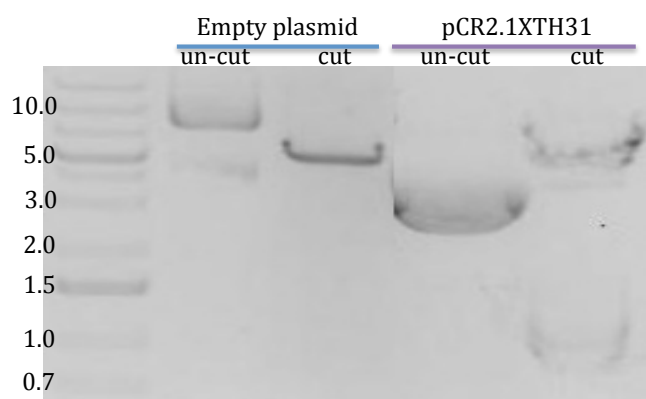


Figure 48 *Bam*HI-*Not*I double digest for pCR2.1XTH31 construct. Size of DNA ladder fragments shown in kb.

The insert fragment was then ligated into pET28b⁺ vector and the ligation mix transformed into DH5 α cells. After selection on kanamycin plates colony PCR was performed using T7 forward primer and reverse primer specific for the insert (primer sequences in Appendix1). This should amplify 1kb fragments if the construct is present. Results from colony PCR reactions (Figure 49) show that with the exception of colonies 10 and 13, all other colonies seem to contain the construct.

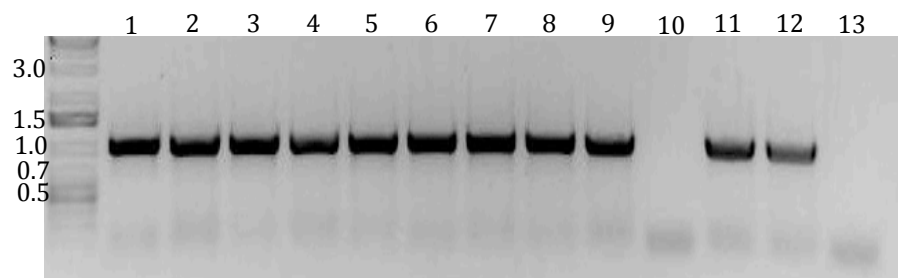


Figure 49 Colony PCR for DH5 α pET28XTH31 (1-12) and DH5 α harbouring pET28 empty plasmid (13) are shown. Size of DNA ladder fragments shown in kb

After performing colony PCR, three clones were selected and grown for plasmid DNA isolation. The DNA samples were digested with *Bam*HI and *Not*I to test whether the insert fragments were still presented. The results shown in Figure 50 suggested that all three clones tested are positive. Clone pET28XTH31#6 was selected for sequencing. This error-free pET28XTH31 construct was then used as the expression construct.

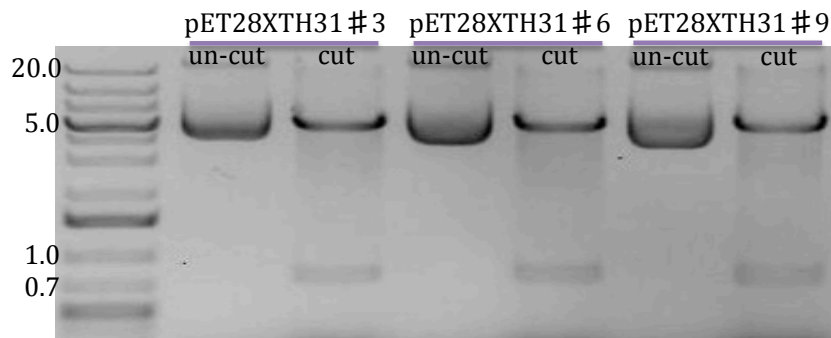


Figure 50 *Bam*HI-*Not*I double digest for pET28XTH31 construct. Size of DNA ladder fragments shown in kb.

The expression construct was transformed into three different *E. coli* strains: BL21, SoluBL21™ and BL21 (DE3). After antibiotic selection and colony PCR, positive strains were used for expression. Expression of *XTH31* from the selected clones was induced with 1 mM IPTG for six hours. Total proteins were isolated and separated via SDS-PAGE as shown in Figure 51.

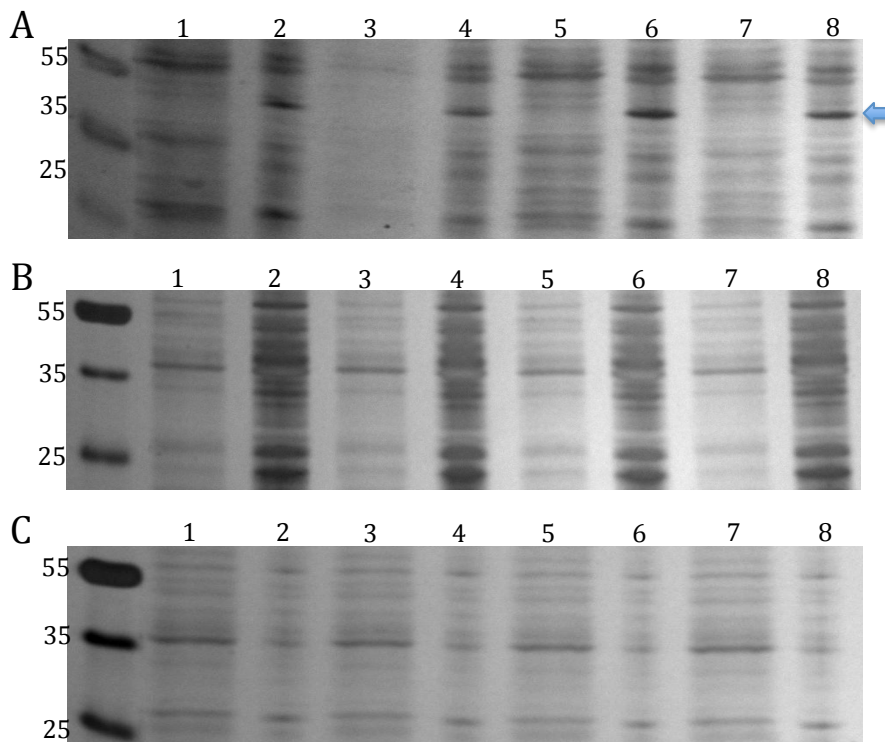


Figure 51 SDS-PAGE analysis for XTH31 expressed in *E.coli* strain BL21 (A), SoluBL21™ (B) and BL21 (DE3) (C). 4 different clones were tested; clone1 = lane 1 and 2, clone2 = lane 3 and 4, clone3 = lane 5 and 6 and clone4 = lane 7 and 8. Samples from 0 hour (lane 1, 3, 5 and 7) and 6 hours (2, 4, 6 and 8) of IPTG induction were shown. Putative XTH31 protein fragments are indicated with an arrow. Size of protein marker shown in kD

The results suggest that the BL21 strain is suitable for expression of XTH31 since treatment with IPTG resulted in appearance of a protein of the expected size. Therefore large-scale cultures were set up for purification and protein isolation. Samples from these cultures were harvested, purified via Nickel Affinity purification and the results of the purification analysed via SDS-PAGE (Figure 52).

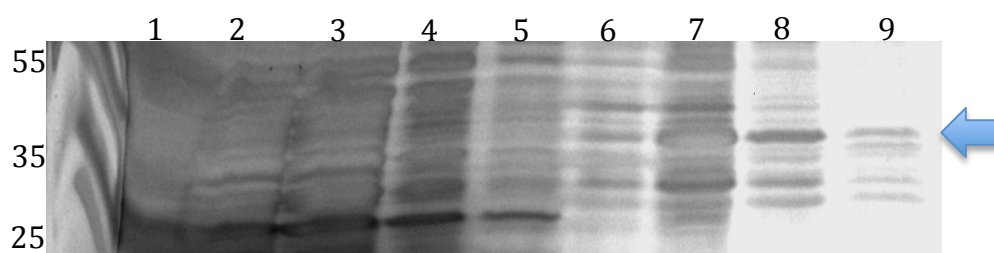


Figure 52 SDS-PAGE for XTH31 after purification for His-tagged proteins. Shown above are cell lysate (1), soluble proteins (2) Flow-through (3), 1th Wash (4), 4th Wash (5), 1st Elution (6), 2nd Elution (7), 3rd Elution (8) and 4th Elution (9). Target proteins were indicated with the arrow. Size of protein marker shown in kD.

The SDS PAGE gel showed that flow through and the initial elution fractions contain large amounts of other proteins. The latter elution fractions seem to be cleaner. Due to time limitations no activity assays using the crude lysate could be performed.

5.4 Isolation of homozygous T-DNA insertion plants

The biological functions of *XTH31* were investigated through a reverse genetics based approach. Instead of backcrossing the mutants with wild type plants to eliminate any background effects that could cause unspecific phenotypes, two independent insertion lines were isolated for *XTH31*. Figure 53 shows PCR-based genotyping to identify plants homozygous for the insertions. Results from PCR reactions for SK_046167 (upper panels) showed all ten plants analysed are homozygous for the insertions. Results for SK_129686 (lower panels) showed that four plants are homozygous (plant number 2, 3, 4 and 11 Figure 53).

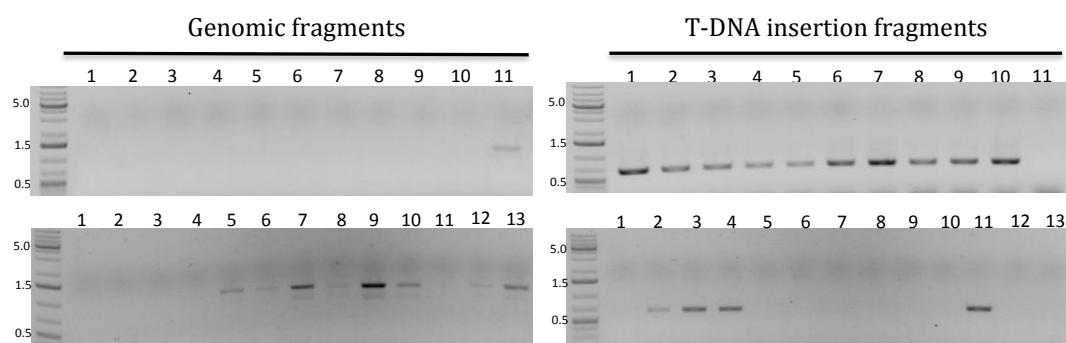


Figure 53 PCR-based genotyping to isolate plants homozygous for insertions into *XTH31*. Gene specific fragments and T-DNA fragments were amplified. For *xth31-1*, PCR reactions from ten independent mutant plants were shown on the upper panels (1-10) and Col-0 (11). For *xth31-2*, PCR reactions from twelve independent mutant plants were shown on the lower panels (1-12) and Col-0 (13). Size of DNA ladder fragments shown in kb.

SK_046167 (plant 7) was designated as *xth31-1* and SK_129686 (plant 11) was designated as *xth31-2*. Both selected insertion lines, SK_046167 (*xth31-1*) and SK_129686 (*xth31-2*), show putative insertion sites at the same predicted active site for the proton donor (Glu-118) within the GH16 domain based on data from SIGNAL and UniProt. Amplified T-DNA insertion fragments for the selected plants were confirmed by sequencing.

5.5 Analysis of *XTH31* expression in the T-DNA insertion lines

Once plants homozygous for the insertions had been isolated, seedlings from these plants and Col-0 wild type controls were mock and isoxaben-treated for 24 hours. Total RNA was isolated and semi-quantitative RT-PCR performed to determine whether the insertions lead to loss of gene expression. Primers specific for *XTH31* were used to amplify the whole coding region as shown in the left panel in Figure 54. Mock-treated Col-0 seedlings seem to exhibit higher expression levels than the isoxaben-treated ones (Figure 54, left panel, samples 1 and 2, respectively). While there is no expression observed in *xth31-1* seedlings (Figure 54, left panel, samples 3 to 4) some expression was detectable in both mock- and isoxaben-treated *xth31-2* seedlings (Figure 54, left panel, samples 5 to 6). In parallel RT-PCR reactions for *UBIQUITIN10* run in parallel, suggest that all samples contain comparable levels of RNA template (Figure 54, right panel). These results suggest that *xth31-1* represents a homozygous knockout allele while *xth31-2* appears to be a heterozygous.



Figure 54 Semi-quantitative RT-PCR for *XTH31* from Col-0, *xth31-1* and *xth31-2*. Gene specific reactions (right panel) and *UBIQUITIN 10* was used as a control (left panel). Col-0 mock and isoxaben-treated samples are shown in (1) and (2), *xth31-1* mock- and isoxaben-treated ones shown in (3) and (4) and *xth31-2* mock- and isoxaben-treated samples shown in (5) and (6). Size of DNA ladder fragments shown in bp.

5.6 Phenotypic characterization of the mutant plants

Seeds for *xth31* mutant alleles and Col-0 controls were germinated on soil to determine if loss/reduction of XTH31 transcription results in growth phenotypes under standard growth conditions. Phenotypes of the plants from various developmental stages throughout the life cycle of plants were determined as described in materials and methods. The results for the time required at each growth stage of Col-0 wild type and *kor2* mutants are shown in Table 10. However, no obvious phenotype was observed.

Table 10 Growth stage-based analysis results for the wild type and *kor2* mutants.

Developmental stages	Col-0	<i>xth31-1</i>	<i>xth31-2</i>
Leaf development	4-5 days	4-5 days	4-5 days
2 rosette leaves > 1 mm in length	6 days	6-8 days	6-8 days
3 rosette leaves > 1 mm in length	7-9 days	8 days	8 days
4 rosette leaves > 1 mm in length	8-10 days	9-10 days	11 days
5 rosette leaves > 1 mm in length	11-12 days	11-12 days	11 days
6 rosette leaves > 1 mm in length	11-13 days	13 days	11-12 days
7 rosette leaves > 1 mm in length	13-14 days	13-14 days	13 days
8 rosette leaves > 1 mm in length	13-15 days	15 days	13-15 days
9 rosette leaves > 1 mm in length	14-15 days	16 days	14-15 days
10 rosette leaves > 1 mm in length	14-16 days	17 days	15-16 days
11 rosette leaves > 1 mm in length	17-18 days	18 days	15-17 days
12 rosette leaves > 1 mm in length	17-19 days	19 days	18-19 days
13 rosette leaves > 1 mm in length	18-19 days	20 days	19-20 days
14 rosette leaves > 1 mm in length	20-21 days	20-21 days	20-21 days
Rosette growth complete	21-22 days	21 days	21-22 days
Flower production	19-20 days	19 days	19-22 days
Senescence complete	60-65 days	58-62 days	63-65 days

Once all the plants were at the growth stage where the rosette growth completed, rosette diameter was measured. Results showed no significantly differences among the wild type and two *xth31* mutants (Figure 55).

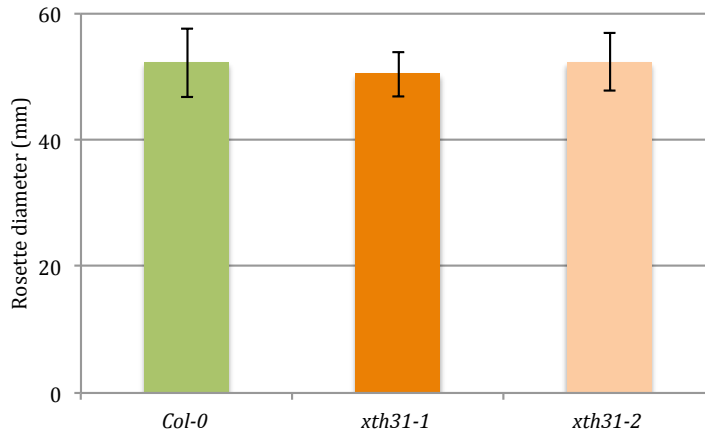


Figure 55 Rosette diameter measurement for *xth31* mutant plants that showed rosette growth complete stage. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average rosette diameter calculated from three biological replicates where the data was analysed using sixteen individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

About four weeks after sowing when the rosette growth is completed, biomass measurement was performed. The results showed that *xth31-1* mutant showed significantly less biomass in both fresh and dry weight than in Col-0 wild type and *xth31-2* (Figure 56).

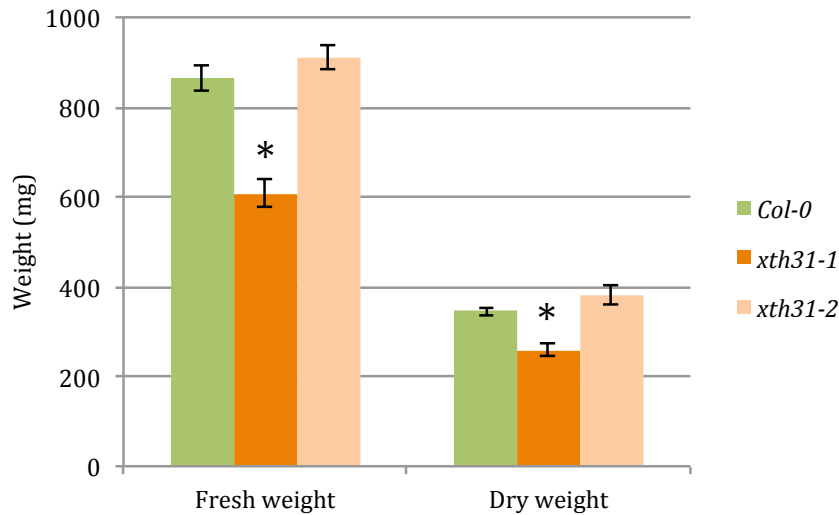


Figure 56 Biomass measurement for four week-old *xth31* mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight of either fresh or dried materials and calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

In addition to the measurement of biomass for plants at the growth stage where the rosette growth completed, the biomass for senescence plants were measure and the dry weights are shown in Figure 57.

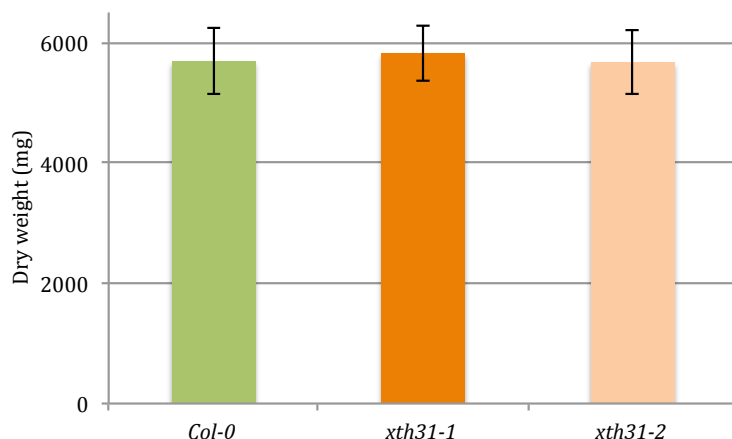


Figure 57 Biomass measurement for senescence *xth31* mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

Although significantly differences were observed for *xth31-1* mutant at the rosette-completed stage, no significantly difference was observed at the senescence stage (Figure 57).

5.6.1 Germination assay

The available microarray derived expression-profiling data deposited in the Genevestigator database suggested that *XTH31* is expressed in imbibed seeds and the radicle of seedlings with particularly high expression observed in the endosperm. Therefore, an experiment was performed to investigate whether *xth31* seeds exhibit germination phenotypes. Figure 58 shows the percentage of germinated seeds for *xth31* and Col-0 during the first three days after vernalisation. The results suggested a delayed germination for *xth31-1* since only approximately 30 % of the seeds germinated on the first day whereas for Col-0 60 % germinated. No significant differences to Col-0 were observed two and three days after germination.

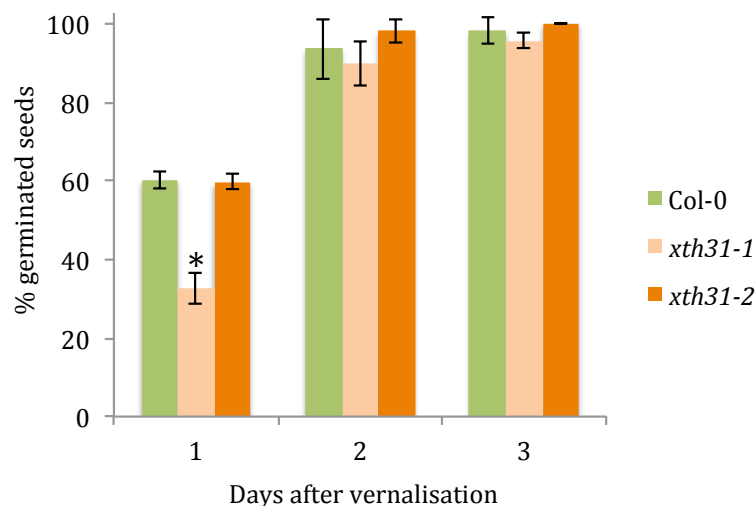


Figure 58 Results from germination assays for *xth31* mutants and Col-0 wild type. Sterile seeds were put on ½ MS plates and stratified at 4°C for forty-eight hours before grown in standard conditions as described. Seeds with radicals emerged were counted as germinated seeds after 1-3 days after plates were put under standard conditions. The values represent average percentage of germination calculated from three biological replicates where the data was analysed using fifty individual seeds ± SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

5.6.2 Root growth and isoxaben response phenotypes

Since it was shown on Genevestigator that is *XTH31* highly expressed in roots (Figure 12), root length of mutant seedlings was measured and compared with Col-0 controls. As shown in Figure 59, *xth31-1* seedling roots exhibit slightly longer root, while *xth31-2* seedling roots show no significant differences compared to wild type.

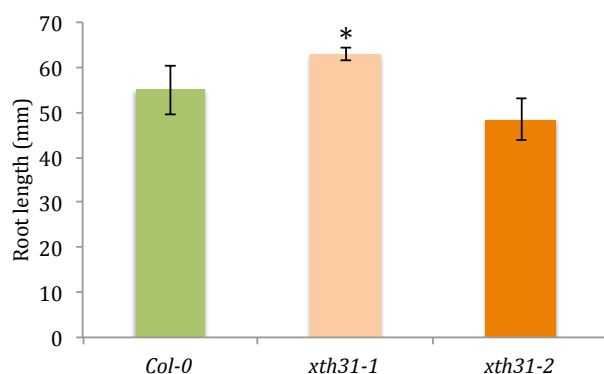


Figure 59 Root of seven day-old *xth31* mutants and Col-0 wild type seedlings grown vertically on ½ MS plates were captured and the photos were analysed using ImageJ software. The values represent the average of root length of seedlings calculated from three biological replicates where the analysis was performed using at least forty individual seedlings ± SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test, $P \leq 0.05$).

To determine if reduction of *XTH31* activity affects the response to isoxaben mutant and Col-0 seedlings were germinated and grown for four days on plates with regular growth media before shifting them onto ½ MS plates with different concentrations of isoxaben and scoring the impact of isoxaben on root growth (Figure 60). It was found that seedling root growth of both mutant alleles examined at 1 nM isoxaben was less inhibited than the Col-0 control. At 2 nM of isoxaben only *xth31-2* seedling root growth differs from Col-0, whereas at 3 nM *xth31-1* seem to be more sensitive while *xth31-2* roots are similar to Col-0. These results suggest that reductions in *XTH31* activity affect root responses to isoxaben treatment. However, they are not consistent between the alleles examined.

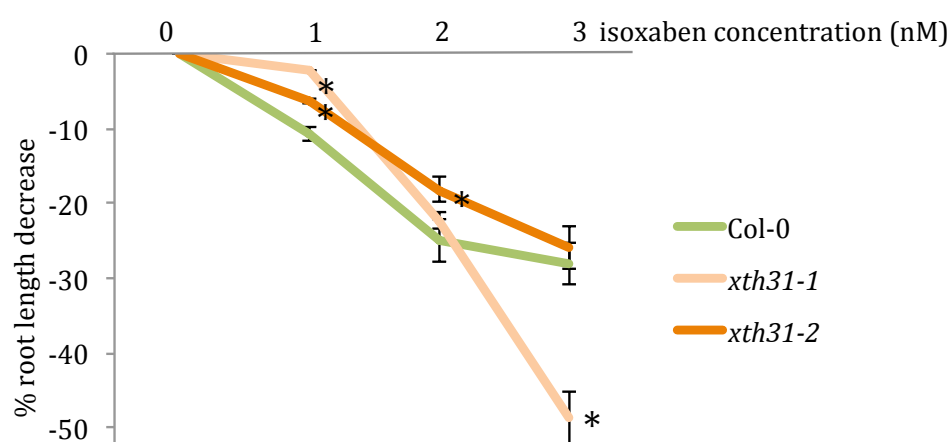


Figure 60 Impact of different isoxaben concentrations on *xth31* and Col-0 seedling root growth. Percentage of root length decreased up on isoxaben treatment. Roots of *xth31* mutants and Col-0 wild type seedlings grown vertically on ½ MS plates for four days before transferred to fresh ½ MS plates in presence or absence of 1, 2 or 3 nM isoxaben were analysed. Photos were captured as described and analysed using ImageJ software. Percentage of root length decreased were calculated from four biological replicates where the analysis was performed using at least ten individual seedlings $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type under the same treatment (student t-test, $n \geq 40$ per line and condition, $P \leq 0.05$).

5.6.3 Lignin deposition after isoxaben treatment

The available microarray derived gene expression for *XTH31* suggests transcript levels change upon isoxaben treatment. As stated earlier that isoxaben treatment has been shown to induce lignin deposition in elongation zone of seedling roots (Hamann et al., 2009). Therefore, *xth31* mutants and Col-0 wild type seedlings were grown in ½ MS liquid medium with 1% sucrose for six days

and then treated with either DMSO (for mock treatment) or 600 nM isoxaben for 12 and 24 hours. The seedlings were stained with phloroglucinol to detect lignin deposition (Figure 61). Figure 61 summarises the results. It seems that both after 12 and 24 hours of treatment the mutant seedling roots contain more ectopic lignin deposition than the Col-0 control as evidenced by the red staining (Figure 61 A). To compare the stained area of the roots quantitatively, the measurement was performed as described (2.2.5.4) and the results are shown in Figure 61, panel B.

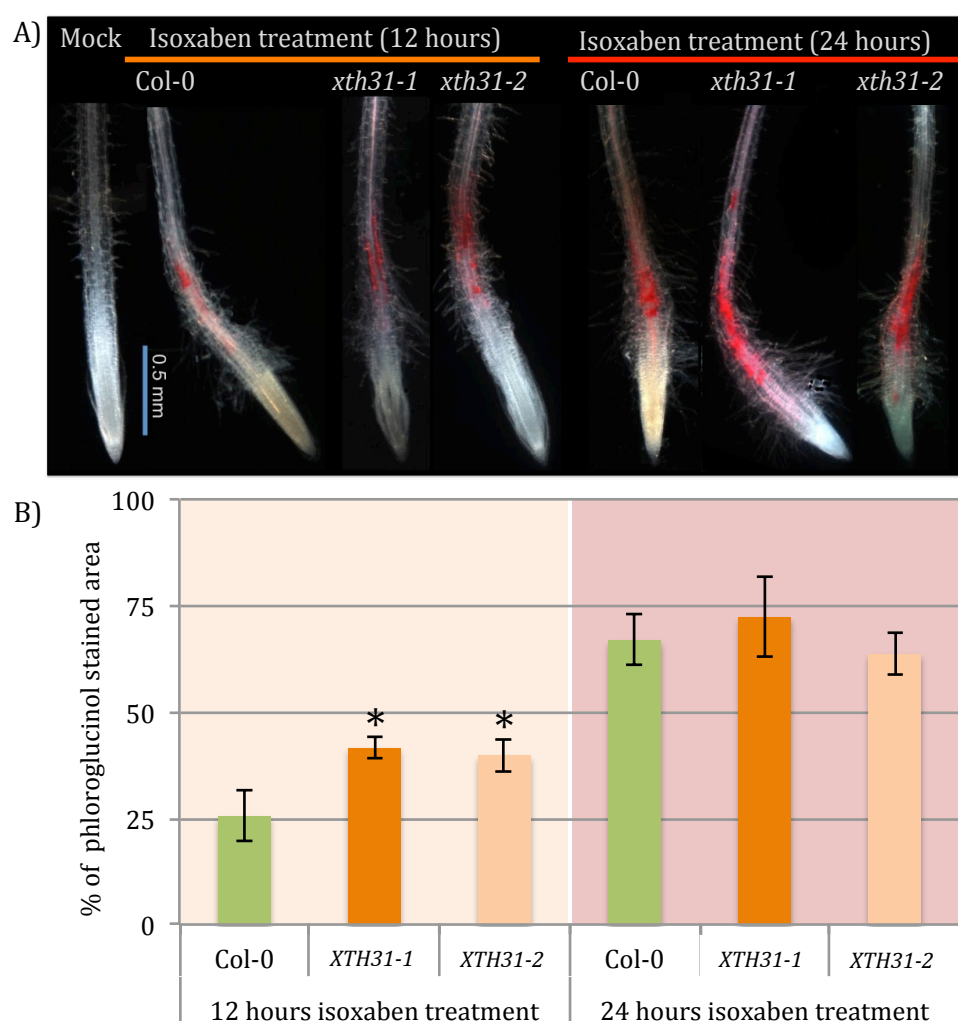


Figure 61 Phloroglucinol stained seven day-old *xth31* and Col-0 seedling root tips after isoxaben treatment for 12 or 24 hours. A) is the representative results of the phloroglucinol staining. Scale bar is 0.5 mm, for each genotype/treatment at least twenty seedling roots were analysed. B) Percentage of the phloroglucinol staining for seven day-old *xth31* and Col-0 seedling root tips after isoxaben treatment for 12 or 24 hours.

The results from 12 hour-treatment showed that both alleles for *xth31* mutants appeared to have significantly more lignin stained than in wild type.

While no significant difference was observed from 24 hours treated samples (Figure 61 B).

In addition to the liquid culture experiments, *xth31* and Col-0 seedlings were germinated on plates without isoxaben then shifted to plates containing different concentrations of isoxaben and lignin deposition and root morphology examined (Figure 62).

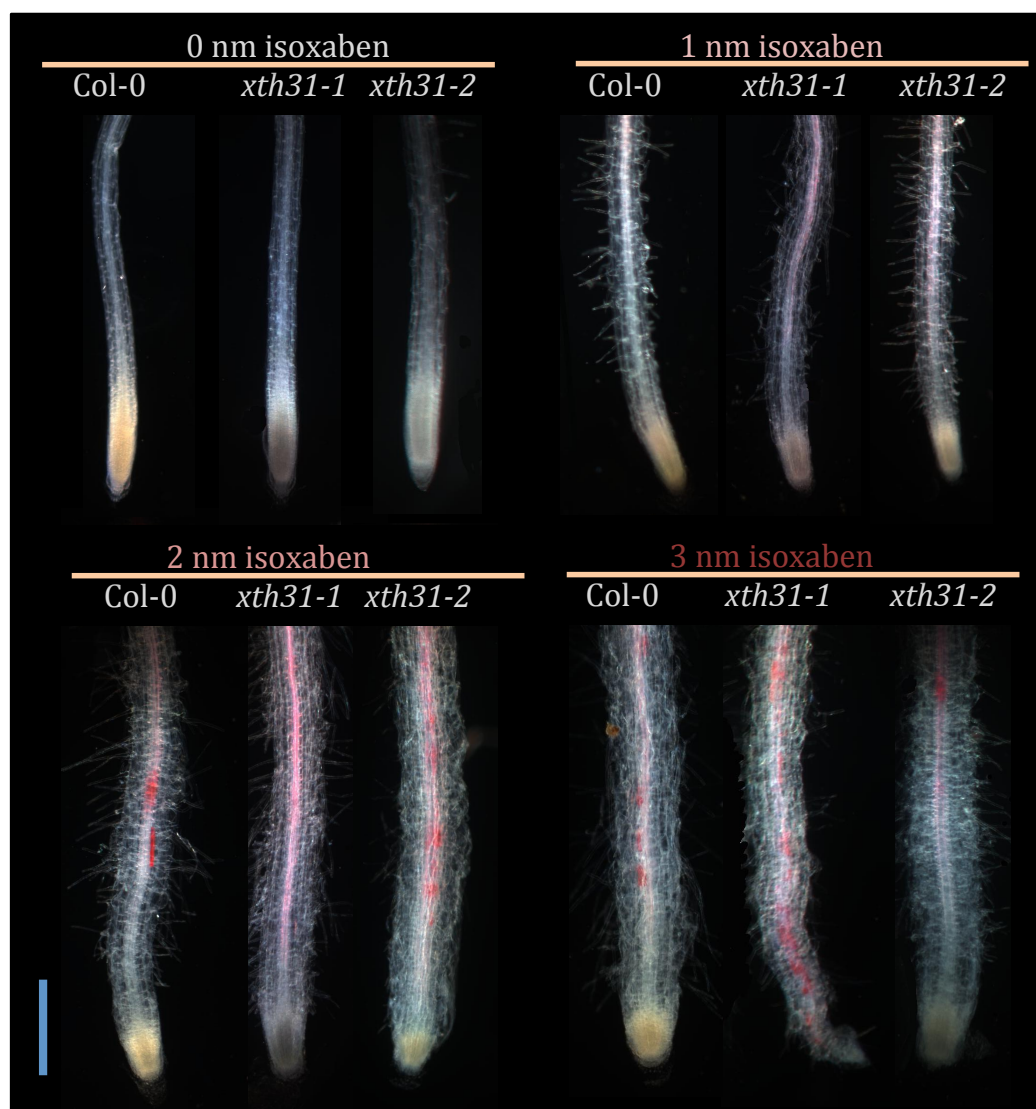


Figure 62 Lignin deposition in root tips of Col-0 and *xth31* seedlings after treatment with 0, 1, 2 and 3 nM isoxaben for four days. Scale bar is 0.5 mm.

Phenotypic analysis of the treated and stained roots suggests that at a concentration of 2 nM and 3 nM isoxaben, some ectopic lignin deposition occurs similarly in *xth31* and Col-0 roots. Cell pattern and root morphology of the *xth31* roots seem to be different compared with wild type after isoxaben treatment as

roots of the mutants appeared to show more swollen cells. The quiescent area at the root tip where cell division occurs seems to be reduced or to disappear upon isoxaben treatment, which is particularly pronounced in mutant root tips. Some degrees of phenotypic variations were observed but representative examples are shown here. In order to quantitatively compare lignin depositions on *xth31* mutants with the Col-0 wild type, lignin depositions were quantified as described (2.2.5.4). Results are shown in Figure 63. Panel A shows percentage values of lignin staining for stele and endodermis tissues, while panel B shows percentage values for lignin staining from cortex and epidermis.

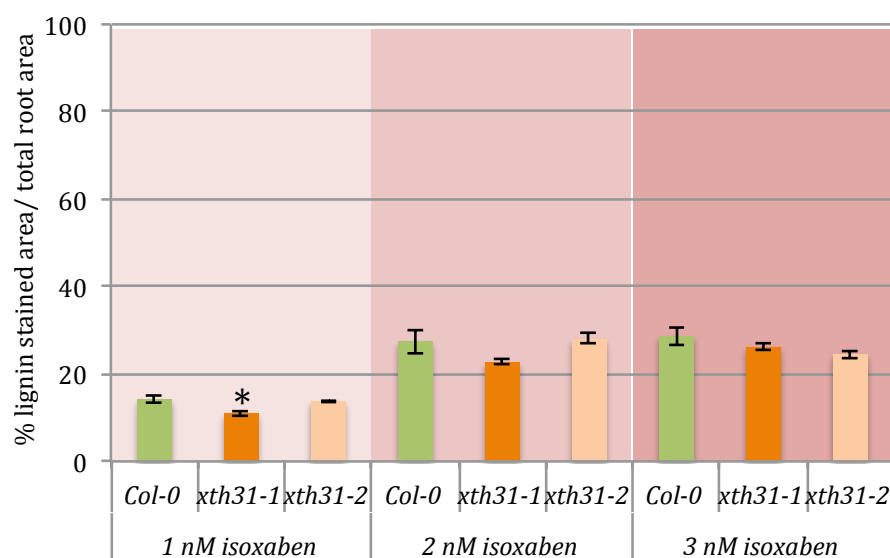


Figure 63 Percentage of phloroglucinol staining for lignin deposition in Col-0 and *xth31* seedling roots treated with 0, 1, 2 and 3 nM isoxaben for three days. A) percentage of phloroglucinol stained for vascular bundle tissue. B) percentage of phloroglucinol stained for other areas along the roots. The values represent average stained area measured via ImageJ and calculated from three biological replicates where the data was analysed using at least twenty individual plants $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

It was shown that the mutants from both alleles showed more lignin staining at 1 nM and 2 nM isoxaben, while at 3 nM no significant difference was observed (Figure 63 A). Panel B shows the percentage lignin staining values for cortex and epidermis of roots. The results for *xth31* mutant showed less staining in every concentration of isoxaben. However, the results for *xth31-2* showed no significant differences to wild type at the 2 nM isoxaben with very high variations among different seedlings for this same population (Figure 63 B).

5.6.4 Analysis of cell wall composition in mutant and wild type plants

Although no striking growth phenotypes were observed for the *xth31* mature plants grown under standard growth conditions, cell wall analysis was performed since changes in cell wall composition do not necessarily lead to obvious growth phenotypes. Cellulose levels from the mature plants were measured as described and the results for *xth31* and Col-0 plants are shown in Figure 64. Plants for both alleles show approximately 10 % higher cellulose levels compared to wild type.

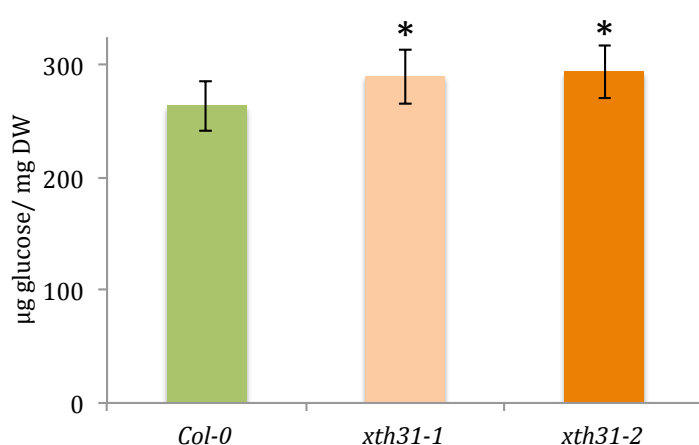


Figure 64 Cellulose measurements from stems of mature Col-0 and *xth31* plants. The values represent average cellulose content in µg glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants. * indicates significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are based on SD values ($n = 3$).

In addition to cellulose content, uronic acids levels were also determined and the results are shown in Figure 65. While *xth31-1* plants show a significant reduction in uronic acids content compared to Col-0, no significant differences were observed for *xth31-2*.

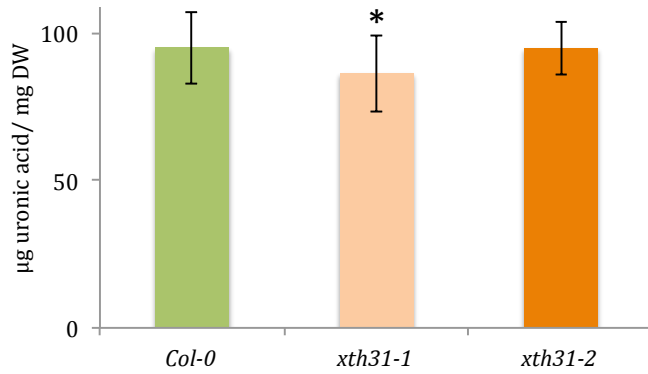


Figure 65 Uronic acids measurement of material from stems of mature Col-0 and *xth31* plants. The values represent average uronic acid content in µg uronic acid per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants. * indicates significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are based on SD values ($n = 3$).

Figure 66 summarizes the results of neutral cell wall sugars quantification in stems from Col-0 and *xth31* plants. The results show that the levels of nearly all neutral cell wall sugars (with exception for xylose) are elevated significantly in stems from mutant plants compared to Col-0.

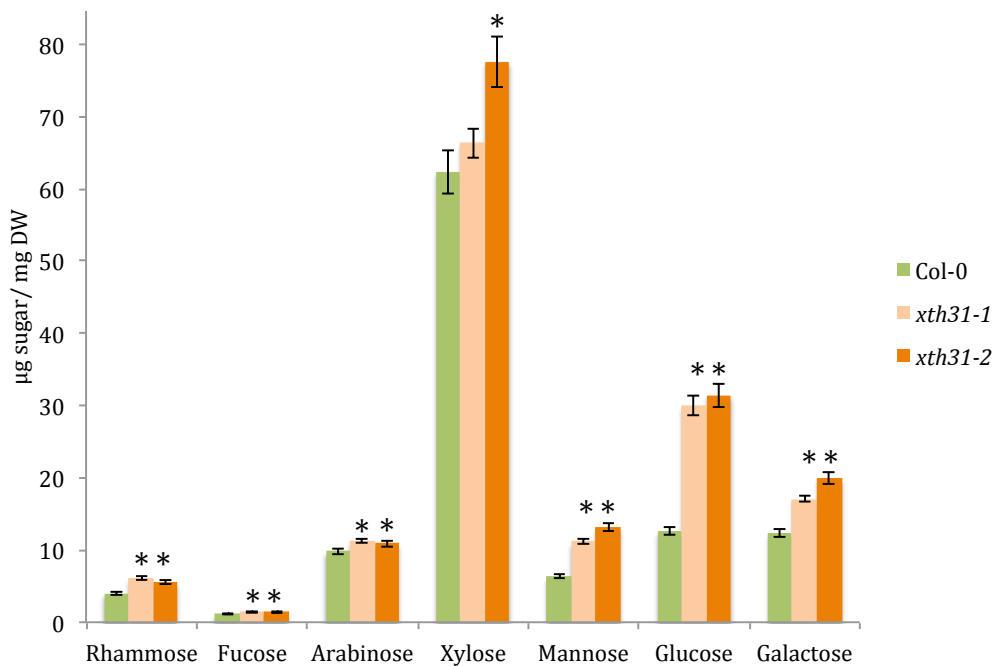


Figure 66 Neutral cell wall sugar measurements in stems from mature Col-0 and *xth31* plants. The values represent average values of different seven neutral cell wall sugar contents in µg sugar per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants. * indicates significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are based on SD values ($n = 3$).

5.7 Discussion and summary

5.7.1 Heterologous expression of XTH31

The analysis of the secondary structure of XTH31 (summarized in

Figure 19) showed that XTH31 could have two disulfide bonds (Cys₂₃₈ with Cys₂₄₆ bond and Cys₂₈₀ with Cys₂₉₃ bond) and eight other phosphorylation sites, suggesting a bacterial expression system may not be the optimal choice here. This is further supported by the success in expressing other XTH enzymes from *Aspergillus sp.* in *P. pastoris* (Bauer et al., 2006). Therefore, initially cloning and expression of *Arabidopsis* XTH31 was attempted in a *Pichia* based expression system. However, the target enzyme could not be expressed. That could be the result of incompatibility of posttranslational protein modifications of the expression system and the target protein and/or native proteolytic activity of the host. Similarly to the KOR2 expression project, correctness of the sequence and transcription of the construct had been confirmed through sequencing and RT-PCR (data not shown). Therefore as an alternative expression system, *E. coli* was chosen for the heterologous expression of the target gene. However, it was shown for XTH31 that the first twenty amino acids appear to represent a signal peptide based on data from the UniProt database (<http://www.uniprot.org/>). In parallel, results from the PredictProtein service suggested that the stretch between Phe₇-Ser₁₆ forms a trans-membrane region. Therefore, a cloning strategy was developed to express the predicted soluble part of the target protein without the signal sequence and the putative transmembrane region in *E. coli*. The pET28 vector from the pET expression system (Novagen) includes a hexahistidine tag, facilitating purification of the target protein using a cobalt or nickel affinity column. Also the vector contains a T7 promoter controlling the expression of the inserted gene (here: *XTH31*), allowing large amounts of target protein to be expressed upon induction with IPTG. Here suitability of three different *E. coli* strains for expression of XTH31 was assessed, with the results summarised in Figure 51. It was found that the BL21 strain was suitable for expression of XTH31 since prominent bands of the expected size (predicted XTH31 sizes 42 kD, observed bands size between 35-55 kD) were detected on

SDS-PAGE gel upon IPTG induction. As shown in Figure 52, other (putatively unspecific) proteins were still present after purification, suggesting optimisation of the purification protocol is required. As a first step towards this goal, more washing steps with low concentrations of imidazole (about 10-50 mM imidazole) could be tried. Although it seems like the target protein can be expressed heterologously via pET28 in BL21 cells, correct folding and activity of the target protein have to be confirmed. This is particularly important taking into account that XTH31 has two predicted disulphide bonds, which could be required for correct folding of the enzyme and generation of activity. If no active protein is produced it is possible to use other pET vectors and/or other host strains to promote the formation of disulfide bonds to obtain correctly folded, active protein (pET manual Novagen). One such vector would be pET32, which allows a thioredoxin fusion protein to be used in order to enhance yields of target proteins. Using the pET32 vector could be combined with alternative host strains such as AD494 or BL21trxB deficient for thioredoxin reductase (competent strains Novagen). Other possible optimizations include using different temperatures and times of induction, buffers used for protein extraction or different cloning sites for expression of different soluble, active forms of the target protein.

Although the biochemical characterisation for XTH31 has not been achieved in this study, at the time of writing up, it was found that another lab had been characterising the enzyme activity from *P. pastoris* protease deficient strain SMDH118H and it has been observed that XHT31 shows XEH activity (Kaewthai, 2011). Notably, it was also found that although the active form of the heterologous enzyme was obtained, mass spectrometry results showed the target protein was expressed as a mixture of some full length and truncated proteins without 6xHis tag or c-myc epitope (Kaewthai, 2011). This is in accordance with previous work for XTH function where it has been proposed that XTH groups 1 and 2 tend to exhibit transglycosylase activity while members of XTH Group3 predominantly show hydrolase activity (Baumann et al., 2007; Rose et al., 2002). This suggests a role for XTH31 in XGs precursor modification and/or turnover of XGs (Scheller & Ulvskov, 2010).

5.7.2 Isolation of the T-DNA insertion lines for *xth31* mutations

To understand the biological functions of *XTH31*, T-DNA insertion lines provided from the stock centre were used as a tool. Lines with putative insertion sites in the exon, SK_046167 (*xth31-1*) and SK_129686 (*xth31-2*), were selected. The former line appears to have the insertion site just before Glu₁₁₈, the proton donor active site; thus, possibly a truncated protein could be expressed which has only the Glu₁₁₄ nucleophile active site. The latter one appeared to have the insertion three bases after the first line, in other words, just after the Glu₁₁₈, which provides both nucleophile and proton donor active sites. However, in both mutants, any truncated protein expressed would not have the predicted disulfide bonds, implying that the truncated versions would be improperly folded.

To isolate plants homozygous for the insertions, PCR-based analysis was used as described before. This high throughput technique allows large-scale analysis of samples. As shown in Figure 53, the PCR genotyping identified plants homozygous for the individual insertions. However, this technique just provides evidence that T-DNA insertion has occurred in the target gene. To determine the precise position of the insertions, fragments from the genotyping reactions were sequenced and the results confirmed the predicted insertion sites. These experiments successfully identified plants homozygous for the insertions selected and confirmed the positions of the insertions.

5.7.3 Analysis of *XTH31* expression disruption using RT-PCR

In order to investigate whether the T-DNA insertions would affect expression of *XTH31*, semi-quantitative RT-PCR reactions using gene specific primers were performed. Reactions with *UBQ10* specific primers were included as controls and the results showed that a comparable amount of total RNA templates were used in these experiments Figure 54. The semi-quantitative RT-PCR data suggested that in Col-0 seedlings, *XTH31* transcript levels were down regulated upon isoxaben treatment. For *xth31-1* seedlings, no expression could be observed while some expression was observed from *xth31-2*, suggesting that *xth31-1* represents a knockout while *xth31-2* is a knockdown for *XTH31*.

However, quantitative RT-PCR should be performed to correctly quantify the expression levels of *XTH31* in the mutant alleles and compare *XTH31* expression levels between mock and isoxaben-treated seedlings over time.

5.7.4 Phenotypic characterisation of *xth31* mutants

In order to investigate the biological functions of genes of interest in plants, many genetics tools like inducible RNAi, overexpression and/ or knockout or knockdown lines can be used or generated.

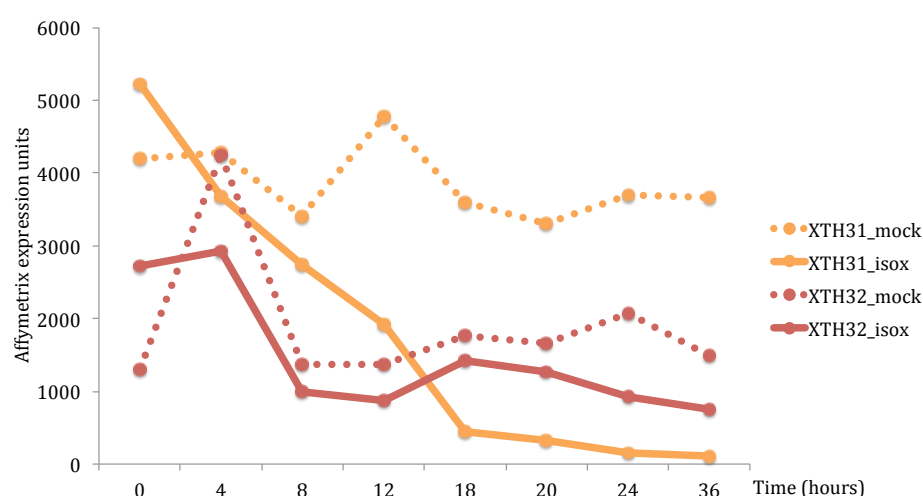


Figure 67 Microarray expression data for *At3g44990* (*XTH31*) (orange) and *At2g36870* (*XTH32*) (red) under mock treatment (dash line) and isoxaben treatment (bold line). Data derived from Hamann et al. (2009).

Mutant phenotypes of loss of function alleles of gene family members can be suppressed by other family members capable of compensating for the loss of gene activity. The available expression data from the previous expression profiling work was used to study the transcript levels of other XTHs that could cause redundancy (Figure 67) (Hamann et al., 2009). The expression profiling data shown in Figure 67 for *XTH32* suggests that isoxaben treatment does not like the treatment affect the expression. Therefore it is conceivable that these family members will not cause any functional redundancy during the response to isoxaben treatment.

In this study, T-DNA insertion lines were isolated to understand the biological function of *XTH31*. For phenotypic analysis, the results from growth

stage-based analysis showed no obvious phenotypes. In accordance with the growth stage-based analysis results, rosette diameter of the mutants from plants that showed rosette growth completed showed no significant difference compare with Col-0 wild type. However, biomass measurement (fresh weight and dry weight) from the plants at the same developmental stage showed lower biomass in the mutant plants than in wild type. However, no significant difference in the biomass phenotype was observed from senescence plants, suggesting the role of *XTH31* gene at early stage of the development. Disruption of *XTH31* clearly affects plant growth as the biomass shown for rosette completed growth stage plants. However, the questions of how the gene involved or regulate early stage of plant growth are still remained. Further investigation and more in depth experiment to observe any phenotypes that could be shown in tissue or cellular level and more imaging and immune labelling techniques could provide better understanding in function of *XTH31* in plant growth and development.

Expression data for the gene deposited in the Genevestigator database suggested that *XTH31* is highly expressed in the endosperm as well as in imbibed seeds and radicles of seedlings, implicating *XTH31* in seed germination. Therefore germination assays were performed to investigate whether disruption of *XTH31* gene expression would affect plant germination. Only seeds for the *xth31-1* allele exhibited delayed germination phenotypes with 30 % of them germinating on day one while about 60 % of Col-0 and the *xth31-2* seeds germinated. After the first day no significant differences were observed suggesting that *XTH31* may play a role during the early stage of germination. The expression data available in the Genevestigator database suggests that *XTH31* is also expressed in roots. Therefore, root lengths were measured in four day-old seedlings. The measurements showed that seedling roots for the *xth31-1* allele were significantly longer than Col-0 and *xth31-2* roots. This implicates *XTH31* in seedling root growth. To investigate whether the target gene is involved in root cell division or cell elongation, cell number and cell size in the root elongation zone could be determined. However, no other obvious phenotypes were observed for seedlings and plants for both *XTH31* alleles grown under standard conditions.

To determine if loss of XTH31 affects the response to isoxaben root growth assays and lignin deposition were characterised in seedlings short and long-term isoxaben treated. It was observed that mutant seedling roots responded to the treatment in a concentration dependent manner that differed from Col-0 controls. As shown in Figure 60, after 1 nM isoxaben treatment, seedlings roots for both alleles seem to be more resistant to the treatment than the wild type. At 2 nM concentration of isoxaben, no significant difference compared to Col-0 was observed for the *xth31-1* allele while the *xth31-2* showed enhanced resistance to isoxaben. At 3 nM isoxaben, *xth31-1* seedling root growth was more sensitive (root length decreased by 50 %) to the treatment while the *xth31-2* allele responded similarly to wild type (root length decreased by 25 % - 30 %).

Analysis of lignin deposition in seedlings treated with high concentrations of isoxaben for 12 or 24 hours showed that both mutants have more lignin deposition in the elongation zone of seedling roots. However, for the seedlings grown on plates of ½ MS with low concentration of isoxaben (0, 1, 2 and 3 nM), just some cells in the inner cell layers were stained. Notably, cell shapes in the roots of the mutant seedlings seem to be affected more than wild type (for example roots treated with 3 nM isoxaben). Moreover, for 3 nM isoxaben treatments, the morphological analysis suggested that the cell division zone near the root cap was absent from *xth31-1* seedling roots. Measurement of the stained area from the roots also confirmed that *xth31* mutants are more sensitive to isoxaben than Col-0 wild type, suggesting the role of *XTH* that involved in reinforcement of the walls during cell wall damage. For the lignin staining, the phenotypes from both *xth31-1* and *xth31-2* could be observed. While no phenotypes showed for *xth31-2* from the growth stage-based analysis, rosette diameter measurement and biomass measurement due to this line appeared to be a heterozygous. The explanation could be that the lignin phenotypes observed probably caused by T-DNA insertion disrupts other genes that involved in lignin deposition when the walls are impaired. However, this hypothesis needs to be confirmed whether it caused by partial disruption of *xth31* or the background genes by rescue the *XTH31* gene in these mutant background.

Although no obvious growth phenotypes have been observed in the mature plants grown under standard conditions, cell wall composition was investigated as described. As stated before that no striking phenotype does not imply to no phenotypes at all such as in the case of *xxt1 xxt2* plants, which are totally deficient in xyloglucan showed only subtle growth phenotypes (Park & Cosgrove, 2012b). In our case, cell wall phenotypes were observed for mature plants. Stems from mutant plants showed significant differences in cellulose (Figure 64), uronic acid (Figure 65) and neutral cell wall sugar contents (Figure 66) compared with wild type controls. For the knockout mutant, *xth31-1*, the mutant showed 10 % more cellulose than Col-0 with approximately 8 % less uronic acid and significantly higher amounts of all neutral sugars except for xylose. While no significant differences were observed for cellulose and uronic levels for *xth31-2*, the mutant showed significantly elevated levels of neutral sugars compared to wild type. These results indicated that disruption of *XTH31* leading to more crystalline cellulose with minor changes in hemicellulose but less pectin. Suggesting possible role of the gene could be involved in regulation of crystallisation of cellulose. One hypothesis could be that the gene is involved in xyloglucan turn over that provides load-bearing structure to support cellulose microfibrils. Knock out the *XTH31* leading to loosing of proper cell wall integrity maintenance, thus, more crystalline cellulose needed to reinforce the walls. Another hypothesis is the gene is involved in pectin biosynthetic or regulation of the pathways related to pectin biosynthesis, hence, knocks out of the gene leading to less pectin in the walls as less uronic acid was observed.

Chapter 6 Functional characterisation of UGT71C3

6.1 Functions of plant UDP-glycosyltransferases (UGTs)

UDP-glucosyltransferases (UGTs) catalyse the transglycosylation of sugar nucleotides to a wide range of acceptors (aglycones) (Ross et al., 2001). Plant UGTs use UDP-glucose for the glycosylation of hormones, secondary metabolites and xenobiotics. As a result of the glycosylation, aglycones could be either activated or inactivated (Gachon et al., 2005). In addition, the ability of the compounds to be transported by active membrane transport systems, which recognise specific glycosides but not aglycones, could be altered through attachment of glycosides (Li et al., 2001). Phylogenetic analysis of the UGT family members showed that the UGTs co-evolved with the plants since the appearance of land plants (Caputi et al., 2012).

6.2 UGTs in *Arabidopsis*

Based on amino acid sequences, a hundred and seven *Arabidopsis* UGTs have been identified and categorised into fourteen distinct groups (A-N), which evolved from fourteen ancestral *UGT* genes (Figure 68) (Ross et al., 2001). The candidate gene *At1g07260* encodes the predicted UGT71C3. Amino acid alignments for UGT71C3 showed that the three *Arabidopsis* UGT71C family members (UGT71C1, C2 and C4) share the greatest sequence similarity with more than 60 % identity across the whole protein (Li et al., 2001; Ross et al., 2001). UGT71C family members are predicted to function as quercetin 3-O-glucosyltransferases (Lim et al., 2004). In addition, the alignment also showed proteins from *Arabidopsis*, poplar, grape, castor, Medicago and soy bean with high sequence similarity to UGT71C3 (50 % - 60 % identity) (Hou et al., 2004). Additionally, the alignment results identified the closest related clades that contain putative UGTs from *Arabidopsis*, Poplar, grape, castor oil and ornamental plants like *Lobelia* and *Dianthus*.

The evolution of substrate recognition and regioselectivity of glycosylation mediated by UGT family 1 has been explored for four of the fourteen groups (group D, E, H and L) (Lim et al., 2003^a). Using the two hydroxycoumarin substrates, esculetin and scopoletin, ninety enzymes have been assayed. The results showed that forty-eight of the ninety enzymes recognised the tested substrates with different efficiencies. Chemical structures of the substrates and possible products are shown in Figure 69.

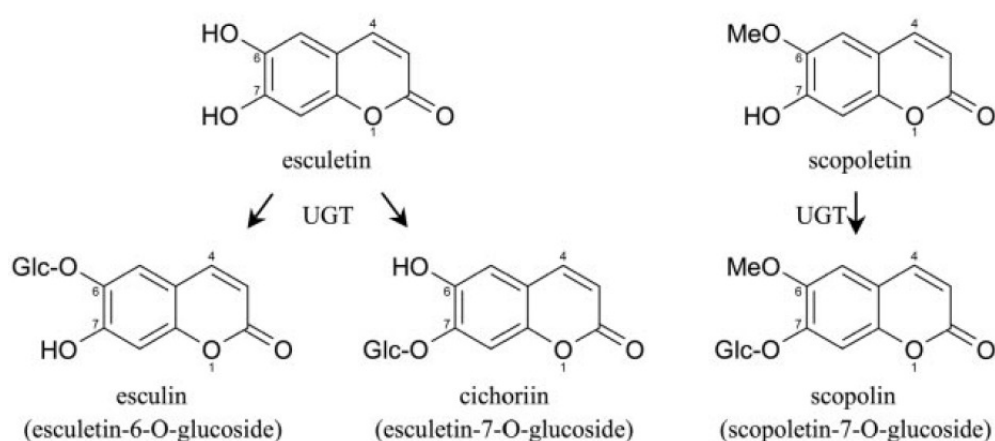


Figure 69 The chemical structures of two hydroxycoumarins, esculetin and scopoletin, the substrates used for functional UGT assays and the three possible products deriving from glycosylation at different hydroxyl positions (Lim et al., 2003^a).

For UGT71C3, a preference for the 7-OH position of esculetin, with 0.5 nkat mg⁻¹ of activity for cichoriin production was detected while less than 0.1 nkat mg⁻¹ was detected for esculin (Lim et al., 2003^a). Similarly, four other members of the UGT71C enzyme family show the same regioselectivity for esculetin. In addition, it was shown that all UGT71C (71C1-C5) family members are active when scopoletin is used as a substrate. Based upon the activity screening of UGT family 1 in *Arabidopsis* by Lim et al. (2003^a), it was suggested that phylogenetic relationships are likely to correlate with substrate recognition. Notably, it is not always the case that family members from the same clade catalyse the same reaction. This is important to bear in mind since previous results show a significantly lower activity level for UGT71C3 compared to other family members. However, both esculetin and scopoletin could be potential substrates for activity assays with UGT71C3 as a first step towards optimisation of the assays. Additionally, to catalyse glycosyltransferases,

standard glycosyltransferase activity can be measured by carrying out free sugar assays, in which transfer of ^{14}C -labelled sugars from nucleotide sugar donors onto acceptor sugars is being detected (Bückner et al., 2000).

6.3 Heterologous expression of UGT71C3

In order to characterize the biochemical properties and functions of the candidate target proteins, heterologous expression is a standard strategy. Bearing in mind the predicted post-translational modification sites and previous attempts at expressing UGT71C3, it was decided to initially attempt to achieve expression in strain GS115 of *P. pastoris*. For cloning of the expression constructs, pPICZ was used for the non-secreted version and pPICZ α for the secreted version. However, the cloning attempt was not successful, therefore, an alternative strategy was initiated. The coding region for UGT71C3 was amplified by PCR and prepared for expression in *E.coli* using the TOPO[®] TA Cloning[®] (Invitrogen) (Figure 70).

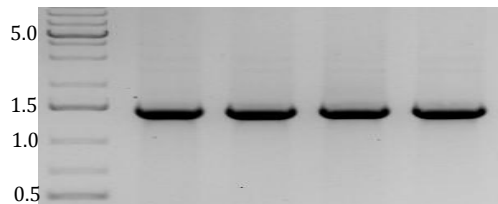


Figure 70 Results from PCR reaction amplifying *UGT71C3* and introducing *Bam*HI and *Not*I restriction sites using Phusion[®] DNA polymerase. Size of DNA ladder fragments shown in kb.

The fragments were extracted from Agarose gel and cloned into the pCR2.1 TOPO vectors as described in manufacturer's protocol. Figure 71 shows results of the colony PCR from ten pCR2.1UGT71C3 clones. All tested clones were positive for the *UGT71C3* fragment.

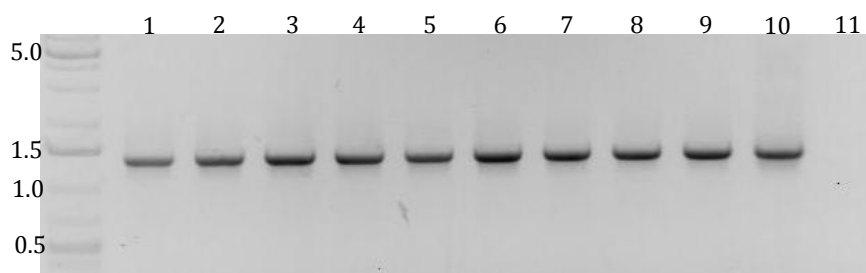


Figure 71 Results from colony PCR to identify pCR2.1UGT71C3 positive colonies. Size of DNA ladder fragments shown in kb, expected fragment size: 1.5kb

Three clones were grown overnight for midi-prep plasmid preparation and restriction digest using *Bam*HI and *Not*I. The results for restriction digest are shown (Figure 72).

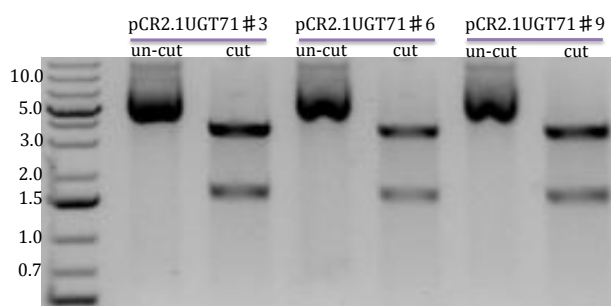


Figure 72 Results of *Bam*HI-*Not*I double digest to generate the pCR2.1UGT71C3 construct. Size of DNA ladder fragments shown in kb.

Construct number 9 (pCR2.1GUT71#3) was ligated into pET28b⁺ vector after digest with *Bam*HI and *Not*I. Colony PCR reactions were performed using T7 forward primer and a reverse primer specific for UGT71C3 primer sequences (Appendix1). The expected fragment size is 1.5 kb whereas no product would be amplified from empty plasmids. Figure 73 shows that all tested clones were positive.

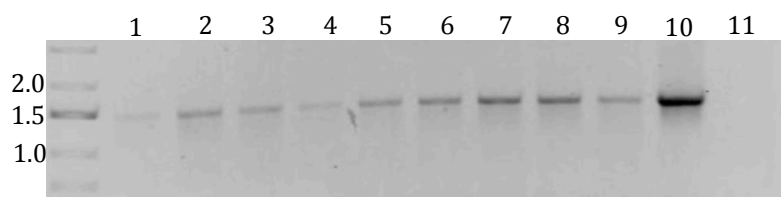


Figure 73 Results from a colony PCR to identify clone containing pET28UGT71C3 (1-10) are shown. Size of DNA ladder fragments shown in kb, expected fragment size: 1.5 kb.

After confirming the presence of the insert via colony PCR, three clones were selected and digested with *Bam*HI and *Not*I to confirm correct insertion of the fragment into the vector (Figure 74).

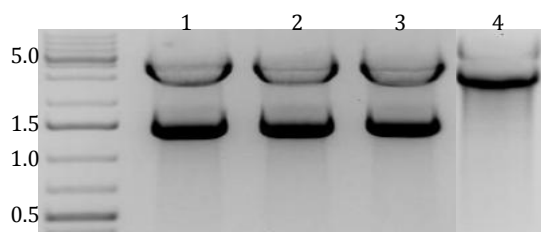


Figure 74 Results from a *Bam*HI and *Not*I double digest of pET28UGT71C3 (1-3) and pET28 empty plasmid (4) are shown. Size of DNA ladder fragments shown in kb, expected fragment size: 1.5 kb.

The results suggested that all three clones tested contain the correct construct. Clone number 3 was selected for sequencing. The sequenced as correct confirmed pET28UGT71C3 construct was then used as expression construct in the following experiments.

The expression construct for UGT71C3 was transformed into three different *E. coli* strains: BL21, SoluBL21™ and BL21 (DE3). After antibiotic selection and colony PCR, positive strains were used for expression trials. The strains were tested for suitability of protein expression with different concentrations of IPTG and for different periods. Figure 75 shows results from an immuno dot blot for pET28UGT71C3 expressed in the BL21 strain.

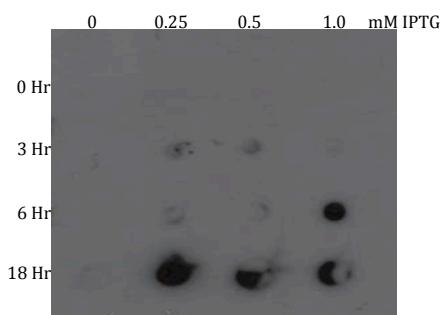


Figure 75 Immuno dot blot for UGT71C3 expressed under 0, 0.25, 0.5 and 1 mM IPTG induction for 0, 3, 6 and 18 hours.

The immuno dot blot results showed the strongest signal from a sample treated with 0.25 mM IPTG for eighteen hours.

In parallel with the immune dot blot screening, SDS-PAGE was used to separate the proteins expressed and test if a prominent band with the correct size (53 kD) for UGT71C3 could be identified. Figure 76 shows results from a representative SDS-PAGE analysis of four selected clones from three different strains used. No obvious bands could be identified in the expected size range in the samples from the induced strains.

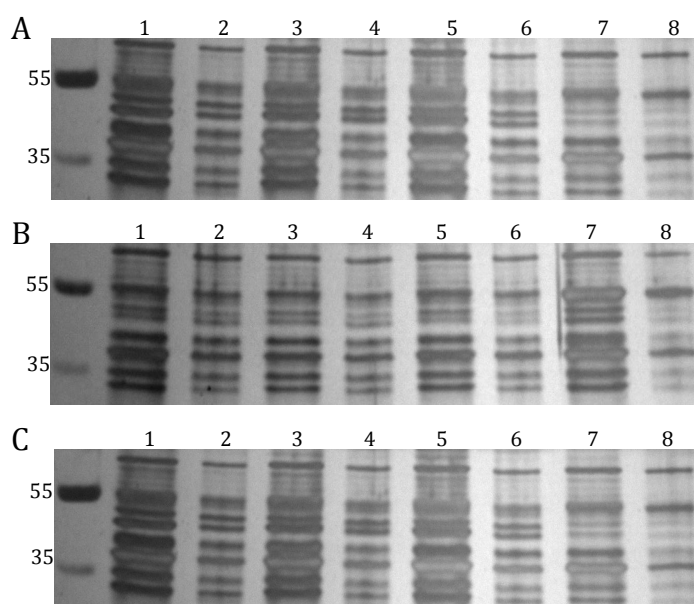


Figure 76 SDS-PAGE analysis of three different *E.coli* strains BL21 (A), SoluBL21™ (B) and BL21 (DE3) (C). Four different clones were tested per strain; clone1 = lane 1 and 2, clone2 = lane 3 and 4, clone3 = lane 5 and 6 and clone4 = lane 7 and 8. Samples from 0 hour (lane 1, 3, 5 and 7) and 6 hours (2, 4, 6 and 8) of IPTG induction were shown. Protein markers shown in kD.

Purification of the target protein was performed using Protino® Ni-TED Packed Columns (MACHEREY-NAGEL GmbH & Co. KG) as described in the manufacturer's protocol.

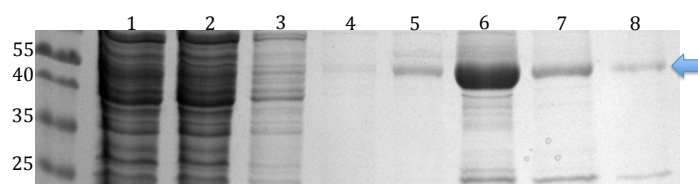


Figure 77 SDS-PAGE for UGT71C3 after purification for His-tagged proteins. Shown above are cell lysate (1), Flow-through (2), 1st Wash (3), 4th Wash (4), 1st Elution (5), 2nd Elution (6), 3rd Elution (7) and 4th Elution (8). Target proteins are indicated with the arrow. Size of protein marker shown in kD.

After performing the purification, some un-specific proteins are still observed but in very small amounts compared to the target protein (Figure 77). By the time of thesis writing up, the activity assay had not yet been achieved.

6.4 Isolation of homozygous insertion lines

Figure 78 shows an example for genotyping results for two insertion lines for *At1g07260*. To identify plants homozygous for the respective insertions, PCR reactions with gene and T-DNA specific primers have been performed as described. From plants homozygous for the insertion, no wild type fragments can be amplified while the fragment for the insertion can be detected.

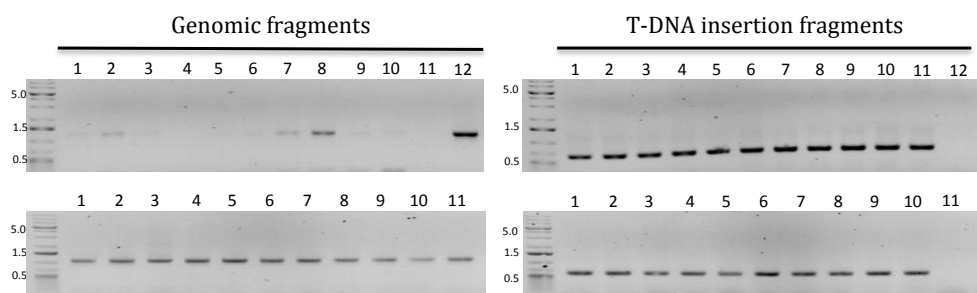


Figure 78 PCR-based screening to isolate homozygous *ugt71c3* plants. Gene specific fragments and T-DNA fragments were amplified from eleven independent mutant plants (2-12) and Col-0 (13). Size of DNA ladder fragments shown in kb.

6.5 Analysis of *UGT71C3* expression using RT-PCR

After isolation of homozygous plants, RT-PCR was performed to determine whether expression of *UGT71C3* is affected. By using primers specific for the gene, the whole coding region was amplified and reactions for *UBIQUITIN10* were performed in parallel as loading control (Figure 79).



Figure 79 Semi-quantitative RT-PCR for *UGT71C3* from Col-0 and the insertion lines (left). *UBIQUITIN 10* was used as loading control (right). Col-0 mock- and isoxaben-treated shown in (1) and (2), *ugt71c3-1* mock and isoxaben treated shown in (3) and (4) and *ugt71c3-2* mock and isoxaben treated are shown in (5) and (6). Size of DNA ladder fragments shown in bp.

The results show the expression of *UGT71C3* only from Col-0 with lower expression in mock treated sample than the isoxaben treated one. No expression at the transcription level can be detected in the mutants by RT-PCR, suggesting both insertion lines are knockouts for *UGT71C3*; therefore, the mutants were used for phenotypic studies in following sections.

6.6 Phenotypic characterization of the mutant plants

After homozygous plants for *ugt71c3* mutants were isolated and loss of *UGT71C3* expression confirmed with RT-PCR, plants were grown on soil under standard conditions along with the Col-0 wild type. Phenotypes of the plants from various developmental stages throughout the life cycle of plants were determined as described in materials and methods (2.2.5.2). The results for the time required at each growth stage of Col-0 wild type and *ugt71c3* mutants are shown (Table 11).

Table 11 Growth stage-based analysis results for the wild type and *ugt71c3* mutants.

Developmental stages	Col-0	<i>ugt71-1</i>	<i>ugt71-2</i>
Leaf development	4-5 days	4-5 days	4-5 days
2 rosette leaves > 1 mm in length	6 days	6-8 days	6-8 days
3 rosette leaves > 1 mm in length	7-9 days	8-9 days	8-9 days
4 rosette leaves > 1 mm in length	8-10 days	10-11 days	10 days
5 rosette leaves > 1 mm in length	11-12 days	11-12 days	11 days
6 rosette leaves > 1 mm in length	11-13 days	13-14 days	12-13 days
7 rosette leaves > 1 mm in length	13-14 days	13-15 days	14-15 days
8 rosette leaves > 1 mm in length	13-15 days	15-16 days	15-16 days
9 rosette leaves > 1 mm in length	14-15 days	15-17 days	17 days
10 rosette leaves > 1 mm in length	14-16 days	18-19 days	17-19 days
11 rosette leaves > 1 mm in length	17-18 days	19 days	18-19 days
12 rosette leaves > 1 mm in length	17-19 days	19-20 days	19 days
13 rosette leaves > 1 mm in length	18-19 days	19-21 days	20 days
14 rosette leaves > 1 mm in length	20-21 days	20-22 days	20-21 days
Rosette growth complete	21-22 days	21-22 days	21 days
Flower production	19-20 days	19-22 days	19-21 days
Senescence complete	60-65 days	60-63 days	60-63 days

The *ugt71c3* mutants grown under standard conditions seemed to grow larger than the corresponding Col-0 controls. However, the variations within each line are observed. Once all the plants were at the stage where the rosette

growth completed, rosette diameter was measured. Results showed no significantly differences among wild type and *ugt71c3* mutants (Figure 80).

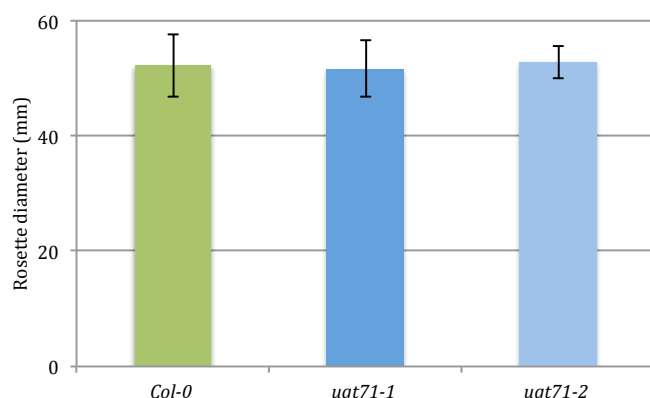


Figure 80 Rosette diameter measurement for *kor2* mutant plants that showed rosette growth complete stage. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average rosette diameter calculated from three biological replicates where the data was analysed using sixteen individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

To determine if the apparent growth phenotype results in increased biomass production, four week-old plants for both *ugt71C3* alleles and the Col-0 control were characterised (Figure 81). Plants for both *ugt71c3* alleles showed significantly lower values for both fresh weight and dry weight. For *ugt71c3-1*, fresh weight of the mutant was about 30 % lower and the dry weight was about 15 % lower than wild type. For *ugt71-2*, the mutant fresh weight was about 25 % lower while the dry weight was about 10 % lower than wild type.

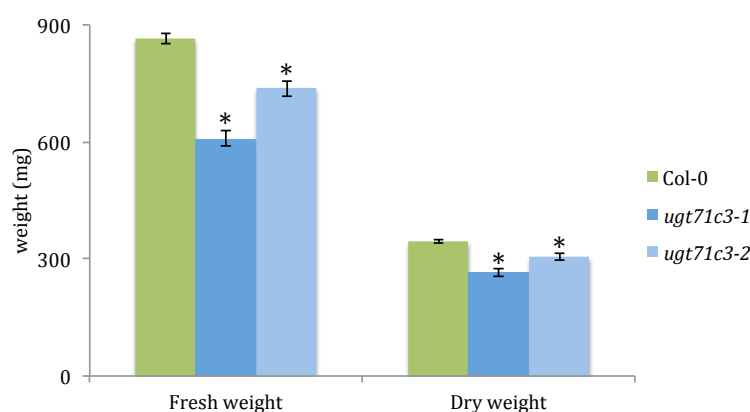


Figure 81 Biomass measurement for four week-old *ugt71-1*, *ugt71-2* and Col-0 wild type plants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight of either fresh or dried materials and calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference for the data set compare with wild type (student t-test, $P < 0.05$).

In addition to the measurement of biomass for plants at the growth stage where the rosette growth completed, the biomass for senescence plants were measure and the dry weights are shown (Figure 82). Although significantly differences were observed for the mutant plants at the rosette-completed stage, no significantly differences was observed for the senescence plants.

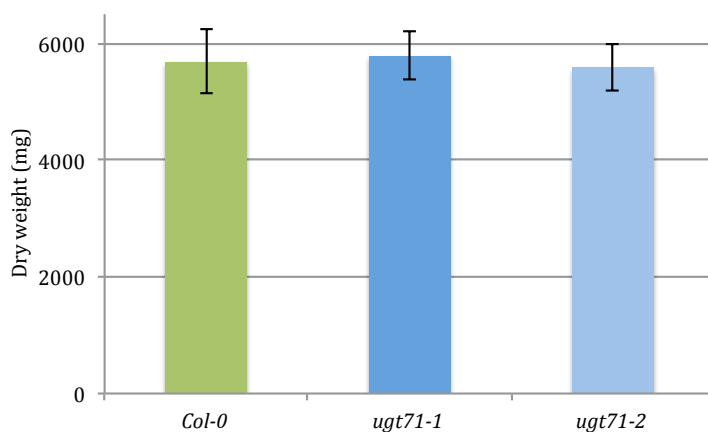


Figure 82 Biomass measurement for senescence *ugt71c3* mutants. Plants were grown under long-day period of light on soil as the growth conditions described. The values represent average weight calculated from three biological replicates where the data was analysed using eight individual plants $\pm$ SE. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

6.6.1 Characterisation of germination rates of *ugt71c3*

The microarray-based expression profiles available in the Genevestigator database suggest that *UGT71C3* is expressed in the radicle and imbibed seeds. To investigate whether knocking out *UGT71C3* will affect germination, the germination rates of mutant and wild type seeds were determined. Figure 83 shows the percentages of germinated seeds for *ugt71c3* and Col-0 for the first three days after vernalisation. Nearly all the *ugt71c3* seeds had germinated after the first day while only 60 % of wild type seeds had at this point in time.

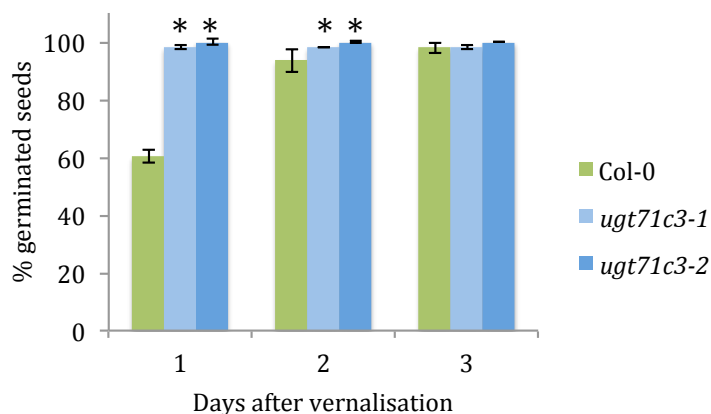


Figure 83 Results from germination assays for *ugt71* mutants and Col-0 wild type. Sterile seeds were put on ½ MS plates and stratified at 4°C for forty-eight hours before grown in standard conditions as described. Seeds with radicals emerged were counted as germinated seeds after 1-3 days after plates were put under standard conditions. The values represent average percentage of germination calculated from three biological replicates where the data was analysed using fifty-two individual seeds ± SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

6.6.2 Analysis of root growth of mutant plants

The available gene expression data on the Genevestigator database also suggested that *UGT71C3* is expressed in seedling roots. Therefore, root length of the mutants were measured and compared with Col-0 wild type. Four day-old seedlings grown on ½ MS medium supplemented with 1% sucrose under the conditions described were used for the measurement. Figure 84 shows the results of these measurements with no significant differences detectable between both mutant alleles and the wild type controls.

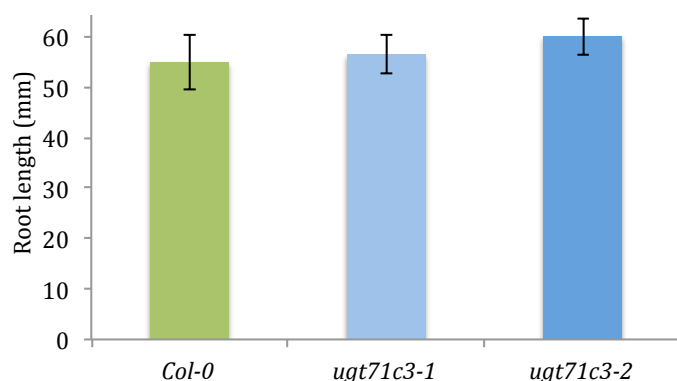


Figure 84 Root of seven day-old *ugt71* mutants and Col-0 wild type seedlings grown vertically on ½ MS plates were captured and the photos were analysed using ImageJ software. The values represent the average of root length of seedlings calculated from three biological replicates where the analysis was performed using at least twenty individual seedlings ± SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test, $P \leq 0.05$).

To determine whether UGT71C3 gene activity is required for the long-term response to cellulose biosynthesis inhibition, the effects of 1, 2 and 3 nM isoxaben on root growth were determined. Figure 85 indicates the effects as percentage decrease in root length. At 1 nM isoxaben, *ugt71c3-1* roots appeared to be more sensitive to the herbicide than Col-0, while, *ugt71c3-2* was slightly more resistant. At 2 and 3 nM isoxaben no significant differences were observed between mutant and wild type seedlings were observed.

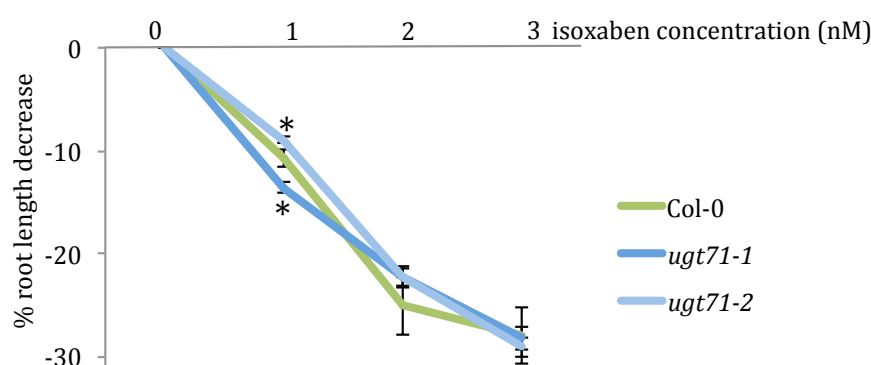


Figure 85 Percentage of root length decreased up on isoxaben treatment. Roots of *ugt71* mutants and Col-0 wild type seedlings grown vertically on ½ MS plates for four days before transferred to fresh ½ MS plates in presence or absence of 1, 2 or 3 nM isoxaben were analysed. Photos were captured as described and analysed using ImageJ software. Percentage of root length decreased were calculated from four biological replicates where the analysis was performed using at least ten individual seedlings ± SE. * indicates significant difference of mean between data sets compared to Col-0 wild type under the same treatment (student t-test, n ≥40 per line and condition, P ≤ 0.05).

In addition to the root growth measurements, staining of ectopic lignification from roots of seedlings were performed since it has been shown in previous study that phloroglucinol can be used as another reliable indicator of cell wall damage (Hamann et al., 2009).

6.6.3 Lignin deposition upon exposure to isoxaben

Microarray-derived expression profiling data suggested that *UGT71C3* transcript levels are increased by isoxaben treatment. Isoxaben treatment induces ectopic lignin deposition in the elongation zone of primary seedling roots (Hamann et al., 2009). To assess the function of *UGT71C3* in the response to isoxaben, mutant and wild type seedlings were grown in liquid medium for six days and were then treated either with DMSO (mock treatment) or 600 nM

isoxaben for 12 and 24 hours. The treated seedlings were then stained with phloroglucinol to detect ectopic lignin deposition. As shown in Figure 39 panel A, the *ugt71c3* seedling root tips appeared to exhibit more lignin deposition than the Col-0 controls (Figure 86 A, compare areas of red staining between *ugt71c3* and Col-0 seedlings). To compare the stained area of the roots quantitatively, the measurement and calculation for percentage of the stained area in total area observed were performed as described (2.2.5.4) and the results are shown (Figure 86 B).

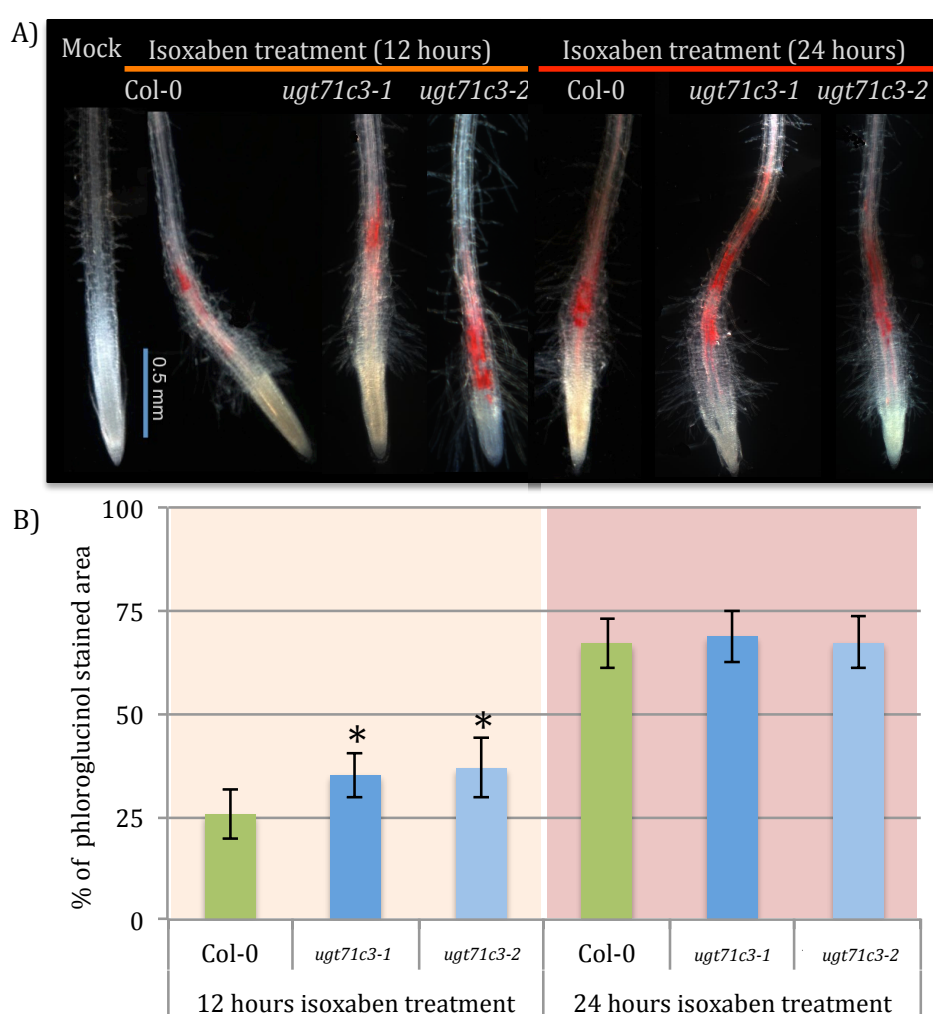


Figure 86 Phloroglucinol stained seven day-old *ugt71c3* and Col-0 seedling root tips after isoxaben treatment for 12 or 24 hours. A) is the representative results of the phloroglucinol staining. Scale bar is 0.5 mm, for each genotype/treatment at least twenty seedling roots were analysed. B) Percentage of the phloroglucinol staining for seven day-old *ugt71c3* and Col-0 seedling root tips after isoxaben treatment for 12 or 24 hours.

The results from 12 hour-treatment showed that all *ugt71c3* seedlings appeared to have significantly more lignin staining than the wild type controls while no significant difference was observed for the seedlings treated for 24 hours (Figure 86 B). In addition to the lignin stained observed, it was shown that seedlings treated with isoxaben were exhibited cells swollen leading to slightly bigger root area compared to ones under mock treatment.

In addition to the analysis of ectopic lignification in seedlings grown in liquid culture system, seedlings grown on plates and treated with 1, 2 and 3 nM isoxaben were also stained with phloroglucinol to detect ectopic lignin deposition (Figure 87). The analysis of the stained seedlings shows that at 1nM no ectopic lignin deposition is detectable. In seedlings treated with 2 nM and 3 nM isoxaben, ectopic lignin deposition was detectable in areas that could be cortex cells. Interestingly, the staining seems to be more intense in the mutant than in the wild type roots. Moreover, the root morphology of the mutants seems to be more sensitive to isoxaben than Col-0. At 2 nM isoxaben, the mutant roots seem to be more swollen and the root tips are smaller than the wild type controls. At 3 nM isoxaben, there was a larger area with more intense staining in the cortex and the area where cell divisions happen at the root tips disappeared in both mutants but not the wild type. Some phenotypic variation was observed among seedlings (n>20) from the same genotype, however here representative results are shown (Figure 87).

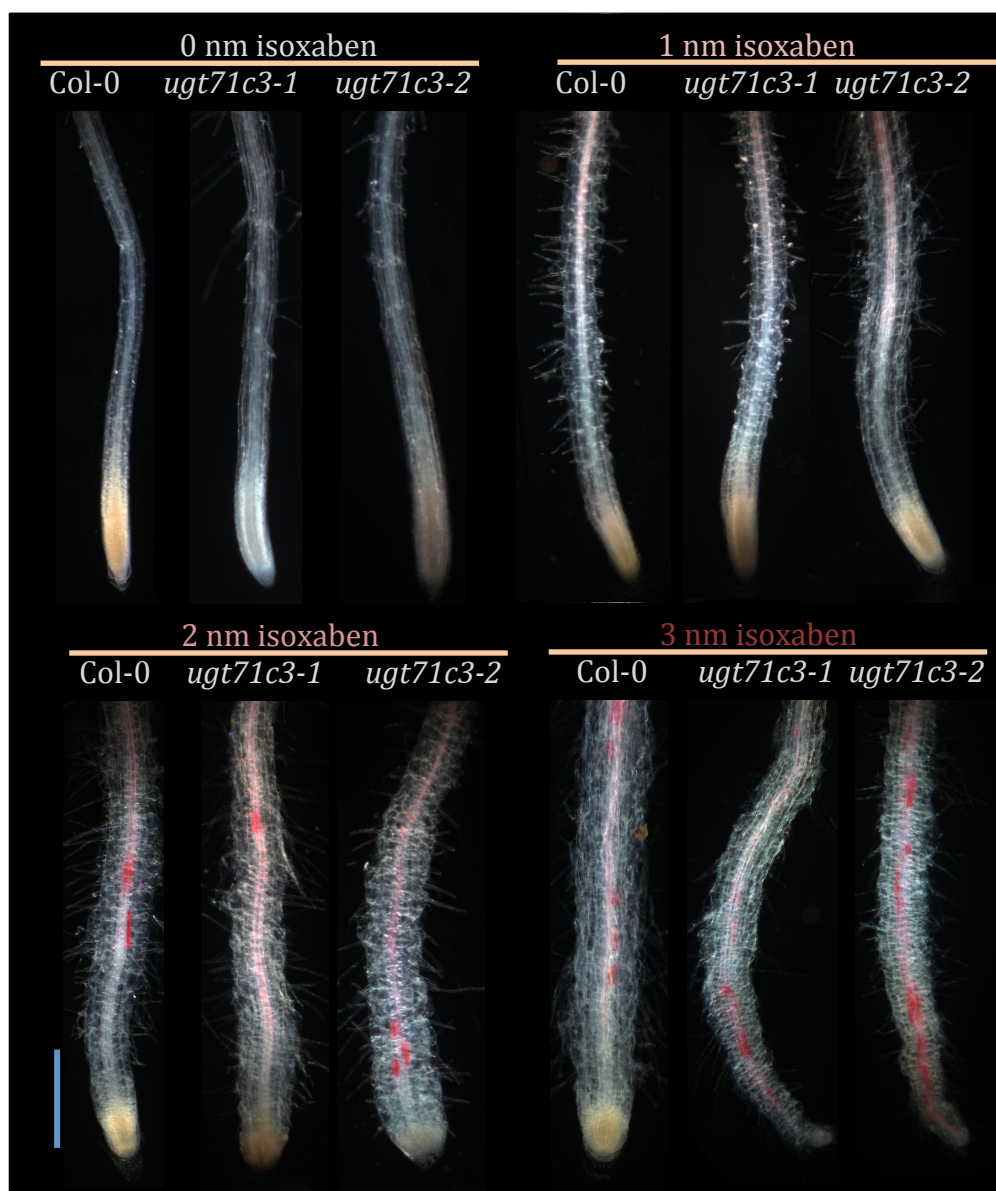


Figure 87 Phloroglucinol staining for lignin deposition from Col-0 and *ugt71c3* mutant seedlings treated with 0, 1, 2 and 3 nM isoxaben. Scale bar is 0.5 mm.

In order to quantitatively compare lignin depositions in *ugt71c3* mutants with the Col-0 wild type, lignin depositions were quantified as described (see 2.2.5.4 for the methods). Results are shown in Figure 88. Panel A shows percentage values of lignin staining for the area of stele and endodermis tissues while panel B shows the percentage lignin staining values for cortex and epidermis.

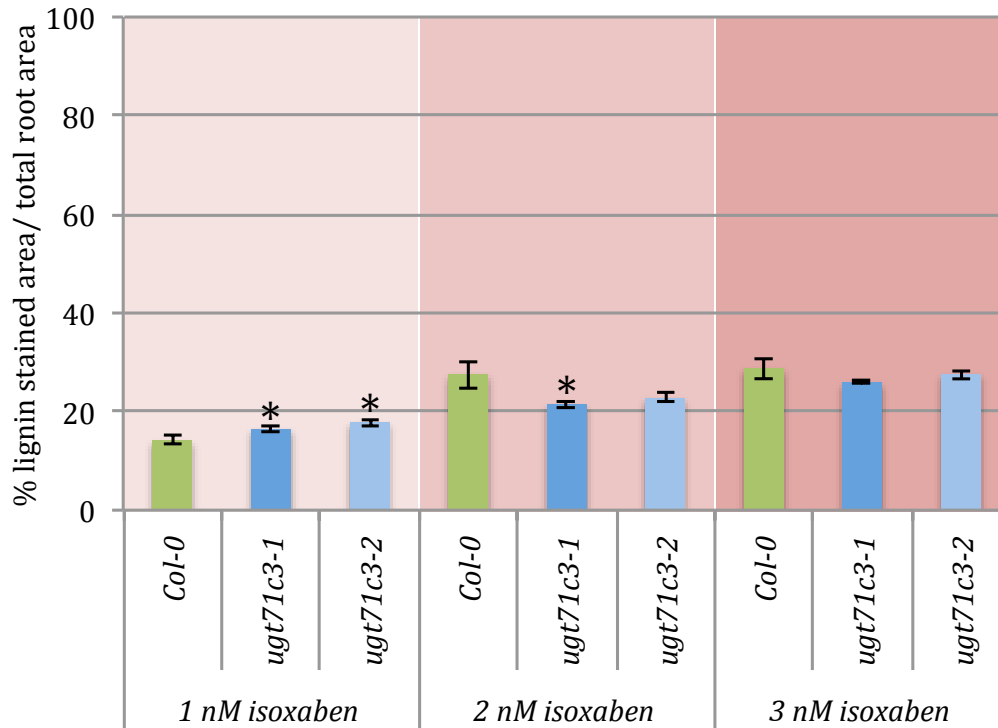


Figure 88 Percentage of phloroglucinol staining for lignin deposition in Col-0, *ugt71c3-1* and *ugt71c3-2* seedling roots treated with 0, 1, 2 and 3 nM isoxaben for three days. A) percentage of phloroglucinol stained for vascular bundle tissue. B) percentage of phloroglucinol stained for other areas along the roots. The values represent average stained area measured via ImageJ and calculated from three biological replicates where the data was analysed using at least twenty individual plants $\pm$ SD. * indicates significant difference of mean between data sets compared to Col-0 wild type (student t-test $P \leq 0.05$).

For panel A, although the variations between different alleles for *ugt71c3* mutants were observed, it was shown that for plants treated with 1-2 nM isoxaben, percentage values of lignin staining on the area of stele and endodermis for all *ugt71c3* mutants were higher than Col-0 wild type. No significant differences were observed on the staining values of stele and endodermis for the seedlings treated with 3 nM isoxaben (Figure 88, A). For the staining values on cortex and epidermis of the roots, all *ugt71c3* mutants showed significantly less lignin staining compared with wild type (Figure 88, B)

6.6.4 Analysis of cell wall composition in mutant and wild type plants

Cellulose contents from mature stems of *ugt71c3* and Col-0 plants were measured and the results are shown in Figure 89. The results showed that both mutant alleles have significantly more cellulose (approximately 10 % more) than Col-0 wild type stems.

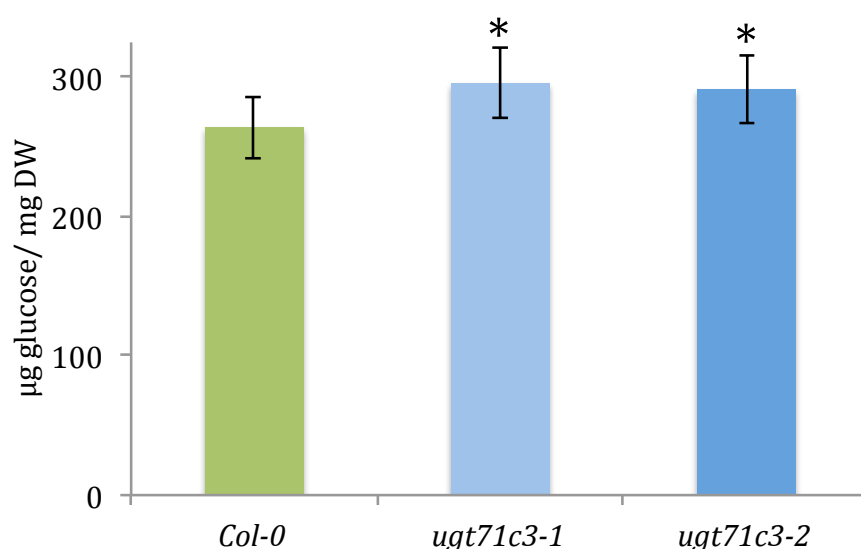


Figure 89 Cellulose measurement from stems of mature Col-0 and *ugt71c3*. The values represent average cellulose content in µg glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants for each biological replicate. * indicates statistical significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are SD values.

Uronic acids contents were also measured in stems from mature *ugt71c3* and Col-0 plants and the results are shown in Figure 90. The results showed that stems from both *ugt71c3-1* and *ugt71c3-2* plants have significantly lower amounts of uronic acids compared to Col-0.

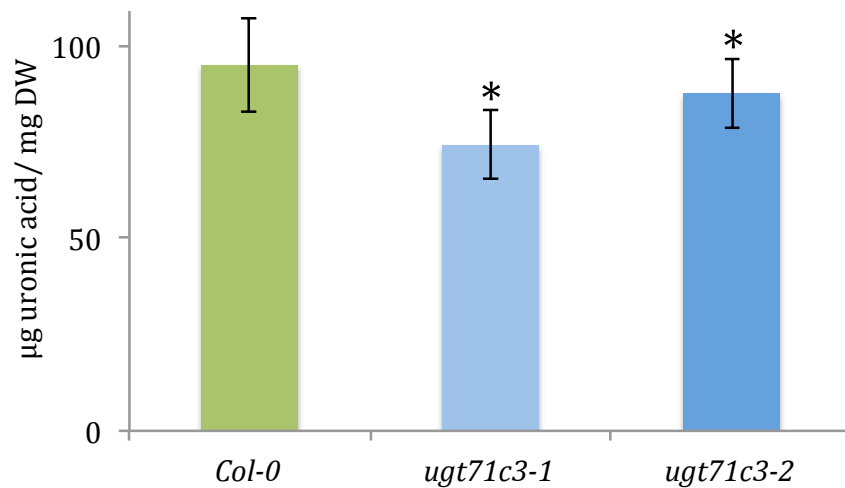


Figure 90 Uronic acids measurements in stems from mature Col-0 and *ugt71c3* plants. The values represent average cellulose content in µg glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants for each biological replicate. * indicates statistical significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are SD values.

Results for neutral cell wall sugar measurements from stems of mature plants are shown in Figure 91. Stems from both mutant alleles have significantly enhanced neutral cell wall sugar levels compared to Col-0.

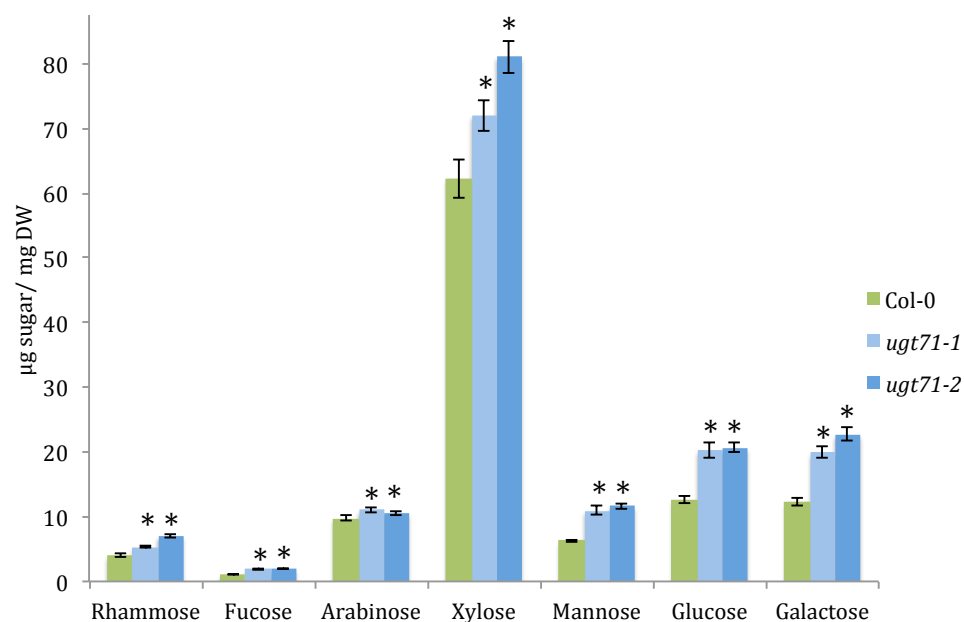


Figure 91 Neutral cell wall sugar measurement from stems of mature Col-0 and *ugt71c3* plants. The values represent average cellulose content in µg glucose per mg dry weight of plant materials which were calculated from three biological replicates where the data was analysed using four individual plants for each biological replicate. * indicates statistical significant differences of the mean value between the data set and Col-0 wild type (student t-test, $P < 0.05$). Error bars are SD values.

6.7 Discussion and summary

6.7.1 Heterologous expression of UGT71C3

In order to characterise the function of *UGT71C3*, different bioinformatics, molecular biology and genetic tools were employed. The secondary structure prediction showed thirteen phosphorylation sites and four N-myristoylation sites with no other post-translational modification sites (Appendix2). Taking into account that the enzyme is from *Arabidopsis* and a eukaryote host may be needed for heterologous expression, initially a *Pichia*-based expression system was employed. However, the enzyme could not be expressed in *Pichia* either in the secreted or non-secreted form. This could be a result of incompatible post-translational modification in the host system since the sequences of the constructs generated were confirmed to be error-free and the RT-PCR reactions performed detected expression of the constructs (data not shown). Therefore, a new attempt was made using an alternative bacterial host. Although two transmembrane regions were predicted by the PredictProtein service, other services and databases including SignalP and UniProt did not show any predicted signal sequences or transmembrane domains. Accordingly, the cloning strategy for protein expression in *E. coli* was designed to express the whole coding region of *UGT71C3*. Although no dominant target band was observed from SDS-PAGE gels, large-scale expression and purification via a cobalt column showed that a target protein of the correct size could be obtained. However, activity assays have not been performed by the time of writing up this thesis. *E. coli* has been used as an expression system for *Arabidopsis* UGTs including *UGT71C3* before and activity assays for quercetin 3-O-glucosyltransferase activity have been performed (Lim et al., 2004). However, no activity has been assigned to *UGT71C3*, implying that the substrate used in the previous *in vitro* study may not be the actual substrate in plants. To identify possible substrates/products of the target enzyme, metabolomics profiles for the knockout and/or overexpression mutants could be analysed and compared with the profile of wild type controls.

6.7.2 Isolation of homozygous insertion lines

In order to determine the biological functions of *UGT71C3*, two independent T-DNA insertion lines were selected; SK_021979 (*ugt71-1*) and SK_042564 (*ugt71-2*). The former possesses the insertion site in *UGT71C3* before the signature conserved region for plant secondary product glycosyltransferase (PSPG box). A possible truncated product of *UGT71C3* could still be expressed and generate a 236 amino acids long protein including six extended β -strand regions and six α -helix regions without any predicted UDP-glucose binding site. The latter insertion line has the insertion site in PSPG box and data analysis suggests that any truncated protein expressed in this line would consist of 353 amino acids including nine extended β -strand regions and nine α -helix regions with two predicted UDP-glucose binding sites. However, no complete PSPG box would be expressed. PCR-based genotyping (Figure 78) identified homozygous plants for *ugt71c3-1* and with RT-PCR experiments confirming loss of transcription. Both insertion lines were subsequently used to determine the biological functions of the gene.

PCR screening for *ugt71c3-2* plants showed both genomic and T-DNA bands in all samples, suggesting that all samples were heterozygous for the insertion. In order to confirm the PCR result, genomic and T-DNA fragments were sequenced. The sequencing of the T-DNA fragment showed the right position of the insertion for *UGT71C3*. The alignment of the genomic fragment sequence against *UGT71C3* showed that the PCR product had only a 72% sequence identity to *UGT71C3*. Therefore, the sequence was blasted against the *Arabidopsis* genome and it became obvious that the fragment was amplified non-specifically from a different family member, *At2g29740/UGT71C2*. These results suggest that the plants could be homozygous.

6.7.3 Analysis of *UGT71C3* expression using RT-PCR

The RT-PCR results confirmed that the disruption of transcription for *UGT71C3* was successful as no fragment can be amplified in both KO lines. In accordance with previous affymetrix expression profiles, it was shown that the

expression of *UGT71C3* in Col-0 is up regulated in response to isoxaben treatment after 24 hours.

6.7.4 Phenotypic characterisation of the mutants

Phenotypic characterisation of the isolated knockout lines has been performed in order to investigate the biological functions of the candidate gene.

Due to the common of functional redundancy of family members that could be a problem for the phenotypic study, microarray expression profiles from seedlings treated with isoxaben were cross-referenced for close family members of *UGT71C3*.

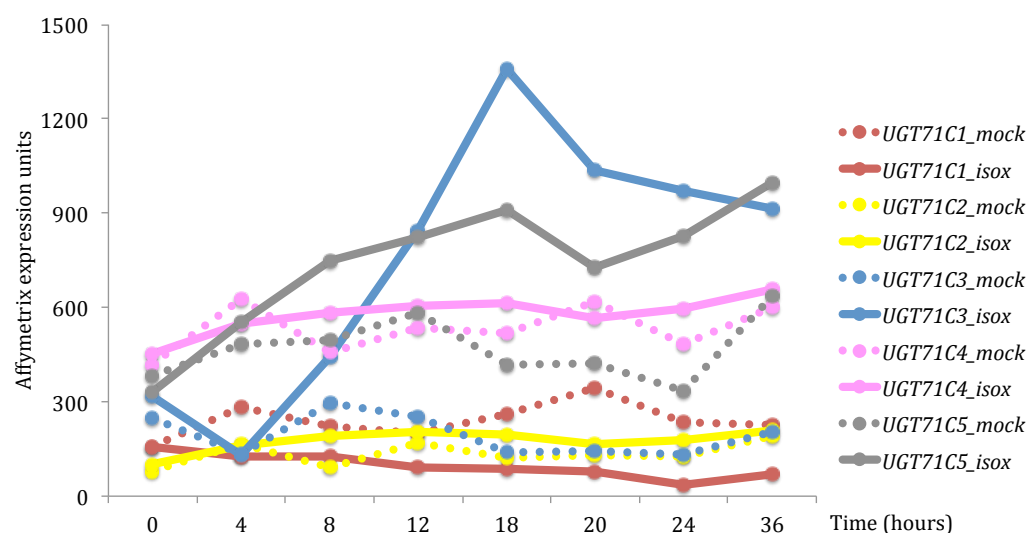


Figure 92 Microarray derived expression data for *UGT71C1* (*At2g29750*) (red), *UGT71C2* (*At2g29740*) (yellow), *UGT71C3* (*At1g07260*) (blue), *UGT71C4* (*At1g07250*) (pink) and *UGT71C5* (*At1g07240*) (grey) under mock treatment (dash line) and isoxaben treatment (bold line). Data derived from Hamann et al. (2009).

As the data shows above (Figure 92), although *UGT71C1*, *UGT71C2* and *UGT71C4* showed high amino acid sequences identities to *UGT71C3* (64 %, 67 % and 62 % identities, respectively) (BlastP analysis; <http://www.arabidopsis.org/cgi-bin/wublast/wublast>), in seedlings treated with isoxaben it was shown that these family members did not respond to isoxaben treatment at all suggesting no functional redundancy at least for isoxaben treated seedlings. However, for *UGT71C5*, which shows 53 % amino acid identities to *UGT71C3*, expression of the gene was up-regulated 0.5-fold

upon isoxaben treatment. This suggests possible functional redundancy could occur.

In addition to the microarray expression profiles from seedlings treated with isoxaben, the family members in UGT71 clad C were cross-referenced on Genevestigator database as shown in Figure 93. From the expression patterns showing in Figure 93, it was shown that all the genes are expressed in hypocotyl and radicle of seedlings and in roots. The expression patterns of *UGT71C1* and *UGT71C2* are very similar to *UGT71C3*, while the patterns of *UGT71C4* and *UGT71C5* are more similar to each other due to their expression observed in shoot, stem and other aerial parts of the plants. The patterns suggest that functional redundancy may occur especially in germinated seeds and roots as the entire clad are expressed across those organs.

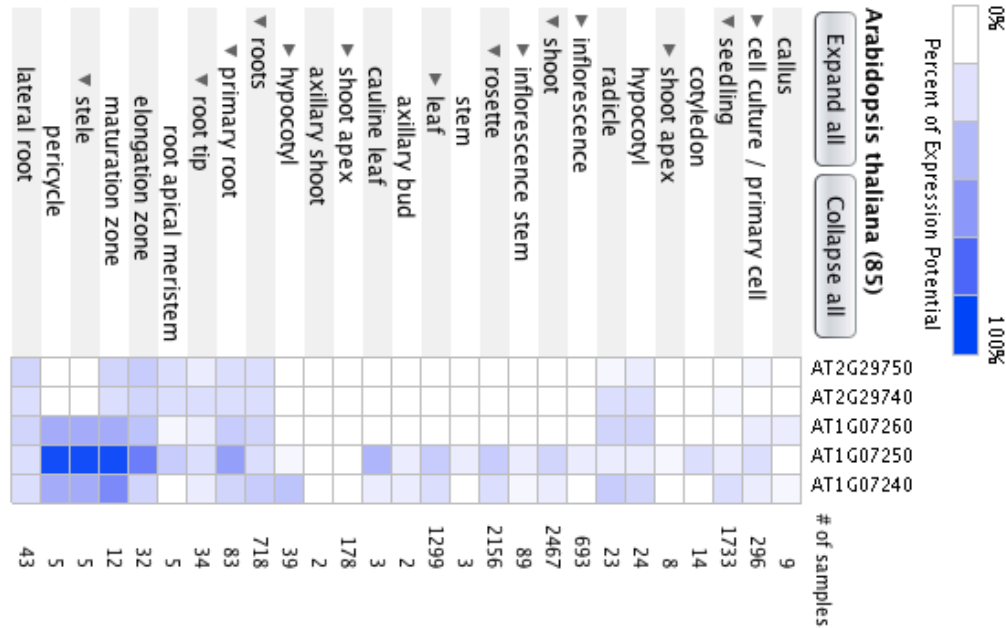


Figure 93 Heat map showing expression patterns of genes belonging to *UGT71* clad C; *At2g29750* (*UGT71C1*), *At2g29740* (*UGT71C2*), *At1g07260* (*UGT71C3*), *At1g07250* (*UGT71C4*) and *At1g07240* (*UGT71C5*) in different plant organs. Data derived from Genevestigator.

According to the available microarray-derived expression data, it was suggested that *UGT71C3* is expressed in radicles, imbibed seeds and roots of seedlings. Therefore, germination assay have been performed to observe whether knockout of the target gene would affect germination rate. As shown in Figure 83, both mutants showed early germination compared with wild type. This implicated the target gene in germination processes or the mechanism/signalling pathway to control the germination such as ABA

signalling. No significant difference was observed for seedling root growth compared to the control, which could be the result of functional redundancy as it was stated before that all members of this clade show expression in roots (Figure 93). At the young rosette stage (four week-old plants), biomass quantification of mutant and wild type showed that mutant plants have lower amounts of biomass (fresh and dry weight), whereas no significant differences were observed in mature plants (eight week-old). This highlights the importance of the target gene at the early growth stages.

Root length measurement and isoxaben hypersensitivity assays showed no significant reproducible differences between *ugt71c3* and wild type seedlings. Previous study on functions of a close family member *UGT71C1* suggested the role of *UGT71C1* in flavonol glycosylation that involved in oxidative stress response as *ugt71c3* mutant was resistance to oxidative stress induced by application of methyl viologen (Lim et al., 2008). Together with previous reported link between oxidative stress and lignin where lignin deposition was found to be induced by oxidative stress during cell wall damage (Denness et al., 2011), hence, lignin staining of mutant seedling roots treated with 600 nM isoxaben were observed. It appeared that interruption of *UGT71C3* expression in the mutants enhanced lignin deposition compared to wild type. In agreement with that, lignin staining for the mutant seedlings grown on the plates with low concentration of isoxaben also appeared to be more sensitive to isoxaben. It could be that lignin deposition in some cells in the cortex or pericycle was more pronounced after treatment with isoxaben. However, more experiments could be performed to reveal which tissues or cell types that lignin deposition occurred. Phloroglucinol staining results suggest more lignin deposition under isoxaben treatment for the mutant seedling roots. In accordance to the observed staining patterns, quantitative analysis also suggests significant difference between the mutants and wild type. However, the measurement was performed to analyse the stained area from 2D photos that hardly cover all possible stained tissues or area in the roots performed the quantification. Thus better imaging techniques or 3D scanning could provide more accurate and better results. However, proper quantitative techniques such as NMR or Klason lignin should be used to quantify the differences exactly. Crossing the knockout lines (or overexpression lines) of

the gene with marker line for the pericycle and cortex could be performed to determine specifically which cell type is exhibiting the enhanced lignin deposition. As it appeared that loss of function *ugt71c3* mutant showed hypersensitive to isoxaben, it could be perceived glycosylation function of the gene involving in regulation of lignin deposition responded to cell wall damage.

Although no obvious phenotypes or even the biomass changes could be observed in mature plants, cell wall analysis of stems was performed. It was found that *ugt71c3* stems exhibit increases in cellulose and all neutral cell wall sugars but decreases in uronic acids levels. The cell wall analysis phenotypes suggested possible role of *UGT71C3* that could be relevant to cell wall integrity maintenance probably via pectin biosynthesis. Based on the uronic measurement observed, it was found that about 20-25 % of the uronic acid decreases when *UGT71C3* was disrupted. The higher cellulose and neutral cell wall sugars observed are probably the results of cell wall reinforcement mechanisms.

Chapter 7 General discussion and summary

7.1 Phenotypic analysis and lignification of mutant seedlings

The plant cell wall forms together with the cuticula the first defence barrier protecting plants from both biotic and abiotic stresses. The wall is a dynamic and complex organ that does not just provide structural support but is also involved in cell-cell interactions, growth and development (Minic & Jouanin, 2006). Due to all the functions needed, the wall must not only have dynamic structures that allow cell growth but also needs to be able to provide structural support to the cell. While wall architecture has been extensively studied, not much has been clarified about the signalling and metabolic processes required to control and maintain integrity of the wall in response to intracellular, intercellular and external stresses during cell growth and development (Hamann, 2012). The main load bearing components of the walls are cellulose microfibrils, which are cross-linked with other matrix polysaccharides. Also, other wall components are present, such as structural proteins and lignin, which are deposited specifically in some cells where more strength is needed (Cosgrove, 2005; Harholt et al., 2010; Park & Cosgrove, 2012a; Scheller & Ulvskov, 2010; Weng & Chapple, 2010). To increase knowledge on the mode of action of the cell wall integrity maintenance mechanism, genetic studies using either forward or reverse approaches have been employed. It was found that impairment of the walls using chemical treatments that specifically affect CESA complexes, such as isoxaben, could trigger plant specific stress response pathways and similar phenotypes could be observed in CESA deficient mutants (Caño-Delgado et al., 2003; Hamann et al., 2009). The phenotypes include ectopic lignification, JA production as early stage responses and carbohydrate distribution changes at the late stage response (Denness et al., 2011; Wormit et al., 2012). In addition, analysis of cell wall composition suggested that not only cellulose content was decreased but also pectin levels were increased in cellulose deficient mutants such as *rsw1* and *kor* or in wild type plants treated with a cellulose biosynthesis

inhibiting herbicide (dichlobenil, DCB) (McCann et al., 2001). Similar results were observed in a tomato cell culture treated with DCB where the cells developed cell walls that lacked cellulose but were enriched in pectins (Shedletsky, Shmuel, Delmer, & Lamport, 1990). These observations suggest that in response to impairment of one load bearing cell wall component, others able to compensate and to maintain cell wall integrity are strengthened. It is reasonable to assume that many regulatory mechanisms with their associated proteins and effectors are required. In parallel, additional data is required in order to understand the biosynthesis and functions of cell wall polysaccharides involved in the compensatory process during plant growth, development and interaction with the environment.

Table 12 Summarised observed phenotypes from *kor2*, *xth31* and *ugt71c3* mutants

Observed phenotypes	<i>kor2-1</i>	<i>kor2-2</i>	<i>kor2-3</i>	<i>xth31-1</i>	<i>xth31-2</i>	<i>ugt71-1</i>	<i>ugt71-2</i>
Growth stage based analysis	N	N	N	N	N	N	N
Rosette diameter measurement	N	N	N	N	N	N	N
Biomass measurement from 4 weeks	S	S	S	S	N	S	S
Biomass measurement from 8 weeks	N	N	N	N	N	N	N
Germination assay	S _L	S _L	N	S _L	N	S _E	S _E
Root growth measurement	S	S	S	S	N	N	N
Root length: isoxaben plates	S	S	S	S	N	N	N
Lignin deposition: 1 nM isoxaben	S	S	S	S	S	S	S
Lignin deposition: 2 nM isoxaben	S	S	S	S	S	S	S
Lignin deposition: 3 nM isoxaben	N	N	N	N	N	N	N
Lignin staining: 12 h, 600 nM isoxaben	S	S	S	S	S	S	S
Lignin staining: 24 h, 600 nM isoxaben	N	N	N	N	N	N	N
Cellulose measurement	N	N	N	S	S	S	S
Uronic acid measurement	S	S	S	S	N	S	S
Neutral cell wall sugar measurement	S	S	S	S	S	S	S

Note: N = Observed phenotypes showed no significant difference from wild type
S = Observed phenotypes showed significant difference from wild type
S_L = Mutants showed significant late germination compared with wild type
S_E = Mutants showed significant early germination compared with wild type

Microarray based expression profiling data for *Arabidopsis* seedlings treated with isoxaben were available in the host lab at the start of the PhD thesis (Hamann et al., 2009). Candidate genes were selected that showed pronounced changes of expression (either up- or down- regulated) upon isoxaben treatment of seedlings and are strongly expressed in plant tissue/ organs where cell wall

remodelling / formation is occurring. From these genes a subset of genes consisting of the following candidates, which encode putative enzymes involved in cell wall metabolism, were targeted in this study: *At1g65610* (*KOR2*), *At3g44990* (*XTH31*) and *At1g65610* (*UGT71C3*). Here, a functional characterization of the candidate genes was performed and the resulting data for *kor2*, *xth31* and *ugt71c3* mutants are summarized in Table 12. On the following pages, these results will be discussed in the context of our current understanding of plant cell wall biology and cell wall integrity maintenance.

Data retrieved from the Genevestigator gene expression database suggested that all the candidate genes are particularly highly expressed in hypocotyls, radicles and roots of seedlings, suggesting that the candidate genes could be involved in cell wall metabolism during cell elongation of fast growing tissues. Similar expression patterns have been reported for other members (*KOR3*, *XTH32* and *UGT76B1*) of the glycosyl hydrolase and glycosyl transferase families based on data deposited in the Genevestigator database (Becnel et al., 2006; Mølhøj et al., 2001^a; von Saint Paul et al., 2011).

KOR1, *KOR2* and *KOR3* are only three members of *EGases* transmembrane proteins in *Arabidopsis* (Mølhøj et al., 2001^a; Nicol et al., 1998). While *KOR1* showed ubiquitous expression throughout most tissues, the expression patterns of *KOR2* and *KOR3* are more similar to each other than to *KOR1*. They are more restricted to specific plant tissues which still actively undergo cell growth and expansion such as cotyledons, young leaves and roots (Mølhøj et al., 2001^a; Mølhøj et al., 2001^b). These observations suggested that the regulation at the transcription level of *KOR2* and *KOR3* could be biologically relevant. However, although both *KOR2* and *KOR3* exhibit similar expression patterns in growing plant organs, differences seem to exist based on data deriving from GUS reporter studies. While *KOR2::GUS* expression was detectable in roots, particularly in developing root hairs and leaf trichomes, *KOR3::GUS* showed particularly strong expression in trichome supporting and bundle sheath cells of leaves (Mølhøj et al., 2001^a). While it has been shown that *kor1* mutant plants exhibit a dwarf growth phenotype, cellulose deficiency and alterations of pectin levels, there is no evidence available regarding the phenotypes of plants mutant for the other

two family members (Krupková & Schmülling, 2009; Lane et al., 2001; Mølhøj et al., 2001^a; Nicol et al., 1998; Szyjanowicz et al., 2004).

Regarding *XTH31* and *XTH32*, both genes belong to subgroup 3 together with another five *XTHs* (*XTH27*, *XTH28*, *XTH29*, *XTH30* and *XTH33*) based on their intron-exon structure and phylogenetic relationships with other *Arabidopsis* *XTHs* (Rose et al., 2002). However, only *XTH31* and *XTH32* have been predicted to have XEH activity based on amino acid sequences similarity to well-characterized enzymes (Baumann et al., 2007; Kaewthai et al., 2010). Microarray-derived expression profiling data available in the Genevestigator suggests that both *XTH31* and *XTH32* are pre-dominantly expressed in expanding and growing tissues such as hypocotyl and roots of seedlings (Becnel et al., 2006). In accordance with the microarray-derived expression data, expression analysis using promoter reporter constructs for *XTH31*, *XTH32* and *XTH33* showed GUS activity in seedling roots (Baumann et al., 2007; Becnel et al., 2006; Ndamukong et al., 2009). However, before this study, no data is available regarding cell wall composition of mutant plants for any of these three family members.

UGT76B1 shows gene expression patterns that are similar to *UGT71C3* based on the available microarray-derived expression data (Genevestigator), with results from isoxaben-treated seedlings suggesting that expression of these two genes is up-regulated (Hamann et al., 2009). *UGT76B1* is the member of the UGT76 clade B whose expression levels are most responsive to stress treatments in *Arabidopsis* based on publically available data (von Saint Paul et al., 2011). Recently, *UGT76B1* has been characterized as a component required for salicylate- (SA) and jasmonate- (JA) mediated signalling pathways that glycosylate the isoleucic acid (ILA) aglycon (von Saint Paul et al., 2011). The phenotypes observed for the *UGT76B1* mutant plants suggested that the gene is required to delay senescence and affects the resistance to *Pseudomonas syringae* and *Alternaria brassicola* (von Saint Paul et al., 2011). Although only subtle mutant phenotypes were observed from the knock-out and over expression mutant plants grown under standard growth conditions, root growth phenotypes from seedlings were observed under application of exogenous ILA (von Saint Paul et al., 2011). In addition, it has been reported that *UGT76B1* plays a role in SA-JA crosstalk by reducing SA-dependent plant defences mechanisms

and inducing JA-dependent responses when no pathogen is present (von Saint Paul et al., 2011). It was found that loss of function of *UGT76B1* SA level and SA-response mechanism as T-DNA insertion mutants showed higher SA levels with the SA-dependent plant defence marker genes being up-regulated with alteration of the expression of JA- response marker genes, while the over expression lines showed opposite phenotypes (von Saint Paul et al., 2011).

Ectopic lignification is one of the effects of impairment of the wall. Several *Arabidopsis* cell wall mutants showed ectopic lignification phenotype, for examples, *ectopic lignification 1 (eli1)*, *constitutive expression of VEGETATIVE STORAGE PROTEINS1 [VSP1] (cev1)*, *root swelling1 (rsw1)* and *de-etiolated3 (det3)* (Caño-delgado et al., 2000; Caño-Delgado et al., 2003; Ellis et al., 2002; Rogers & Campbell, 2004). It had been reported previously that *Arabidopsis* seedlings treated with isoxaben and DCB show ectopic lignification (Desprez et al., 2002; Hamann et al., 2009). In this study, it has been shown that the mutant seedlings treated with isoxaben exhibit enhanced lignin deposition suggesting all mutants in this study including *kor2*, *xth31* and *ugt71c3* are more sensitive to isoxaben than wild type. More severe alterations of morphology and cell expansion were observed in mutant roots treated with isoxaben. Upon treatment with 3nM isoxaben, mutant roots exhibited very short (compared to wild type) or even complete lack of cell elongation zones (see Figure 40, Figure 62 and Figure 87). This part of the study suggested that the candidates are involved in mechanisms mediating the responses to cell wall damage and could be a negative regulator of lignin deposition induced by cell wall damage. To examine the effect of isoxaben on root growth, in addition to lignin deposition staining, root length of seedlings treated with different concentration of isoxaben were measured. For *xth31* mutants, inconsistent results were observed. As shown in Figure 60, both alleles seemed to be somewhat more resistant to isoxaben compared to wild type. However, at 3 nM, the knockout alleles seemed to be more sensitive than wild type while the knockdown allele still exhibited resistance. Higher concentrations of isoxaben could be used to examine the affect on *xth31* root growth. In accordance with the qualitative staining results, quantitative analysis also suggests significant difference. However, measuring the stained area from photos of the roots could be incorrect if some tissues or some areas of the roots

are not correctly imaged. Therefore, better imaging analysis techniques and other quantitative methods such as immune labelling, Klason chemical assays or NMR-based analysis should be employed for more accurate quantification of lignin deposition induced by isoxaben treatment (Chabannes et al., 2001; Marita et al., 1999).

7.2 KORRIGAN 2

It has been shown that most of the members of the endo 1,4- β -glucanase gene family encode enzymes that catalyse hydrolysis of non-cellulosic glucans and are involved in cell expansion, dehiscence or abscission (Williamson et al., 2002). These observations implicate KOR2 in cell wall loosening during cell expansion. Employing forward and reverse genetic approaches, different mutant alleles have been isolated for a close family member, *KORRIGAN*, and extensively studied. Among these alleles are *acw1*, *dec*, *irx2-1*, *korrigan*, *rsw2* and *tsd1* (Frank et al., 2002; Krupková & Schmülling, 2009; Lane et al., 2001; Nicol et al., 1998; Sato et al., 2001; Szyjanowicz et al., 2004; Turner & Somerville, 1997). Phenotypic analysis of the *korrigan* mutants showed that hypocotyl elongation was affected and cellulose levels in the primary cell walls were decreased compared to those in wild type (Nicol et al., 1998). In this study, germination rate analysis showed slightly delayed germination for *kor2*. In parallel root growth analysis showed the roots are longer for *kor2* mutant alleles than wild type. For *kor2*, it was also detected that root growth of all the knockout alleles is more sensitive to isoxaben treatment compared to wild type. Data from Genevestigator also suggest that *KOR2* is expressed in seedlings and also expressed at a medium expression level throughout all developmental stages with a slightly higher expression than other stages during young flower development. However, no obvious phenotypes from the flowers could be observed in this study. More detailed analysis of flower development using imaging techniques like electron microscopy might provide additional data regarding possible flower phenotypes (Smyth et al., 1990). Although no obvious growth phenotypes were observed for mature *kor2* plants cell wall composition analysis for mature stems from all alleles found deviations from wild type

controls for uronic acids and neutral cell wall sugar levels. For *kor2* knockouts, detected cellulose contents showed no significantly different to that in wild type. One explanation that no changes in cellulose levels are observed could be that the Updegraff technique employed here (Updegraff, 1969) limits the measurement to only cellulose molecules that are resistance to acid hydrolysis during the preparation of the wall materials and during the Updegraff's procedure itself. Therefore, changes in other soluble glucans that could be the products of KOR2 would not be detected via the Updegraff procedure. In addition to the cellulose measurement, uronic acids and neutral cell wall sugars were investigated here and it was observed that the *kor2* stems exhibited lower uronic acids content (indicative of changes in pectic polysaccharide backbones) and higher levels of certain neutral cell wall sugars (represents hemicellulose and decorated branches of xylogalacturonan or rhamnogalacturonans) compared to those in wild type. These results suggest that loss of KOR2 induces changes possibly adaptations in cell wall composition and structure to reinforce the wall to permit normal functions. A similar possibly the same situation had been reported before in a study of pectin composition of *kor1*. This study had been carried out employing microscopic techniques and immunolabelling for different kinds of pectins and cellobiose (His et al., 2001). Although no alteration was observed for pectin methyl-esterification, the walls of the mutants showed more homogalacturonan epitopes and less rhamnogalacturonan epitopes compared with those presented in wild type (His et al., 2001). It was also found that epidermal cell walls of *kor1* hypocotyls contained less cellobiohydrolase epitopes indicating a lower level of cellulose crystallinity (His et al., 2001). Similarly, a study of the aspen homolog (*PttCel9A1*) of *KOR1* has revealed that overexpression of *PttCel9A1* leads to high trifluoroacetic acid hydrolysable glucan content and lower crystallinity levels of cellulose (measured by ^{13}C NMR) whereas the Updegraff cellulose contents were similar in the overexpression and control lines (Takahashi et al., 2009). These results provided further support that a feedback mechanism exists that upon impairment of the biosynthesis of a particular plant cell wall component production or structure of another is modified to maintain integrity of the wall. The results in this study indicated that knockout of *KOR2* gene does not affect Updegraff cellulose content but causes

changes in other matrix polysaccharides as shown by lower pectin levels and increased hemicellulose content. Although the observed phenotypes of *kor1* and *kor2* are different, it appeared that loss of function mutants of both genes showed similar effects on pectin composition of the walls. Therefore, *KOR2* and *KOR1* could possibly be membrane anchored EGases which each having distinctive functions to balance biosynthesis of cellulose/hemicellulose network and that functions correlate to pectin metabolism. However, more experimental evidence is needed like examining in detail the structure and composition of pectic polysaccharides in *kor2* seedlings and *KOR2* subcellular localisation needs to be determined. In addition, the actual substrates for either *KOR1* or *KOR2* have not been characterised yet which would be a key to prove whether *KOR2* is involved in cellulose-hemicellulose network biosynthesis and possibly pin point the pathway that specifically requires *KOR2* activity.

7.3 Xyloglucan tranglycosylase/ hydrolase 31

Arabidopsis XTH31 is like *XTH32* a member of the group 3-A *XTH* genes (Baumann et al., 2007). Recently the activities of both *XTH31* and *XTH32* were characterized suggesting they are capable of hydrolysing xyloglucan (Kaewthai et al., 2013; Zhu et al., 2012). The results from a germination rate analysis performed in this study showed slightly delayed germination for *xth31* seeds. On the other hand, an analysis of genes involved in *Arabidopsis* germination reported *xth31* mutants as showing slightly early germination as percentage of germinated seeds since the germination rates of mutants were about three fold higher than that of wild type controls on day two after imbibition (Endo et al., 2012). However, after day 2, the percentage of germination from the mutants were similar to that shown by the wild type control (Endo et al., 2012). Notably, the methods that were used to observe germination rates differ between this study and the one carried out by Endo et al. (2012). The study from Endo et al. (2012) used seeds that had been kept at room temperature for two weeks to grow under standard growth conditions without vernalisation step, while, our study used seeds that had been harvested and kept at 4°C for more than three months. In addition two days of vernalisation at 4°C had also been performed

before germinating the seeds at standard conditions. Therefore the different results observed in both studies could be caused by differences in age, storage methods of the seeds and time differences regarding seed vernalisation.

Different algorithms predict the first nineteen amino acids at the N-terminus of XTH31 act as a signal peptide targeting the protein to the plasma membrane (Figure 19). Cell biological studies using a XTH31-GFP fusion protein have provided support for this prediction (Zhu et al., 2012). To investigate spatial and temporal expression patterns of *AtXTH genes*, Becnel et al. (2006) generated *XTH::GUS* transgenic plants for each *AtXTH* gene and showed that the expression of *XTH31* is specific to roots. The root growth analysis performed in this study found that *xth31* seedling roots are longer than wild type controls. Previously it has been reported that no significant root growth phenotypes were observed in *xth31*, *xth32* or the *xth31*, *xth32* seedlings (root length determined every 24 hours for four days) (Kaewthai et al., 2013). In addition opposite results had been reported previously for *xth31* seedlings showing shorter roots than wild type (Zhu et al., 2012). The study from Zhu et al. (2012) used media supplemented with 0.5 mM CaCl₂ with the aim of the study being to understand the function of AlCl₃. Although our study employed similar methods as the method used by Kaewthai et al. (2013), the developmental stages of seedlings analysed seem to be different. The different results observed are probably caused by the different methods employed for plant growth and to perform root growth analysis. These results highlight the importance of clear methods description to allow comparability of results and to avoid any confusion. However, taken together germination analysis and root assay results implicate *XTH31* in cell wall metabolism of *Arabidopsis* seedlings during early stages of germination and roots of seven day-old *Arabidopsis* seedlings.

Although no obvious growth phenotypes were observed for mature *xth31* mutant plants in this study, cell wall composition analysis showed some interesting differences from wild type controls with respect to cellulose, uronic acids and neutral cell wall sugar levels. For *xth31* mutant stems, both alleles showed elevated cellulose levels. The knockouts for *xth31-1* also showed lower amounts of uronic acids with higher levels of neutral cell wall sugars compared to wild type, suggesting that knockout of the candidate gene promoting cellulose

synthesis. One possible explanation could be that the knockout of the gene is resulting in more precursors being available for cellulose biosynthesis. However, this hypothesis needs more in-depth detailed investigation. Another possibility could be based on the reported XEH enzyme activity of the candidate (Kaewthai, 2011; Kaewthai et al., 2013; Zhu et al., 2012). These reports suggest a role of the enzyme in the turnover process of xyloglucan. Therefore, knocking out of the gene could interrupt turnover of xylose into xyloglucan. However, no changes in xylose content in the walls were detected even though all other matrix components (pectins and other neutral cell wall sugars) were altered. An alternative hypothesis regards the rhamnogalacturonan I. It was found that at least 50 % of the xyloglucan in *Arabidopsis* walls is linked to RG-I (Popper & Fry, 2008). The disruption of xyloglucan turnover process during plant development by knocking out *XTH31* could change the structure or amount of RG-I which would fit with the observation reported here that uronic acids levels are reduced.

7.4 UDP- glycosyl tranferase 71C3

Based on expression data available in the Genevestigator database, *UGT71C3* and other family members in clade C showed preferential expression in hypocotyl, radicle and roots of seedlings (see Figure 93). It has been reported that *UGT78D1* encoding a flavonol 3-*O*-rhamnosyltransferase is involved in flavonol biosynthesis in the endosperm and embryo, but no phenotypic data related to germination phenotypes for the mutants for *UGT78D1* has been reported (Endo et al., 2012). *UGT71C3* has been described as a quercetin 3-*O*-glycosyltransferase since it recognised quercetin as a substrate and glycosylate as the aglycon while showing regioselectivity over 3-OH position of quercetin (E.-K. Lim et al., 2003, 2004). As antioxidant secondary metabolites, quercetin derivatives have been found to accumulate in *Arabidopsis* seeds which the major one of those apparently being quercetin-3-*O*-glycoside (Routaboul et al., 2006). However, although flavonol accumulation has been clearly detected in *Arabidopsis* seeds during germination, the biological functions of this process are still unknown (Endo et al., 2012; Routaboul et al., 2006). A study of another family member *UGT76B1*, which showed very similar expression patterns, is involved in SA-JA

cross talk (von Saint Paul et al., 2011), it had been studied for at the beginning of this project, however, neither in the knock-out mutants nor over-expression lines significant germination phenotypes were observed. Despite the fact that UGTs are very likely to have roles in seed germination probably due to their involvement in flavonol biosynthesis and/or accumulation, for none of known *ugt* mutants germination phenotypes have been described (Endo et al., 2012). Mutants showing germination phenotypes are primarily affecting genes encoding cell wall modifying enzymes that are involved in the wall weakening/loosening during, for example, *man5*, *man6* and *man7* mutants showed late germination, while, *xth31* mutants showed early germination compared with wild type plants (Endo et al., 2012; Iglesias-Fernández et al., 2011). In this study, germination rate analysis showed a slightly early germination phenotype for *ugt71c3* seeds suggesting that *UGT71C3* could probably has flavonol glycosylation function which is required for seed germination since the expression from *Arabidopsis* eFP browser showed the gene expression is dominant in mature seeds and less expression showed after 24 hours imbibition (Winter et al., 2007).

No obvious phenotypes were observed for *ugt71c3* knockout plants, this could be explained by functional redundancy caused by other family members, since the UGT gene family is large and there are five members in UGT clade C, which exhibit a amino acid sequence similarity higher than 50 %. As discussed in Chapter 6, there is a high probability that redundancy could occur, especially for *ugt71c5*, which shows expression patterns very similar to *ugt71c3* in both seedlings treated with isoxaben and throughout growth and development based on expression data deposited in the Genevestigator database (see Figure 92 and Figure 93). Again no obvious growth phenotypes have been observed in mature mutant plants grown under standard growth conditions. Cell wall analysis showed that uronic acids levels were reduced and neutral cell wall sugars were elevated, while cellulose content values were consistent comparing mutants with wild type. These cell wall analysis results indicated that although knockout of the candidate gene did not cause any striking growth phenotypes, the *UGT71C3* KOs affected wall composition and possibly the wall structure. These results are similar to those observed for *xth31* mutants, indicating that knock-outs for these

genes affect the cellulose-hemicellulose network of the wall. To investigate that in more detail, imaging techniques to study the structure of the wall in more detail and/ or other methods to test the strength could be employed. One example for such a study of cellulose microfibril structure is a study of overexpression of *sucrose synthase* (*SuSy*) genes from *Gossypium hirsutum* using hybrid poplar (Coleman et al., 2006). The study on crystallinity and microfibril angle in the walls together with microscopic techniques involving histochemical staining of cellulose and lignin, it was shown that overexpression of *SuSy* leads to increases in secondary cell wall cellulose content and crystallinity of the walls while no significant changes were observed for both secondary wall cellulose microfibril angle and morphology of the transgenic plants (Coleman et al., 2010; Coleman et al., 2006). Another methodology for the study of cell wall structure employs two- and three- dimensional (2D and 3D) magic angle spinning (MAS) solid-state nuclear magnetic resonance (NMR) spectroscopy on intact and insoluble primary cell wall to reveal structure of the walls and dynamic interactions between cellulose, hemicelluloses and pectins (Dick-Pérez et al., 2011). These methods enabled scientists to gain a better understanding of cellulose-pectin networks and that xyloglucan formed the network with cellulose microfibrils by having only limited entrapped parts in the microfibrils thus forming an intermediate substance between cellulose and pectins in the wall networks (Dick-Pérez et al., 2011). In addition, by combining in-depth structural studies of specific types of pectins/ hemicelluloses with phenotypic studies clearer insights into the biological functions of the candidates and certain cell wall components will be generated. These specific structures can be detected using antibodies with over hundred monoclonal antibodies specific for distinct epitopes being available including pectins, xyloglucans, glucans, mannans, xylans and other molecules linked to proteins and other wall polysaccharides (Marcus et al., 2008; McCartney et al., 2005; Pattathil et al., 2010). High sensitivity together with high specificity between the individual antibodies and the epitopes allow accurate and specific detection of particular cell wall polysaccharides. For example, two different antibodies (LM10 and LM11) both show activities against xylans, however the former one exhibits specificity to un-substituted or low-substituted xylans while the latter binds to arabinoxylan in addition to un-

substituted xylans (McCartney et al., 2005). The use of antibodies not only permits detection/quantification of specific wall polysaccharides but can also be used to determine their localisation for better understanding of how these wall components associate with each other (Marcus et al., 2008). Applying these techniques to investigate wall composition and structure in more detail in the candidate genes KOs will provide insights into functions and roles of the candidate genes in plant cell wall metabolism.

7.5 Perspectives and possible contributions to facilitated lignocellulosic biofuel production

Despite the fact that none of the mutants for the three candidate genes showed obvious phenotypes from the plants grown under standard growth conditions, it was found that all knock out mutants exhibited at least 10 % less biomass compared with wild type for the four week-old plants, suggesting possible roles of these genes in plant growth at early stages of development. Interestingly, even though less biomass phenotypes were observed at the early stage of development, no significant differences for total biomass of the senescence plants were detectable, implying that these candidate genes appeared to be function. Although the functions at early growth stages were suggested, they did not appeared to be crucial and could be retrieved probably by other family members at later developmental stages. It would be interesting to determine whether the overexpression of these genes would enhance growth of seedlings or alter the biomass. In mature plants, cell wall analysis from stems of mature plants showed that cellulose content values were significantly higher in *xth31* and *ugt71c3* mutants than in wild type. The mutants for all three genes showed significantly lower levels of pectic polysaccharides (based on uronic acid measurement results) with higher hemicellulose content (based on neutral cell wall sugar measurement results). These results clearly showed that the candidate genes were involved in cell wall metabolism. However, the precise functions remain to be determined. Furthermore, saccharification assays performed by Dr. Szostkiewicz using stem material from mature mutant plants

revealed that the KOs reduce the saccharification performance of the plant material suggesting that overexpression of the candidate genes may improve saccharification efficiency (Szostkiewicz, per. communication). Moreover, heterologous expression and pull down assays of the purified target proteins still need to be achieved for KOR2 and UGT71C3 to examine enzymatic activities, biochemical properties and actual substrates.

Due to the fact that genes encoding cell wall modifying enzymes are very conserved across plant species, characterisation of these genes in the model plant like *Arabidopsis* will provide information that can be applied to manipulate other crops for lignocellulosic biomass, for example, to genetically modified perennial crops such as switchgrass or woody plants such as poplars and willows (Carroll & Somerville, 2009; Simmons et al., 2008; Sticklen, 2008; Vanholme et al., 2010). In switchgrass (*Panicum virgatum*), it was found that down-regulation expression of the caffeic acid 3-O- methyltransferase (*COMT*) gene altered lignin composition, providing the feedstock that required three to four time-less enzymes for ethanol fermentation and increased about 25 % to 38 % ethanol yield (Fu et al., 2011). The transgenic switch grass plants exhibited normal growth and development with no obvious phenotypes while cell wall chemical analysis showed less lignin for about 6 % to 12 % less than wild type control, with minor effect on cellulose content that varied from -1 % to +3 % of the values from wild type (Fu et al., 2011). Study of a glycosyltransferase *PoGT47C* gene in hybrid poplar (*Populus alba* × *tremula*) showed that down-regulation of the gene caused reduction of glucuronoxytan, providing better digestible biomass feedstock for biofuel (Lee et al., 2009). Although inhibition of *PoGT47C* could clearly overcome cell wall recalcitrance, it affected plant fitness as it caused reduction of secondary cell wall thickening and collapsed vessels (Lee et al., 2009). In addition, these candidates can be targeted as biological markers for breeding and selection of other lignocellulosic crops. Very recent studied on secondary cell wall development in willow (*Salix purpurea*) using immune labelling and RNA-sequencing showed co-expression of candidate secondary cell wall genes homologue to *Arabidopsis* *IRX9*, *IRX10* and *GUX1* over six-week study, with similar secondary cell wall development to poplar,

suggesting strong potential for using available genetic tools on willows to improve biomass feedstock for lignocellulosic biofuels (Wan et al., 2014).

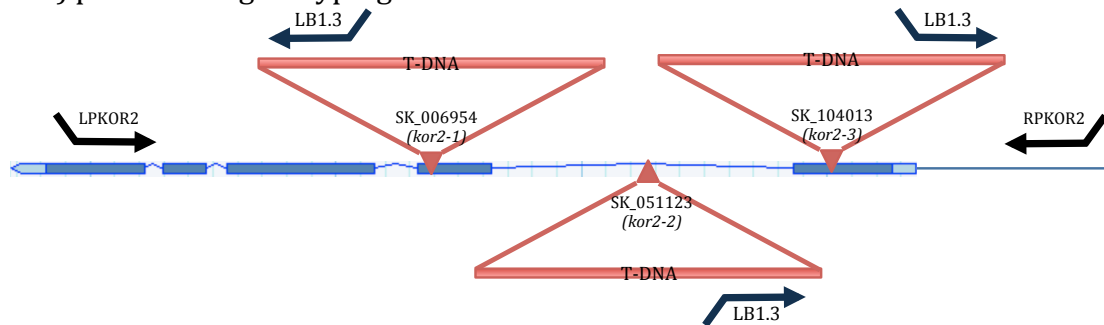
The data gain here could be also employed together through other genetic manipulation approaches including (1) the attempts to enhance cellulose productions either by directly increase cellulose content or modify other pathways involved in the matrix wall components and (2) the attempts to overcome the wall degradation recalcitrant for example, reducing lignin and xylose to provide more digestible biomass feedstock. However, the transgenic crops with alteration of target gene(s) must not drastically affect plant fitness, growth and development. As the mutants presented here exhibited significantly enhanced cellulose and/or hemicellulose levels without any obviously morphological changes or causing any obvious growth phenotypes, they represent attractive leads for facilitation of biofuel production from future bioenergy crops. In addition, even though it does not appear to be the case that the target enzymes here are primary modules in cellulose and/or matrix components and lignin degradations, developing for the successful heterologous expression of these enzymes could provide an efficient tools for specific degradation to study plant cell wall and probably could be employed together with other available enzymes to adjust degradation and modification of the walls. The challenges for other future works in developing lignocellulosic biofuel productions are involving in lower cost and time of the production, specially the pre-treatment and hydrolysis of the biomass where gene manipulation for the lignocellulosic crops and the use of cell wall degrading enzymes could help improving the process. However, to achieve the ultimate goal, more studies need to be performed, not only to improve the biomass and enzyme production but also to develop microorganisms and optimise better fermentation conditions for lignocellulosic biofuel production.

Appendices

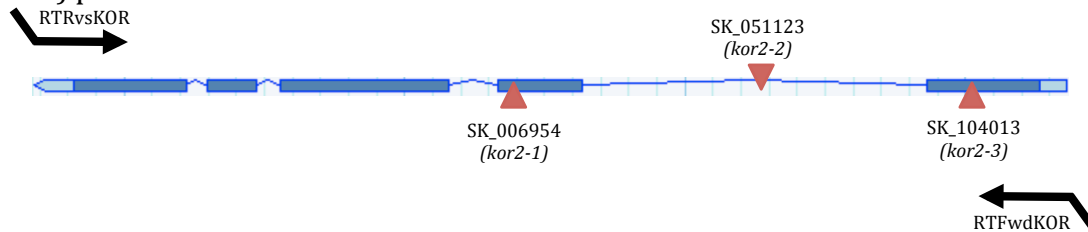
Appendix 1 Primers: annealing sites and sequences

In this research, different primers were designed and used for many aspects including genotyping of the T-DNA insertion mutants and cloning of the target proteins for heterologous expression. Some examples are shown below via schematic depicts represents gene models or protein coding regions of the candidates with annealing sites where oligo primers were designs (shown as arrows).

A-1) primers for genotyping of *kor2* mutants



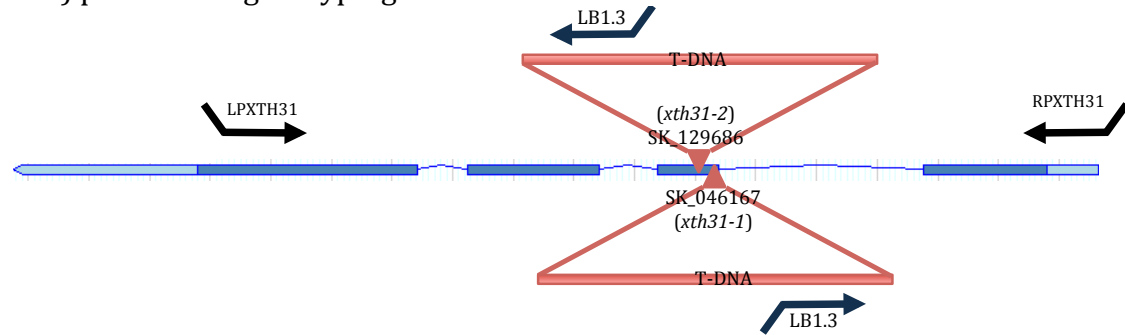
A-2) primers for RT-PCR of *kor2* mutants



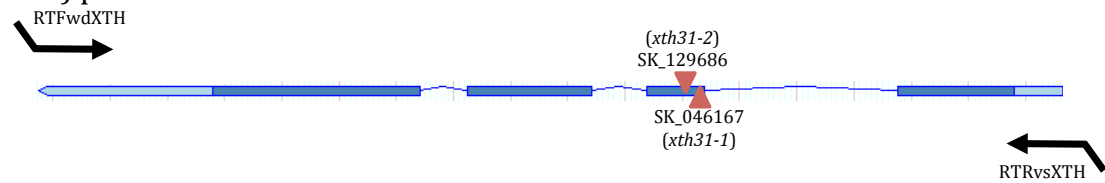
A-3) primers for amplification of KOR2 molecular cloning for heterologous expression



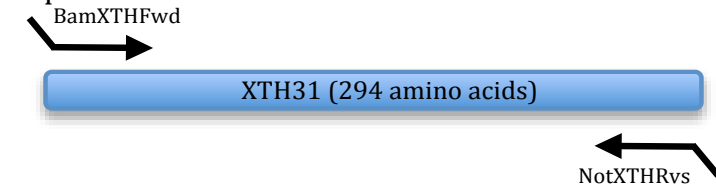
B-1) primers for genotyping of *xth31* mutants



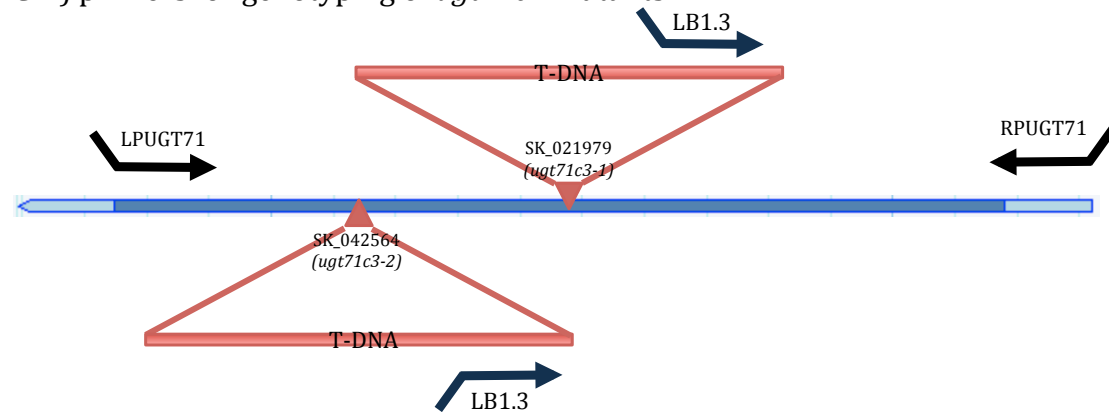
B-2) primers for RT-PCR of *xth31* mutants



B-3) primers for amplification of XTH31 molecular cloning for heterologous expression



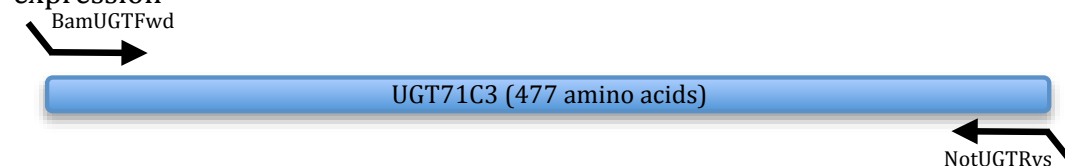
C-1) primers for genotyping of *ugt71c3* mutants



C-2) primers for RT-PCR of *ugt71c3* mutants



C-3) primers for amplification of UGT71C3 molecular cloning for heterologous expression



The rest of the primers used throughout this study are listed in Table 13.

Table 13 Other primers used throughout this study.

Primers	Oligo sequences
07260F3	CCCTGAGAGACATCGCATAACCAC
07260F5	CGTACAAATCCGACGGAGAAAAGCG
3'ColUGT71C3	TACGGGCTCGCTTTCTCCGTC
3'KOR2	GCTCTAGAGGCTTCCAGGCTTTAGGAGGT
3'UBQ10	CCACGGAGCCTGAGGACCAAGTGG
3'UGT71C3	GCTCTAGAACATCTCCGCCGATCAACTCG
3'XTH31	GCTCTAGAACACATTCTGGTGTTTGGGTA
5'ColUGT71C3	CGAGCCGCGGATGAAAGCAGA
5'KOR2	GGACCGCGGAATGCATCCCGGAATGTAT
5'UBQ10	AACCTTCTCTCAATTCTCTCAACC
5'UGT71C3	GGACCGCGGCATGAAAGCAGAAGCAGAGA
5'XTH31	GACTGCAGGGATGGCTTTGTCTCTTATCT
71stpNotIR	ATAAGAATGCGGCCGCTAAACATCTCCGCCG
76stpNotIR	ATAAGAATGCGGCCGCTTAGAAAGACAATATA
AOX1F	GACTGGTTCCAATTGACAAGC
AOX1R	GCAAAATGGCATTCTGACATCC
At1g07260FRT	GAAGACGATGAAAGCAGAAGCAGAG
At1g07260P2r_490	GTCTCTCAGGGAGATACTTCATC
At1g07260P3f_480	CCCTGAGAGACATCGCATAAACAC
At1g07260P4r_966	CGTCGGATTGTACGTATTGACC
At1g07260P5f_952	CGTACAAATCCGACGGAGAAAAGCG
At1g07260RRT	AACACGCACGAACGACGACGT
At3g44990P2r_432	GACATTTTCGGTCGCCACTTCCCC
At3g44990P3f_416	GTGGCGACCGAAATGTCATTGG
BamHI71F	CGGGATCCGATGAAAGCAGAAGCAGAGATC
BamHI76F	CGGGATCCGATGGAGACTAGAGAAACAAAA
ecoExtKOR+aF:	TCCGAATTCAATGGTCAAGTCGTTGCCTC
ExtKORF	CCGGAATTCATGGTCAAGTCGTTGCCTCG
ExtKORR	AGGCCGCGGCAGGCTTCCAGGCTTAGGA
GT71NSF	GGAATTCATGAAAGCAGAAGCAGAGATCA
GT71NSR	GCCGCGGCTAACATCTCCGCCGATCAACT
KORins2	CGGGGGATACTACGACGGAGG
KORins3	CGACGAGTTCATGTGGGCTGG
KORins4	GAGCTATGTGGTTGGATTCCGGC
pET28HisR	TCAGTGGTGGTGGTGGTGGTgc
pstXTHF	AACTGCAGGAATGCAACGTTCTCCTTCACC
pXTHxF	GACTGCAGGGATGGCTTTGTCTCTTATCT
pXTHxR	GCTCTAGAACACATTCTGGTGTTTGGGTA
RTKORF	CCTAGGTTGTGTTTCCGTG
RTKORR	CAAGGCTTCCAGGCTTTAG
Sfi71F	TGGCCCAGCCGCCCATGAAAGCAGAAGCA
sKORxF	GGACCGCGGAATGCATCCCGGAATGTAT
sKORxR	GCTCTAGAGGCTTCCAGGCTTTAGGAGGT
UNI51_P1f_-31&start	GGCCAGAAGGAGATATAACCATG
UNI51_P2r_stop&+31	CTAGTCGACGGCCCATGAGTTA
XTHXbalR	GCTCTAGAACATTCTGGTGTTTGGG
XTHxbaR	GCTCTAGAAAACATTCTGGTGTTTGGGTAT

Appendix 2 Results from Prosite analysis for the candidates

Table 14 Predicted protein domains, families and functional sites for the candidates from Prosite database (adapted from Sigrist et al., 2010).

Protein query	Predicted pattern	Predicted site
KOR2	N-glycosylation site	29 (NTTE)
		116 (NYTL)
		221 (NNSA)
		328 (NSTS)
		349 (NKTY)
		412 (NATG)
		429 (NRTS)
		464 (NSTG)
		548 (NITG)
		571 (NASE)
	cAMP and cGMP-dependent protein kinase phosphorylation site	314 (KRYYS)
		528 (KRRS)
	Protein kinase C phosphorylation site	16 (SDR)
		40 (SQR)
		133 (SGK)
		143 (SWR)
		193 (SHK)
		287 (TDK)
		311 (SRR)
		364 (TAK)
		531 (SCR)
	Casein kinase II phosphorylation site	31 (TEWD)
		43 (SELD)
		188 (SLIE)
		225 (TRLD)
		243 (SPDD)
		267 (SAAD)
		331 (SMFD)
		476 (TFVE)
		489 (SQID)
	Tyrosine kinase phosphorylation site	531 (SCRE)
		252 (KPEDMSY)
		297 (KKGAETLY)
	N-myristoylation site	530 (RSCREGLKY)
		4 (GNVWGG)
		36 (GAIHSQ)
		149 (GTKDGL)
		159 (GGLVGG)
		237 (GLRDSE)
		272 (GAEVSA)
		319 (GQPTAQ)
		415 (GITMCA)
		551 (GAMVGG)
		582 (GLVAAL)
	Amidation site	508 (FGKK)
	Cell attachment sequence	145 (RGD)
	Glycosyl hydrolase family 9 active sites signature 1	503 (SYVVGFGKKFPRRVHHR)

	Glycosyl hydrolase family 9 active sites signature 2	563 (FHDLRNNYNASEPTLSGNA)
XTH31		20 (SQR)
		31 (SSR)
	Protein kinase C phosphorylation site	73 (SLR)
		85 (SIK)
		245 (SCR)
	Casein kinase II phosphorylation site	37 (SPFD)
		203 (SIWD)
	Tyrosine kinase phosphorylation site	216 (RIKADYRY)
	N-myristoylation site	95 (GVDTSL)
		121 (GTTPGK)
UGT71C3	cAMP and cGMP-dependent protein kinase phosphorylation site	112 (RKES)
		93 (STK)
		110 (SSR)
		117 (SVR)
	Protein kinase C phosphorylation site	254 (SLK)
		266 (SDR)
		316 (SIR)
		322 (TEK)
		438 (TPR)
		111 (SRKE)
		167 (TTSE)
		226 (TCLE)
	Casein kinase II phosphorylation site	254 (SLKD)
		266 (SDRD)
		326 (SPYD)
		371 (SVLE)
		431 (SLMD)
	N-myristoylation site	219 (GILVNS)
		316 (GGFVSH)
		379 (GVPIAT)
		427 (GAIRSL)
		348
	UDP-glycosyltransferases signature	(WAPQVEVLAHKALGGFVSHCGWNSVLESLWFG VPIATWPMYAEQ)

Appendix 3 Preliminary studies of eight candidate genes

Appendix 3.1 T-DNA insertion mutants for the candidate genes

Table 15 Overview of T-DNA insertion lines for phenotypic study

Genes	Insertion lines	Insert region	References
<i>At3g44990, XTH31</i>	SK_046167	Exon	This study
	SK_129686	Exon	This study
<i>At3g49220, PME34</i>	SK_062058	Exon	L.Denness
	SK_003562	Exon	This study
<i>At3g53190, PLY12</i>	SK_000463	Exon	L.Denness
	SK_132916	UTR 5'	L.Denness
<i>At1g04680, PLY1</i>	SK_129959	Exon	P.Madhou
	SK_147295	Exon	This study
<i>At1g64390, GUN6</i>	SK_063722	Exon	P.Madhou
	SK_041266	UTR 5'	This study
<i>At1g65610, GUN7, KOR2</i>	SK_006954	Exon	This study
	SK_051123	Intron	This study
	SK_104013	Exon	This study
<i>At1g07260, UGT71C3</i>	SK_042564	Exon	This study
	SK_021979	Exon	This study
<i>At3g11340, UGT76B1</i>	SAIL_1171_A11	Exon	This study
	SK_104508	Exon	This study
	SK_016588	Promoter	This study

Note: Homozygous plants for some insertion lines had been isolated before this study started, by Dr. Lucinda Denness and Dr. Priyadharshini Madhou as indicated in the references.

Appendix 3.2 Heterologous expression of target enzymes in *Pichia*

To predict native signal peptides, SignalP 3.0 server was employed since any signal peptides need to be removed to ensure successful expression using the yeast secretion factor of the pPICZ α vectors. Based on predicted signal peptides and trans-membrane domains, the putative regions for secreted soluble protein expression were used to design the cloning strategy. For the cloning of each construct, two restriction enzymes that have no recognition site in the cDNA but in the multi-cloning site of the pPICZ α vector in both version A and B were selected using NEBcutter v.2 program (<http://tools.neb.com/NEBcutter2/>). The ORF of the final constructs will harbour the cassette that posses the yeast α signal sequence at the N-terminus, followed by sequences of the candidate gene and c-myc and hexa-histidine tags at the C-terminus. In addition, Rare Codon Analysis Tools from GenScript (http://www.genscript.com/cgi-bin/tools/rare_codon_analysis) was employed to analyse the expression cassettes based on codon usage in *P. pastoris*. The results including adaptation index (CAI), GC content and low frequency codons (codon usage in *P. pastoris* is lower than 30%) are shown in Table 16.

Table 16 Codon usage analysis using Rare Codon Analysis Tools from GenScript

Constructs	CAI	% GC	% low frequency codons
pPICZ α B_XTH31	0.65	46.74	10
pPICZ α A_PME8	0.64	48.99	11
pPICZ α A_PLL17	0.69	45.6	7
pPICZ α A_PLL26	0.7	47.5	7
pPICZ α A_KOR2	0.7	46.6	8
pPICZ α B_GH9C2	0.7	48	6
pPICZ α A_UGT71C3	0.65	49.29	10
pPICZ α A_UGT76B1	0.69	44.33	7

The CAI for all the queries is in the range from 0.65 to 0.7 with GC contents of 45 to 50 percent. The constructs contain 6 to 10 percent of low frequency codons. While the ideal CAI is 1.0 and the acceptable range for GC content is in between 30 to 70 percent. Data gained from Rare Codon Analysis Tools suggests that all the ORF designed are in an acceptable range and should be suitable to be expressed in *P. pastoris*.

Based on cloning strategy shown in Table 16, coding region without any predicted signal sequences (see Table 8) for the candidate genes was cloned into selected expression vector and consequently transformed into *E. coli* then *P. pastoris*.

Genomic DNA of each candidate was prepared and used as templates for PCR reaction to amplify the AOX1 region. The reactions using AOX1 primers can amplify WT

AOX1 fragment (2.2 kb; Figure 94, B2 and C5) and pPICZ α A (588 bp; Figure 94, A3). Notably, the recombinant strains have Mut⁺ genotype; therefore, the results should show both a 2.2 kb WT fragment and an insert fragment. For pPICZ α A_UGT71C3 (2 kb; Figure 94, A1) and pPICZ α A_UGT76B1 (1.9 kb; Figure 94, A2), sizes of the inserted fragments are similar to WT that cannot be separated on 1% agarose gel. Therefore 2% agarose was used to analyse other constructs as shown in Figure 94 B and C. The PCR for pPICZ α A_PME8 resulted in 2.4 kb for inserted fragment and 2.2 kb for WT fragment (Figure 94, B1). For pPICZ α A_PLL17, the result showed 1.9 kb for inserted fragment with 2.2 kb WT fragment (Figure 94, C1). For pPICZ α A_PLL26, the result showed 1.8 kb inserted fragment and the 2.2 kb WT band (Figure 94, C2). For pPICZ α A_KOR2, the result showed 2.5 kb inserted fragment and 2.2 kb WT band (Figure 94, C3). For pPICZ α B_GH9C2, the PCR resulted in 2.4 kb for inserted fragment and 2.2 WT fragment (Figure 94, C4)

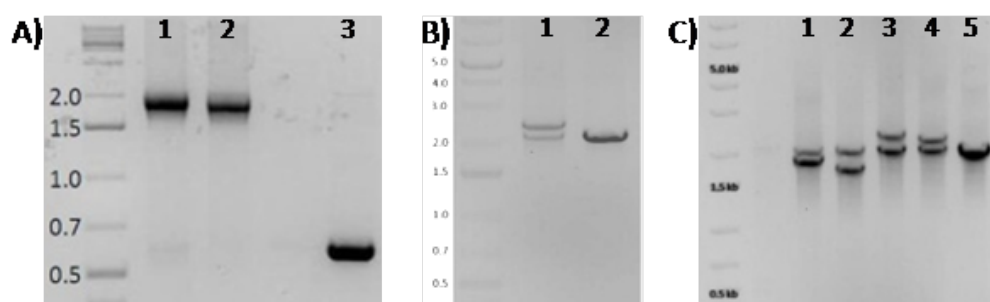


Figure 94 PCR to amplify AOX1 region using genomic DNA templates from *Pichia* clones, pPICZ α A_UGT71C3 (A, 1), pPICZ α A_76B1 (A, 2), Empty plasmid (A, 3), pPICZ α A_PME (B, 1), GS115 harbouring empty plasmid (B, 2), pPICZ α A_PLL17 (C, 1), pPICZ α A_PLL26 (C, 2), pPICZ α A_KOR2 (C, 3), pPICZ α B_GH9C2 (C, 4) and GS115 (C, 5).

At this step, *Pichia* expression strains for seven candidate enzymes had selected and confirmed by PCR using gDNA as templates. Additionally, the amplified DNA fragments were sent for sequencing to confirm the integration of the whole expression cassette including the secretion signal, c-myc and 6XHis tags into *Pichia* genome.

However, the trial expression of the selected clones was not successful as no target protein could be detected from SDS-PAGEs nor Western blots. The optimisation had been done in variations of cell density, media, temperature and time. Still no expression had been detected. Preliminary investigation was performed and confirms the expression of the target genes at transcriptional level as shown in semi-quantitative RT-PCR. By the time limit, no further experiment had been carried out to determine the causes of unsuccessful heterologous of the target enzymes in *Pichia*.

References

- Achnine, L., Huhman, D. V., Farag, M. a., Sumner, L. W., Blount, J. W., & Dixon, R. a. (2005). Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. *The Plant Journal : For Cell and Molecular Biology*, 41(6), 875–87.
- Atanasov, I. I., Pittman, J. K., & Turner, S. R. (2009). Elucidating the mechanisms of assembly and subunit interaction of the cellulose synthase complex of *Arabidopsis* secondary cell walls. *The Journal of Biological Chemistry*, 284(6), 3833–41.
- Bauer, S., Vasu, P., Persson, S., Mort, A. J., & Somerville, C. R. (2006). Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. *Proceedings of the National Academy of Sciences of the United States of America*, 103(30), 11417–22.
- Baumann, M. J., Eklöf, J. M., Michel, G., Kallas, A. M., Teeri, T. T., Czjzek, M., & Brumer, H. (2007). Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo-transglycosylases: biological implications for cell wall metabolism. *The Plant Cell*, 19(6), 1947–63.
- Becnel, J., Natarajan, M., Kipp, A., & Braam, J. (2006). Developmental expression patterns of *Arabidopsis* XTH genes reported by transgenes and Genevestigator. *Plant Molecular Biology*, 61(3), 451–67.
- Blakeney, A. B., Harris, P. J., Henry, R. J., & Stone, B. A. (1983). A simple and rapid preparation of alditol acetates. *Carbohydrate Research*, 113, 291–299.
- Bleecker, a B., & Kende, H. (2000). Ethylene: a gaseous signal molecule in plants. *Annual Review of Cell and Developmental Biology*, 16, 1–18.
- Blumenkrantz, N., & Asboe-Hansen, G. (1973). New method for quantitative determination of uronic acids. *Analytical Biochemistry*, 54(2), 484–489.
- Borriss, R., Buettner, K., & Maentsaelae, P. (1990). Structure of the beta-1,3-1,4-glucanase gene of *Bacillus macerans* : Homologies to other beta-glucanases. *Molecular and General Genetics*, 34, 278–283.
- Boyes, D. C., Zayed, a M., Ascenzi, R., McCaskill, a J., Hoffman, N. E., Davis, K. R., & Görlach, J. (2001). Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants. *The Plant Cell*, 13(7), 1499–510.
- Caffall, K. H., & Mohnen, D. (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate Research*, 344(14), 1879–900.
- Cannon, M. C., Terneus, K., Hall, Q., Tan, L., Wang, Y., Wegenhart, B. L., Chen, L., Lamport, D. T. A., Chen, Y., & Kieliszewski, M. J. (2008). Self-assembly of the plant cell wall requires an extensin scaffold. *Proceedings of the National Academy of Sciences of the United States of America*, 105(6), 2226–31.
- Caño-delgado, A. I., Metzlaß, K., & Bevan, M. W. (2000). The *eli1* mutation reveals a link between cell expansion and secondary cell wall formation in *Arabidopsis thaliana* Screening procedure. *Screening*, 3405, 3395–3405.

- Caño-Delgado, A., Penfield, S., Caño-Delgado, A., Smith, C., Catley, M., & Bevan, M. W. (2003). Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*. *The Plant Journal*, 34, 351–362.
- Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S., & Martens, S. (2012). A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. *The Plant Journal : For Cell and Molecular Biology*, 69(6), 1030–42.
- Carpita, N. C. (1996). Structure and Biogenesis of the Cell Walls of Grasses. *Annual Review of Plant Physiology and Plant Molecular Biology*, 47, 445–476.
- Carpita, N. C. (2011). Update on mechanisms of plant cell wall biosynthesis: how plants make cellulose and other (1->4)- β -D-glycans. *Plant Physiology*, 155(1), 171–84.
- Carroll, A., Mansoori, N., Li, S., Lei, L., Vernhettes, S., Visser, R. G. F., Somerville, C., Gu, Y., & Trindade, L. M. (2012). Complexes with mixed primary and secondary cellulose synthases are functional in *Arabidopsis thaliana* plants. *Plant Physiology*.
- Carroll, A., & Somerville, C. (2009). Cellulosic biofuels. *Annual Review of Plant Biology*, 60, 165–82.
- Chabannes, M., Ruel, K., Yoshinaga, A., Chabbert, B., Jauneau, A., Joseleau, J. P., & Boudet, A. M. (2001). *In situ* analysis of lignins in transgenic tobacco reveals a differential impact of individual transformations on the spatial patterns of lignin deposition at the cellular and subcellular levels. *The Plant Journal*, 28(3), 271–82.
- Ciuffetti, L. M., Tuori, R. P., & Gaventa, J. M. (1997). A single gene encodes a selective toxin causal to the development of tan spot of wheat. *The Plant Cell*, 9(2), 135–44.
- Coleman, H. D., Beamish, L., Reid, A., Park, J.-Y., & Mansfield, S. D. (2010). Altered sucrose metabolism impacts plant biomass production and flower development. *Transgenic Research*, 19(2), 269–83.
- Coleman, H. D., Ellis, D. D., Gilbert, M., & Mansfield, S. D. (2006). Up-regulation of sucrose synthase and UDP-glucose pyrophosphorylase impacts plant growth and metabolism. *Plant Biotechnology Journal*, 4(1), 87–101.
- Coppolino, M. G., & Dedhar, S. (2000). Bi-directional signal transduction by integrin receptors. *The International Journal of Biochemistry & Cell Biology*, 32(2), 171–88.
- Cosgrove, D. J. (2000). Expansive growth of plant cell walls. *Plant Physiology and Biochemistry*, 38(1-2), 109–24.
- Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nature Reviews. Molecular Cell Biology*, 6(11), 850–61.
- Cosgrove, D. J., & Jarvis, M. C. (2012). Comparative structure and biomechanics of plant primary and secondary cell walls. *Frontiers in Plant Science*, 3(August), 1–6.
- D'Souza, S., Ginsberg, M., & Plow, E. (1991). Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *Trends in Biochemical Sciences*, 16, 246–250.
- Denness, L., McKenna, J. F., Segonzac, C., Wormit, A., Madhou, P., Bennett, M., Mansfield, J., Zipfel, C., & Hamann, T. (2011). Cell wall damage-induced lignin biosynthesis is regulated by a reactive oxygen species- and jasmonic acid-dependent process in *Arabidopsis*. *Plant Physiology*, 156(3), 1364–74.

- Desprez, T., Juraniec, M., Crowell, E. F., Jouy, H., Pochylova, Z., Parcy, F., Höfte, H., Gonneau, M., & Vernhettes, S. (2007). Organization of cellulose synthase complexes involved in primary cell wall synthesis in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, 104(39), 15572–7.
- Desprez, T., Vernhettes, S., Fagard, M., Refrégier, G., Desnos, T., Aletti, E., Py, N., Pelletier, S., & Höfte, H. (2002). Resistance against herbicide isoxaben and cellulose deficiency caused by distinct mutations in same cellulose synthase isoform CESA6 1. *Plant Physiology*, 128(February), 482–490.
- Dick-Pérez, M., Zhang, Y., Hayes, J., Salazar, A., Zabortina, O. A., & Hong, M. (2011). Structure and interactions of plant cell-wall polysaccharides by two- and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry*, 50(6), 989–1000.
- Egelund, J., Petersen, B. L., Motawia, M. S., Damager, I., Faik, A., Olsen, C. E., Ishii, T., Clausen, H., Ulvskov, P., & Geshi, N. (2006). Arabidopsis thaliana RGXT1 and RGXT2 encode golgi-localized (1,3)-alpha-D-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II. *The Plant Cell*, 18(10), 2593–607.
- Eklöf, J. M., & Brumer, H. (2010). The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling. *Plant Physiology*, 153(2), 456–66.
- Ellis, C., Karafyllidis, I., Wasternack, C., & Turner, J. G. (2002). The *Arabidopsis* mutant *cev1* links cell wall signaling to jasmonate and ethylene responses. *The Plant Cell*, 14(July), 1557–1566.
- Endo, A., Tatematsu, K., Hanada, K., Duermeyer, L., Okamoto, M., Yonekura-Sakakibara, K., Saito, K., Toyoda, T., Kawakami, N., Kamiya, Y., Seki, M., & Nambara, E. (2012). Tissue-specific transcriptome analysis reveals cell wall metabolism, flavonol biosynthesis and defense responses are activated in the endosperm of germinating *Arabidopsis thaliana* seeds. *Plant and Cell Physiology*, 53(1), 16–27.
- Etzioni, A. (1999). Integrins: the glue of life. *The Lancet*, 353(9150), 341–3.
- Fornalé, S., Capellades, M., Encina, A., Wang, K., Irar, S., Lapierre, C., Ruel, K., Joseleau, J. P., Berenguer, J., Puigdomènech, P., Rigau, J., & Caparrós-Ruiz, D. (2011). Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. *Molecular Plant*, 5(4), 1–14.
- Frank, M., Guivarc'h, A., Krupková, E., Lorenz-Meyer, I., Chriqui, D., & Schmölling, T. (2002). Tumorous shoot development (TSD) genes are required for co-ordinated plant shoot development. *The Plant Journal : For Cell and Molecular Biology*, 29(1), 73–85.
- Fu, C., Mielenz, J. R., Xiao, X., Ge, Y., Hamilton, C. Y., Rodriguez, M., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R. D., & Wang, Z. Y. (2011). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proceedings of the National Academy of Sciences of the United States of America*, 108(9), 3803–8.
- Gachon, C. M. M., Langlois-Meurinne, M., & Saindrenan, P. (2005). Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends in Plant Science*, 10(11), 542–9.
- Geshi, N., Petersen, B. L., & Scheller, H. V. (2010). Toward tailored synthesis of functional polysaccharides in plants. *Annals of the New York Academy of Sciences*, 1190, 50–7.
- Gille, S., Hänsel, U., Ziemann, M., & Pauly, M. (2009). Identification of plant cell wall mutants by means of a forward chemical genetic approach using hydrolases. *Proceedings of the National Academy of Sciences of the United States of America*, 106(34), 14699–704.

- Hamann, T. (2012). Plant cell wall integrity maintenance as an essential component of biotic stress response mechanisms. *Frontiers in Plant Science*, 3(April), 77.
- Hamann, T., Bennett, M., Mansfield, J., & Somerville, C. (2009). Identification of cell-wall stress as a hexose-dependent and osmosensitive regulator of plant responses. *The Plant Journal : For Cell and Molecular Biology*, 57(6), 1015–26.
- Harholt, J., Suttangkakul, A., & Vibe Scheller, H. (2010). Biosynthesis of pectin. *Plant Physiology*, 153(2), 384–95.
- Heinisch, J. J., & Dufrêne, Y. F. (2010). Is there anyone out there: single-molecule atomic force microscopy meets yeast genetics to study sensor functions. *Integrative Biology : Quantitative Biosciences from Nano to Macro*, 2(9), 408–15.
- Hématy, K., Cherk, C., & Somerville, S. (2009). Host-pathogen warfare at the plant cell wall. *Current Opinion in Plant Biology*, 12(4), 406–13.
- Henrissat, B., Coutinho, P. M., & Davies, G. J. (2001). A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. *Plant Molecular Biology*, 47(1-2), 55–72.
- Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., & Foust, T. D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*, 315(5813), 804–7.
- His, I., Driouich, a, Nicol, F., Jauneau, a, & Höfte, H. (2001). Altered pectin composition in primary cell walls of korrigan, a dwarf mutant of *Arabidopsis* deficient in a membrane-bound endo-1,4-beta-glucanase. *Planta*, 212(3), 348–58.
- Holmgren, A., Norgren, M., Zhang, L., & Henriksson, G. (2009). On the role of the monolignol gamma-carbon functionality in lignin biopolymerization. *Phytochemistry*, 70(1), 147–55.
- Huang, C. C. J., Damasceno, L. M. L., Anderson, K. a., Zhang, S., Old, L. J., & Batt, C. A. (2011). A proteomic analysis of the *Pichia pastoris* secretome in methanol-induced cultures. *Applied Microbiology and Biotechnology*, 90(1), 235–47.
- Humphrey, T. V, Bonetta, D. T., & Goring, D. R. (2007). Sentinels at the wall: cell wall receptors and sensors. *The New Phytologist*, 176(1), 7–21.
- Husar, S., Berthiller, F., Fujioka, S., Rozhon, W., Khan, M., Kalaivanan, F., Elias, L., Higgins, G. S., Li, Y., Schuhmacher, R., Krska, R., Seto, H., Vaistij, F. E., Bowles, D., & Poppenberger, B. (2011). Overexpression of the UGT73C6 alters brassinosteroid glucoside formation in *Arabidopsis thaliana*. *BMC Plant Biology*, 11:51.
- Iglesias-Fernández, R., Rodríguez-Gacio, M. C., Barrero-Sicilia, C., Carbonero, P., & Matilla, A. (2011). Three endo-β-mannanase genes expressed in the micropylar endosperm and in the radicle influence germination of *Arabidopsis thaliana* seeds. *Planta*, 233(1), 25–36.
- Jensen, J. K., Sørensen, S. O., Harholt, J., Geshi, N., Sakuragi, Y., Møller, I., Zandleven, J., Bernal, A. J., Jensen, N. B., Sørensen, C., Pauly, M., Beldman, G., Willats, G. T., & Scheller, H. V. (2008). Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in *Arabidopsis*. *The Plant Cell*, 20(5), 1289–302.
- Kaewthai, N. (2011). *Plant and microbial xyloglucanases: Function, Structure and Phylogeny*. PhD thesis, School of Biotechnology, Royal Institute of Technology.

- Kaewthai, N., Gendre, D., Eklöf, J. M., Ibatullin, F. M., Ezcurra, I., Bhalerao, R. P., & Brumer, H. (2013). Group III-A XTH genes of *Arabidopsis* encode predominant xyloglucan endohydrolases that are dispensable for normal growth. *Plant Physiology*, 161(1), 440–54.
- Kaewthai, N., Harvey, A., & Hrmova, M. (2010). Heterologous expression of diverse barley XTH genes in the yeast *Pichia pastoris*. *Plant Biotechnology*, 258, 251–258.
- Katembe, W. J., Swatzell, L. J., Makaroff, C. a, & Kiss, J. Z. (1997). Immunolocalization of integrin-like proteins in *Arabidopsis* and *Chara*. *Physiologia Plantarum*, 99(1), 7–14.
- Knepper, C., Savory, E. A., & Day, B. (2011). *Arabidopsis* NDR1 is an integrin-like protein with a role in fluid loss and plasma membrane-cell wall adhesion. *Plant Physiology*, 156(1), 286–300.
- Krupková, E., & Schmülling, T. (2009). Developmental consequences of the *tumorous shoot development1* mutation, a novel allele of the cellulose-synthesizing *KORRIGAN1* gene. *Plant Molecular Biology*, 71(6), 641–55.
- Laemmli, U. (1970). Cleavage of Structural Proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680–685.
- Lamport, D. T. a, Kieliszewski, M. J., Chen, Y., & Cannon, M. C. (2011). Role of the extensin superfamily in primary cell wall architecture. *Plant Physiology*, 156(1), 11–9.
- Lane, D. R., Wiedemeier, a, Peng, L., Höfte, H., Vernhettes, S., Desprez, T., Hocart, C. H., Birch, R. J., Baskin, T. I., Burn, J. E., Arioli, T., Betzner, A. S., & Williamson, R. E. (2001). Temperature-sensitive alleles of *RSW2* link the *KORRIGAN* endo-1,4-beta-glucanase to cellulose synthesis and cytokinesis in *Arabidopsis*. *Plant Physiology*, 126(1), 278–88.
- Lee, C., Teng, Q., Huang, W., Zhong, R., & Ye, Z.-H. (2009). Down-regulation of *PoGT47C* expression in poplar results in a reduced glucuronoxylan content and an increased wood digestibility by cellulase. *Plant & Cell Physiology*, 50(6), 1075–89.
- Levin, D. E. (2005). Cell Wall Integrity Signaling in *Saccharomyces cerevisiae*. *Society*, 69(2), 262–291.
- Levin, D. E. (2011). Regulation of cell wall biogenesis in *Saccharomyces cerevisiae*: the cell wall integrity signaling pathway. *Genetics*, 189(4), 1145–75.
- Li, X., Weng, J. K., & Chapple, C. (2008). Improvement of biomass through lignin modification. *The Plant Journal*, 54(4), 569–81.
- Li, Y., Baldauf, S., Lim, E. K., & Bowles, D. J. (2001). Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. *The Journal of Biological Chemistry*, 276(6), 4338–43.
- Liepmann, A. H., Wightman, R., Geshi, N., Turner, S. R., & Scheller, H. V. (2010). *Arabidopsis*: a powerful model system for plant cell wall research. *The Plant Journal*, 61(6), 1107–21.
- Liepmann, A. H., Wilkerson, C. G., & Keegstra, K. (2005). Expression of *cellulose synthase-like (Csl)* genes in insect cells reveals that *CslA* family members encode mannan synthases. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), 2221–6.

- Lim, C. E., Choi, J. N., Kim, I. A., Lee, S. A., Hwang, Y., & Lee, C. H. (2008). Improved resistance to oxidative stress by a loss-of-function mutation in the *Arabidopsis* *UGT71C1* gene. *Molecules and Cells*, 25(3), 368–375.
- Lim, E. K., Ashford, D. a, Hou, B., Jackson, R. G., & Bowles, D. J. (2004). *Arabidopsis* glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. *Biotechnology and Bioengineering*, 87(5), 623–31.
- Lim, E. K., Baldauf, S., Li, Y., Elias, L., Worrall, D., Spencer, S. P., Jackson, R. G., Taguchi, G., Ross, J., & Bowles, D. J. (2003). Evolution of substrate recognition across a multigene family of glycosyltransferases in *Arabidopsis*. *Glycobiology*, 13(3), 139–45.
- Lim, E. K., Doucet, C. J., Hou, B., Jackson, R. G., Abrams, S. R., & Bowles, D. J. (2005). Resolution of (+)-abscisic acid using an *Arabidopsis* glycosyltransferase. *Tetrahedron: Asymmetry*, 16(1), 143–147.
- Lopez-Casado, G., Urbanowicz, B. R., Damasceno, C. M. B., & Rose, J. K. C. (2008). Plant glycosyl hydrolases and biofuels: a natural marriage. *Current Opinion in Plant Biology*, 11(3), 329–37.
- Lü, B., Chen, F., Gong, Z. H., Xie, H., Zhang, J. H., & Liang, J. S. (2007). Intracellular localization of integrin-like protein and its roles in osmotic stress-induced abscisic acid biosynthesis in *Zea mays*. *Protoplasma*, 232(1-2), 35–43.
- Lü, B., Chen, F., Gong, Z., Xie, H., & Liang, J. (2007). Integrin-like protein is involved in the osmotic stress- induced abscisic acid biosynthesis in *Arabidopsis thaliana*. *Journal of Integrative Plant Biology*, 49(4), 540–549.
- Marcus, S. E., Verhertbruggen, Y., Hervé, C., Ordaz-Ortiz, J. J., Farkas, V., Pedersen, H. L., Willats, W. G. T., & Knox, J. P. (2008). Pectic homogalacturonan masks abundant sets of xyloglucan epitopes in plant cell walls. *BMC Plant Biology*, 8, 60.
- Marita, J. M., Ralph, J., Hatfield, R. D., & Chapple, C. (1999). NMR characterization of lignins in *Arabidopsis* altered in the activity of ferulate 5-hydroxylase. *Proceedings of the National Academy of Sciences of the United States of America*, 96(22), 12328–32.
- Mark, P., Baumann, M. J., Eklöf, J. M., Gullfot, F., Michel, G., Kallas, A. M., Teeri, T. T., Brumer, H., & Czjzek, M. (2009). Analysis of nasturtium TmNXG1 complexes by crystallography and molecular dynamics provides detailed insight into substrate recognition by family GH16 xyloglucan endo-transglycosylases and endo-hydrolases. *Proteins*, 75(4), 820–36.
- Martínez-Alonso, M., García-Fruitós, E., Ferrer-Miralles, N., Rinas, U., & Villaverde, A. (2010). Side effects of chaperone gene co-expression in recombinant protein production. *Microbial Cell Factories*, 9, 64.
- McCann, M. C., Bush, M., Milioni, D., Sado, P., Stacey, N. J., Catchpole, G., Defernez, M., Carpita, N. C., Höfte, H., Ulvskov, P., Wilson, R. H., & Roberts, K. (2001). Approaches to understanding the functional architecture of the plant cell wall. *Phytochemistry*, 57(6), 811–21.
- McCartney, L., Marcus, S. E., & Knox, J. P. (2005). Monoclonal antibodies to plant cell wall xylans and arabinoxylans. *The Journal of Histochemistry and Cytochemistry*, 53(4), 543–6.
- Meinhardt, S., Cheng, W., Kwon, C. Y., Donohue, C. M., & Rasmussen, J. B. (2002). Role of the Arginyl-Glycyl-Aspartic motif in the action of Ptr toxA produced by *Pyrenophora tritici-repentis*. *Plant Physiology*, 130(3), 1545–1551.

- Miedes, E., & Lorences, E. P. (2009). Xyloglucan endotransglucosylase/hydrolases (XTHs) during tomato fruit growth and ripening. *Journal of Plant Physiology*, 166(5), 489–98.
- Minic, Z. (2008). Physiological roles of plant glycoside hydrolases. *Planta*, 227(4), 723–40.
- Minic, Z., & Jouanin, L. (2006). Plant glycoside hydrolases involved in cell wall polysaccharide degradation. *Plant Physiology and Biochemistry*, 44(7-9), 435–49.
- Mølhøj, M., Jørgensen, B., Ulvskov, P., & Borkhardt, B. (2001). Two *Arabidopsis thaliana* genes, *KOR2* and *KOR3*, which encode membrane-anchored endo-1,4-beta-D-glucanases, are differentially expressed in developing leaf trichomes and their support cells. *Plant Molecular Biology*, 46(3), 263–75.
- Mølhøj, M., Ulvskov, P., & Degan, F. D. (2001). Characterization of a functional soluble form of a heterologously expressed in *Pichia pastoris*. *Society*, 127(2), 674–684.
- Naseer, S., Lee, Y., Lapierre, C., Franke, R., Nawrath, C., & Geldner, N. (2012). Casparian strip diffusion barrier in *Arabidopsis* is made of a lignin polymer without suberin. *Proceedings of the National Academy of Sciences of the United States of America*, 109(25), 10101–6.
- Ndamukong, I., Chetram, A., Saleh, A., & Avramova, Z. (2009). Wall-modifying genes regulated by the *Arabidopsis* homolog of trithorax, *ATX1*: repression of the *XTH33* gene as a test case. *The Plant Journal : For Cell and Molecular Biology*, 58(4), 541–53.
- Nicol, F., His, I., Jauneau, a, Vernhettes, S., Canut, H., & Höfte, H. (1998). A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. *The EMBO Journal*, 17(19), 5563–76.
- Notley, S. M., Pettersson, B., & Wågberg, L. (2004). Direct measurement of attractive van der Waals' forces between regenerated cellulose surfaces in an aqueous environment. *Journal of the American Chemical Society*, 126(43), 13930–1.
- O'Neill, M. a, Eberhard, S., Albersheim, P., & Darvill, a G. (2001). Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth. *Science*, 294(5543), 846–9.
- Oritani, T., & Kiyota, H. (2003). Biosynthesis and metabolism of abscisic acid and related compounds. *Natural Product Reports*, 20(4), 414.
- Osmani, S. a, Bak, S., & Møller, B. L. (2009). Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. *Phytochemistry*, 70(3), 325–47.
- Paquette, S., Møller, B. L., & Bak, S. (2003). On the origin of family 1 plant glycosyltransferases. *Phytochemistry*, 62(3), 399–413.
- Park, S., Baker, J. O., Himmel, M. E., Parilla, P. a, & Johnson, D. K. (2010). Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnology for Biofuels*, 3, 10.
- Park, Y. B., & Cosgrove, D. J. (2012a). A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant Physiology*, 158(4), 1933–43.
- Park, Y. B., & Cosgrove, D. J. (2012b). Changes in cell wall biomechanical properties in the xyloglucan-deficient xxt1/xtt2 mutant of *Arabidopsis*. *Plant Physiology*, 158(1), 465–75.

- Pattathil, S., Avci, U., Baldwin, D., Swennes, A. G., McGill, J. a, Popper, Z., Bootten, T., Albert, A., Davis, R. H., Chennareddy, C., Dong, R., O'Shea, B., Rossi, R., Leoff, C., Freshour, G., Narra, R., O'Neil, M., York, W. S., & Hahn, M. G. (2010). A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiology*, 153(2), 514–25.
- Patten, A. M., Jourdes, M., Cardenas, C. L., Laskar, D. D., Nakazawa, Y., Chung, B. Y., Franceschi, V. R., Davin, L. B., & Lewis, N. G. (2010). Probing native lignin macromolecular configuration in *Arabidopsis thaliana* in specific cell wall types: further insights into limited substrate degeneracy and assembly of the lignins of *ref8*, *fah 1-2* and *C4H::F5H* lines. *Molecular BioSystems*, 6(3), 499–515.
- Peaucelle, A., Braybrook, S., & Höfte, H. (2012). Cell wall mechanics and growth control in plants: the role of pectins revisited. *Frontiers in Plant Science*, 3(June), 121.
- Popper, Z. A., & Fry, S. C. (2008). Xyloglucan-pectin linkages are formed intra-protoplasmically, contribute to wall-assembly, and remain stable in the cell wall. *Planta*, 227(4), 781–94.
- Priest, D. M., Jackson, R. G., Ashford, D. A., Abrams, S. R., & Bowles, D. J. (2005). The use of abscisic acid analogues to analyse the substrate selectivity of UGT71B6, a UDP-glycosyltransferase of *Arabidopsis thaliana*. *FEBS Letters*, 579(20), 4454–8.
- Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P. F., Marita, J. M., Hatfield, R. D., Ralph, S. A., Christensen, J. H., & Boerjan, W. (2004). Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenyl- propanoids. *Phytochemistry Reviews*, 3(1/2), 29–60.
- Reiter, W. D., & Vanzin, G. F. (2001). Molecular genetics of nucleotide sugar interconversion pathways in plants. *Plant Molecular Biology*, 47(1-2), 95–113.
- Reiter, W. D. (2008). Biochemical genetics of nucleotide sugar interconversion reactions. *Current Opinion in Plant Biology*, 11(3), 236–43.
- Rodicio, R., & Heinisch, J. (2010). Together we are strong: cell wall integrity sensors in yeasts. *Yeast*, 27(8), 531–540.
- Rogers, L. A., & Campbell, M. M. (2004). The genetic control of lignin deposition during plant growth and development. *New Phytologist*, 164(1), 17–30.
- Rose, J. K. C., Braam, J., Fry, S. C., & Nishitani, K. (2002). The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. *Plant & Cell Physiology*, 43(12), 1421–35.
- Ross, J., Li, Y., Lim, E., & Bowles, D. J. (2001). Higher plant glycosyltransferases. *Genome Biology*, 2(2), REVIEWS3004.
- Routaboul, J. M., Kerhoas, L., Debeaujon, I., Pourcel, L., Caboche, M., Einhorn, J., & Lepiniec, L. (2006). Flavonoid diversity and biosynthesis in seed of *Arabidopsis thaliana*. *Planta*, 224(1), 96–107.
- Ruoslahti, E., & Pierschbacher, M. (1987). New perspectives in cell adhesion: RGD and integrins. *Science*, 238(4826), 491–497.
- Sato, S., Kato, T., Kakegawa, K., Ishii, T., Liu, Y. G., Awano, T., Takabe, K., Nishiyama, Y., Kuga, S., Sato, S., Nakamura, Y., Tabata, S., & Shibata, D. (2001). Role of the putative membrane-bound endo-1,4-beta-glucanase *KORRIGAN* in cell elongation and cellulose synthesis in *Arabidopsis thaliana*. *Plant & Cell Physiology*, 42(3), 251–63.

- Scheller, H. V., & Ulvskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology*, 61, 263–289.
- Schiermeier, Q., Tollefson, J., & Scully, T. (2008). Energy alternatives: electricity without carbon. *Nature*, 454(7206), 816–823.
- Seifert, G. J. (2004). Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. *Current Opinion in Plant Biology*, 7(3), 277–284.
- Shao, H., He, X., & Achnine, L. (2005). Crystal structures of a multifunctional triterpene/ flavonoid glycosyltransferase from *Medicago truncatula*. *The Plant Cell*, 17(11), 3141–3154.
- Shedletzky, E., Shmuel, M., Delmer, D. P., & Lampert, D. T. (1990). Adaptation and growth of tomato cells on the herbicide 2,6-dichlorobenzonitrile leads to production of unique cell walls virtually lacking a cellulose-xyloglucan network. *Plant Physiology*, 94(3), 980–987.
- Shipp, M., Nadella, R., Gao, H., Farkas, V., Sigrist, H., & Faik, A. (2008). Glyco-array technology for efficient monitoring of plant cell wall glycosyltransferase activities. *Glycoconjugate Journal*, 25(1), 49–58.
- Sigrist, C. J. A., Cerutti, L., de Castro, E., Langendijk-Genevaux, P. S., Bulliard, V., Bairoch, A., & Hulo, N. (2010). PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Research*, 38(Database issue), D161–6.
- Simmons, B. A., Loque, D., & Blanch, H. W. (2008). Next-generation biomass feedstocks for biofuel production. *Genome Biology*, 9(12), 242.
- Singh, K. (2002). Transcription factors in plant defense and stress responses. *Current Opinion in Plant Biology*, 5(5), 430–436.
- Sinha, J., Plantz, B. a, Inan, M., & Meagher, M. M. (2005). Causes of proteolytic degradation of secreted recombinant proteins produced in methylotrophic yeast *Pichia pastoris*: case study with recombinant ovine interferon-tau. *Biotechnology and Bioengineering*, 89(1), 102–112.
- Smyth, D., Bowman, J., & Meyerowitz, E. (1990). Early Flower Development in *Arabidopsis*. *Plant Cell*, 2, 755–767.
- Somerville, C. (2006). Cellulose synthesis in higher plants. *Annual Review of Cell and Developmental Biology*, 22, 53–78.
- Stephens, C. H., Whitmore, P. M., Morris, H. R., & Bier, M. E. (2008). Hydrolysis of the amorphous cellulose in cotton-based paper. *Biomacromolecules*, 9(4), 1093–9.
- Sterling, J. D., Atmodjo, M. A., Inwood, S. E., Kumar Kolli, V. S., Quigley, H. F., Hahn, M. G., & Mohnen, D. (2006). Functional identification of an *Arabidopsis* pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, 103(13), 5236–41.
- Sticklen, M. B. (2008). Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. *Nature Reviews. Genetics*, 9(6), 433–43.
- Szyjanowicz, P. M. J., McKinnon, I., Taylor, N. G., Gardiner, J., Jarvis, M. C., & Turner, S. R. (2004). The *irregular xylem 2* mutant is an allele of *korrigan* that affects the secondary cell wall of *Arabidopsis thaliana*. *The Plant Journal*, 37(5), 730–740.

- Takahashi, J., Rudsander, U. J., Hedenström, M., Banasiak, A., Harholt, J., Amelot, N., Immerzeel, P., Ryden, P., Endo, S., Ibatullin, F. M., Brumer, H., del Campillo, E., Master, E. R., Scheller, H. V., Sundberg, B., Teeri, T. T., & Mellerowicz, E. J. (2009). KORRIGAN1 and its aspen homolog *PttCel9A1* decrease cellulose crystallinity in *Arabidopsis* stems. *Plant & Cell Physiology*, 50(6), 1099–115.
- Taylor, N. G. (2008). Cellulose biosynthesis and deposition in higher plants. *The New Phytologist*, 178(2), 239–52.
- Taylor, N. G., Howells, R. M., Huttly, A. K., Vickers, K., & Turner, S. R. (2003). Interactions among three distinct Cesa proteins essential for cellulose synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 100(3), 1450–5.
- Turner, S. R., & Somerville, C. R. (1997). Collapsed xylem phenotype of *Arabidopsis* identifies mutants deficient in cellulose deposition in the secondary cell wall. *The Plant Cell*, 9(5), 689–701.
- Udvardi, M. K., Czechowski, T., & Scheible, W. R. (2008). Eleven golden rules of quantitative RT-PCR. *The Plant Cell*, 20(7), 1736–7.
- Updegraff, D. M. (1969). Semimicro determination of cellulose in biological materials. *Analytical Biochemistry*, 32(3), 420–4.
- Valentin, R., Cerclier, C., Geneix, N., Aguié-Béghin, V., Gaillard, C., Ralet, M. C., & Cathala, B. (2010). Elaboration of extensin-pectin thin film model of primary plant cell wall. *Langmuir: The ACS Journal of Surfaces and Colloids*, 26(12), 9891–8.
- Van Sandt, V. S. T., Guisez, Y., Verbelen, J. P., & Vissenberg, K. (2006). Analysis of a xyloglucan endotransglycosylase/hydrolase (XTH) from the lycopodiophyte *Selaginella kraussiana* suggests that XTH sequence characteristics and function are highly conserved during the evolution of vascular plants. *Journal of Experimental Botany*, 57(12), 2909–22.
- Vanholme, R., Demedts, B., Morreel, K., Ralph, J., & Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiology*, 153(3), 895–905.
- Vanholme, R., Van Acker, R., & Boerjan, W. (2010). Potential of *Arabidopsis* systems biology to advance the biofuel field. *Trends in Biotechnology*, 28(11), 543–7.
- Vanzin, G. F., Madson, M., Carpita, N. C., Raikhel, N. V., Keegstra, K., & Reiter, W. D. (2002). The *mur2* mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. *Proceedings of the National Academy of Sciences of the United States of America*, 99(5), 3340–5.
- Vissenberg, K., Van Sandt, V., Fry, S. C., & Verbelen, J. P. (2003). Xyloglucan endotransglucosylase action is high in the root elongation zone and in the trichoblasts of all vascular plants from *Selaginella* to *Zea mays*. *Journal of Experimental Botany*, 54(381), 335–344.
- Von Saint Paul, V., Zhang, W., Kanawati, B., Geist, B., Faus-Kessler, T., Schmitt-kopplin, P., & Schäffner, A. R. (2011). The *Arabidopsis* glucosyltransferase UGT76B1 conjugates isoleucic acid and modulates plant defense and senescence. *The Plant Cell*, 23(11), 4124–45.
- Wan, Y., Gritsch, C., Tryfona, T., Ray, M. J., Andongabo, A., Hassani-Pak, K., Jones, H. D., Dupree, P., Karp, A., Shewry, P. R., & Mitchell, R. A. C. (2014). Secondary cell wall composition and candidate gene expression in developing willow (*Salix purpurea*) stems. *Planta*. Feb 7, Epub ahead of print

- Wang, Y., Liu, C., Li, K., Sun, F., Hu, H., Li, X., Zhao, Y., Han, C., Zhang, W., Duan, Y., Liu, M., & Li, X. (2007). *Arabidopsis* EIN2 modulates stress response through abscisic acid response pathway. *Plant Molecular Biology*, 64(6), 633–44.
- Wang, Y., Wang, T., Li, K., & Li, X. (2007). Genetic analysis of involvement of ETR1 in plant response to salt and osmotic stress. *Plant Growth Regulation*, 54(3), 261–269.
- Weng, J. K., Akiyama, T., Ralph, J., & Chapple, C. (2011). Independent recruitment of an O-methyltransferase for syringyl lignin biosynthesis in *Selaginella moellendorffii*. *The Plant Cell*, 23(7), 2708–24.
- Weng, J. K., & Chapple, C. (2010). The origin and evolution of lignin biosynthesis. *The New Phytologist*, 187(2), 273–85.
- Wilkie, G. S., Dickson, K. S., & Gray, N. K. (2003). Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. *Trends in Biochemical Sciences*, 28(4), 182–8.
- Williamson, R. E., Burn, J. E., & Hocart, C. H. (2002). Towards the mechanism of cellulose synthesis. *Trends in Plant Science*, 7(10), 461–7.
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., & Provart, N. J. (2007). An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PloS One*, 2(8), e718.
- Wormit, A., Butt, S., Chairam, I., McKenna, J. F., Nunes-Nesi, A., Kjaer, L., O'Donnelly, K., Fernie, A. R., Woscholski, R., Barter, M. C. & Hamann, T. (2012). Osmosensitive changes of carbohydrate metabolism in response to cellulose biosynthesis inhibition. *Plant Physiology*, 159(1), 105–17.
- Yonekura-Sakakibara, K., & Hanada, K. (2011). An evolutionary view of functional diversity in family 1 glycosyltransferases. *The Plant Journal*, 66(1), 182–93.
- Zhong, R., & Ye, Z. H. (2007). Regulation of cell wall biosynthesis. *Current Opinion in Plant Biology*, 10(6), 564–72.
- Zhu, X. F., Shi, Y. Z., Lei, G. J., Fry, S. C., Zhang, B. C., Zhou, Y. H., ... Zheng, S. J. (2012). *XTH31*, encoding an in vitro XEH/XET-active enzyme, regulates aluminum sensitivity by modulating in vivo XET action, cell wall xyloglucan content, and aluminum binding capacity in *Arabidopsis*. *The Plant Cell*, 24(11), 4731–47.