Functional testing of xylan-associated genes in *Arabidopsis* and application of CRISPR/Cas9 gene editing in hybrid poplar trees

Nganeavhutshilo Nangammbi

Submitted in partial fulfilment
of the requirements for the degree

Magister Scientiae

In the Faculty of Natural and Agricultural Sciences
Department of Biochemistry, Genetics and Microbiology
University of Pretoria
Pretoria
August 2021

Under the supervision of **Dr Victoria J. Maloney**

and co-supervision of **Prof Alexander A. Myburg** and **Prof Eshchar Mizrachi**

DECLARATION

I, Nganeavhutshilo Nangammbi, declare that this dissertation, which I hereby submit for the degree MSc Biotechnology at the University of Pretoria, is my own work and has not been submitted by me for a degree at this or any other tertiary institution.

Nganeavhutshilo Nangammbi

August 22, 2021

DISSERTATION SUMMMARY

Functional testing of xylan-associated genes in *Arabidopsis* and application of CRISPR/Cas9 gene editing in hybrid poplar trees

Nganeavhutshilo Nangammbi

Supervised by **Dr Victoria J. Maloney**

Co-supervised by **Prof Alexander A. Myburg** and **Prof Eshchar Mizrachi**

Submitted in partial fulfilment of the requirements of the degree *Magister Scientiae*

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

The xylem cells of vascular plants are surrounded by a tough, hydrophobic, thickened secondary cell wall that is predominantly comprised of a polysaccharide matrix of cellulose and hemicellulose impregnated with a complex polyphenolic biopolymer, lignin. These biopolymers make up the bulk of lignocellulosic biomass, the most abundant renewable raw material on earth. In the bio-based economy, lignocellulosic biomass can be exploited for a wide range of applications such as the production of timber, textiles, wood-composites, paper, and other bio-composite materials. In addition, lignocellulosic biomass from fast-growing plantation trees such as *Populus* and *Eucalyptus* species holds a promise as a raw material in the production of second-generation biofuels and other forms of bioenergy. Generating these bio-based products from lignocellulosic biomass often requires complete separation of the wood cell wall matrix into individual biopolymer constituents (e.g. cellulose, xylan and lignin) during the pulping process, or the individual sugar and phenolic monomers of the polymers for use in biorefinery approaches. However, the physical structure and interaction of biopolymers hamper the efficient deconstruction of wood which means that large amounts of energy and/or

chemicals are required, often leading to reduced product yields. Novel genetic engineering approaches have been devised to engineer cell walls that are easier to deconstruct. The success of this approach requires a good understanding of the biological roles of cell wall biosynthetic genes and how altering such genes can potentially alter plant physiology, cell wall chemistry and plant growth and health.

The work presented in this MSc dissertation aimed to assess the impact of gene knock-out mutations on selected cell wall-related genes on the cell wall morphology and chemistry in *Arabidopsis* and poplar plants. It was hypothesized that altering the selected endogenous genes will have a specific impact on plant growth and cell wall development. Using data generated from a previous systems genetics study, five genes, namely, *XYN1*, *MAN2*, *UXS6*, *EMP70*, and *ARAF1*, were selected for this functional testing work, the basis of this selection is discussed in detail later. Analyses of *Arabidopsis* T-DNA insertional mutants in these genes showed that *Atxyn1* mutants have slow growth which can be attributed to the significantly shorter roots observed during early growth, furthermore, mutants showed increased lignin content in dry stems. Conversely, *emp70* mutants germinated significantly earlier and continued to show growth advantage compared to wild-type, although such growth advantage was not accompanied by any alterations in secondary cell chemistry. Thus, from the selected genes, consistent growth phenotypes were observed in *xyn1* and *emp70* *Arabidopsis* mutants, however, only *xyn1* mutants had altered cell wall chemistry.

Finally, an additional goal of this work was to use CRISPR/Cas9 gene-editing to target cell wall genes in *Populus* and assess the effect of the mutation on plant growth and development. To address this objective, *XYN1* and *MAN2* homologs were CRISPR-targeted in *Populus alba* x *P.tremula* (P717) hybrid. The basis for the selection of these gene homologs is discussed in the relevant research chapter. Following several rounds of P717 *Agrobacterium*-mediated transformation, few CRISPR/Cas9 *XYN1* and *MAN2* putative transgenics were successfully generated, however, no genetic modification was

observed in these transgenic plants. Altogether, this work presents initial steps towards elucidating the precise biological roles of the putative cell wall genes not only in *Arabidopsis* but also in woody energy crops. A comprehensive understanding of the exact roles of these genes in cell wall development will be highly beneficial in uncovering their complex functional interactions with other genes. This will enable the establishment of specific relationships between the organism's genes and some of the phenotypes observed here and elsewhere.

PREFACE

Lignocellulosic biomass is important for many essential applications, ranging from the production of biomaterials to being used as a major feedstock for the production of second-generation biofuels. Various global biorefinery companies produce high-quality products for pharmaceutical, textile, food, and industrial applications. Amongst others, about 1.4 million tons of dissolving pulp (DP) are produced per year, accounting for about 16% of the global market share (https://www.sappi.com/dissolving-pulp). In addition to DP, a wide range of papers, varying from graphic and release papers, to packaging, and casting, are also produced from lignocellulosic biomass. Notably, the process of papermaking consumes only half of the raw wood feedstock for paper and DP production, with the remainder being used to generate the power necessary to power the mill, thus, utilizing bioenergy and reducing carbon emissions while enhancing the cost-efficiency. Furthermore, the biorefineries are involved in the production of vital biomaterials and chemicals including chemical cellulose, nanocellulose, lignin, and sweeteners, which are generated from hemicelluloses. In addition to these traditional products, novel biorefinery processes capable of extracting even more value from wood to generate some specialty chemicals and biomaterials have been developed. One such specialty product is the Sappi Symbio, which is a new composite material generated by combining high-quality cellulose and thermoplastics, it can be used in processing equipment, consumer electronics, and the automotive industry.

This significant reliance on wood biomass as a natural feedstock is ideal not only because of its renewable nature but also the fact that the biomass cycle continuously sequesters carbon dioxide, which significantly reduces the carbon footprint on the environment. Altogether, wood biomass is necessary for the production of important bio-based materials and bioenergy and reducing the impact of carbon dioxide emissions, which largely contributes to a sustainable bio-based economy. To realize these benefits requires efficient means to extract value from wood with minimal wastage. This

primarily entails the complete separation of wood biopolymers into their units. Furthermore, value is further extracted from waste and by-products to produce high-energy chemicals, thereby maximizing yield while minimizing biorefinery costs. However, the structure and strong interactions between wood biopolymers lead to resistance of wood to bioconversion, creating challenges for efficient industrial bioprocessing. Currently, the most efficient biorefinery approaches in South Africa are only capable of converting about 47% of wood into paper and pulp products, with the rest being lost as waste (https://www.csir.co.za/biorefinery). The impact of wood structure on bioprocessing is thus of major concern in the bio-based economy, and subsequently, significant efforts have been made to understand the specifics of how wood structure contributes to biomass recalcitrance.

Within the secondary cell wall, xylan and lignin are the major players in biomass recalcitrance. Xylan backbone is often variably substituted with D-glucuronic acid (GlcA) which can further be methylated at specific backbone positions. GlcA substitutions are necessary for the formation of the functionally mature cell wall and are thus necessary for normal plant growth and development. GlcA substitutions at the xylan backbone can either be evenly spaced, forming the major xylan domain or, unevenly spaced, to yield the minor domain. These distinct domains dictate how xylan interacts with cellulose and lignin. During dissolving pulp processing (DPP), lignin and lignin-carbohydrate complexes physically block hydrolytic enzymes' access to xylan, which leads to inefficient xylan removal from wood biomass. Furthermore, complex lignin linkages make lignin highly resistant to degradation, often requiring a suite of enzymes for its complete deconstruction. Besides GlcA substitutions, the separation of secondary cell wall biopolymers is further significantly hindered by the presence of *O*-acetyl groups on the xylan backbone. Like the complex interaction between lignin and xylan, acetyl moieties also contribute to the steric hindrance of hydrolytic enzymes, which reduces xylan removal and cellulose digestibility during DPP. Furthermore, acetyl groups lead to significant pH alteration when acetyl is released into the media, this inhibits the activity of fermentative microbes during fermentation. Thus,

xylan acetyl groups and the interaction between xylan and lignin as predefined by GlcA substitution pattern has a profound effect on efficient bioprocessing. Accordingly, additional pre-treatments such as deacetylation and delignification are necessary for improved biopolymer separation and complete deconstruction. However, these often involve the use of harsh chemicals which create a non-ideal environment for the optimal functioning of hydrolytic enzymes, furthermore, additional forms of recalcitrance are often introduced due to significant alteration of the native cell wall structure. Together, these ultimately lead to compromised biomass yield coupled with substantial bioprocessing costs.

Alternative approaches aimed at improvement of bioprocessing efficiency are thus necessary, indeed, with a significant understanding of the secondary cell wall development, genetic engineering of plants' secondary cell walls is a promising strategy for the improvement of bioprocessing efficiency. By modifying cell wall biosynthetic genes necessary for the biosynthesis of cell wall biopolymers, we can engineer plants with easier-to-deconstruct cell walls or any other favourable bioprocessing traits. Several previous studies have modified xylan backbone biosynthesis genes, GlcA substitutions, methylation, and acetylation genes to produce trees with improved saccharification yield with no or minimal growth penalties. Similarly, targeting lignin biosynthetic genes has led to plants with reduced lignin and whose cell walls can be easily deconstructed. The success of the genetic engineering approach requires a thorough understanding of the biological roles of cell wall biosynthetic genes. More importantly, the correlation between the expression of such genes with specific bioprocessing traits also has to be established to aid in identifying which genes are worth targeting to potentially improve industrial bioprocessing efficiency.

With this in mind, this MSc **aimed** to determine how selected xylan-associated genes impact cell wall development, plant growth, and physiology in *Arabidopsis* and *Populus*. This involved selecting the appropriate genes that when genetically altered, are more likely to yield plants with modified cell walls that can be easily deconstructed during bioprocessing. The selection of genes relied on a previous systems genetics study that used *Eucalyptus* backcross populations to not only identify putative xylan modification genes but also establish the correlation of their expression to some important bioprocessing traits. Of interest to us were genes whose expression was negatively correlated with pulp yield and positively correlated with xylan acetylation; these will be genes that when knocked out, xylan acetylation could be reduced, which would theoretically increase pulp yield as predicted by correlation analyses. To further substantiate the significance of these genes in cell wall development, we perused through literature and online databases to understand what is currently known about the involvement of these genes in cell wall biosynthesis. After selecting the five genes (namely, *XYN1, ARAF1, EMP70, UXS6,* and *MAN2*), we obtained T-DNA insertion mutants of each gene and performed a growth trial followed by in depth phenotyping. Furthermore, we investigated the impact of the mutation on each gene on secondary cell wall development by integrating cell wall microscopy and chemistry analyses. We finalize by scaling up the functional testing study to *Populus* to investigate how targeting these genes can alter plant growth in woody plants. Here, we first perform phylogenetic analyses to establish the evolutionary relationships amongst *Eucalyptus grandis* (from which the genes were originally identified), *Arabidopsis thaliana,* and *Populus trichocarpa* homologs. We further analysed the expression pattern of selected genes during secondary cell wall biosynthesis in poplar. Finally, the CRISPR/Cas9 protocol involving guide RNA design, off-target identification, and *Agrobacterium*-mediated transformation was followed for targeting selected genes and ultimately transforming hybrid poplar (*Populus alba* x *P. tremula*). We generated few putative transgenic plants which we screened for possible mutations.

Altogether, preliminary growth phenotypes corresponding to *Arabidopsis* mutants of selected xylan-associated genes were established, the impact of such mutations on secondary cell wall morphology and chemistry was further dissected. Importantly, we apply the CRISPR/Cas9 gene editing approach to target cell wall biosynthetic genes in the model woody plant poplar. The observations made in this study will aid in subsequent studies aimed at functionally characterizing the candidate genes. Furthermore, the preliminary phenotypes obtained here provide an initial clue on the biological roles of selected genes. Finally, by integrating the growth phenotypes and overall impact of the mutation on cell wall chemistry, it would be possible to precisely ascertain the precise impact of gene modification on lignocellulose bioprocessing.

This dissertation describes the functional analyses of *Arabidopsis* and *Populus* putative mutant phenotypes corresponding to the selected xylan-associated genes. The dissertation follows the structure below:

Chapter 1 is a literature review , focusing extensively on the biosynthesis of cell wall biopolymers, the overall structure of the secondary cell wall, and how this physical structure potentially affects the efficiency of industrial bioprocessing. I describe genes that are currently known to play major roles in SCW biosynthesis. To do this, some of the significant functional analyses studies that helped identify and/or functionally characterize these genes are revisited and discussed thoroughly. I further describe the metabolic requirements for biopolymer biosynthesis, such as sugar nucleotide precursors necessary for cellulose and xylan backbone synthesis, and donor metabolites required for modification of the cell wall biopolymers. Next I focus on understanding how the cell wall structure and its modifications contribute to biomass recalcitrance. Here, I specifically discuss some of the measures that have been employed to perturb the effects of biomass recalcitrance in industrial bioprocessing. Furthermore, I review some genetic engineering studies that have successfully manipulated SCW genes to potentially

improve saccharification efficiency. Finally, I discuss how CRISPR and CRISPR-based second-generation gene editing techniques present a promise for a more robust and reliable means of bioprocessing improvement through manipulation of the cell wall.

In **Chapter 2**, I begin by summarizing the overall procedure that was followed to select the five xylan-associated genes which were functionally tested in this study. Next, all notable mutant growth phenotypes corresponding to the mutants of selected genes were described, specifically focusing on what these phenotypes possibly suggest regarding the overall plant growth and physiology. Notably, *xyn1* and *emp70* mutant phenotype were discussed in more detail as the work done here was the very first to attempt to decipher the growth phenotypes of these so-far-uncharacterized genes in *Arabidopsis*. Finally, I discuss in detail the impact or lack thereof of gene mutations on cell wall morphology by referring to the microscopic analyses of the xylem vessel structure and subsequent analyses of cell wall chemistry.

Chapter 3 focuses on the essentials of the CRISPR/Cas9 gene editing for targeting cell wall biosynthetic genes in woody plants. I establish the phylogenetic maps and gene expression profiles for *XYN1* and *MAN2* poplar homologs which were selected for CRISPR/Cas9 editing in hybrid poplar. Then, I focus on the specific steps we followed for guide RNA design, cloning of gRNAs into CRISPR/Cas9 vectors, and, the ultimate transformation of poplar. I finalise by discussing the outcome of transformations and how we screened for mutations in putative transgenic plants, moreover, I briefly discuss the limitations of CRISPR/Cas9 and how these may contribute to poor recovery of transformants and lack of mutations in transgenic plants.

Chapter 4 is the conclusion, here, I provide brief conclusions to some of the major findings from all the research chapters. First, the implications and limitations of selecting genes from a gene set generated through population-based identification methods are discussed. Then I review the lessons learned from functional testing of uncharacterized candidate genes by focusing on the major preliminary phenotypes observed in *Arabidopsis* mutants. I specifically focus on the *EMP70* mutant phenotype and suggest a comprehensive model which describe the possible role of EMP70 in the crosstalk between Cu and Abscisic acid (ABA), and how this may ultimately affect seed germination as observed in the current study. Finally, with reference to the work involving CRISPR/Cas9 targeting of *XYN1* and *MAN2* in poplar, the issues and novel solutions around poor mutation generation and transformation efficiency in woody plants are dissected.

ACKNOWLEDGEMENTS

Mr Ntodeni Nangammbi and **Ms Thinamano Tshipani**, my parents. I know it took an extremely long time to reach this stage, you gave up on waiting for me to complete this stage of life, but mom and dad, this is it, I am done (for now). Thank you for all the slim patience you had left until this point, more importantly, the emotional support and all the love you showed while I was on it. I promise this is the start of something different and worthwhile for me and both of you.

Dr Victoria Maloney, my supervisor. I have no doubt that if I was not under your supervision I would not have done it the way I have. You were always open to any questions, even when the aspect is perhaps obvious to the experienced, that meant a lot. I hope you remember that you and I never had to set up meetings, instead, you were always available for me, anytime, any day. Your insistence on perfection and attention to detail has been the most important characteristic I kept in mind throughout this work. I finish by thanking you for believing in my potential perhaps more than I believed in myself, which has made it possible for me to make it this far, I highly appreciate it.

Prof Alexander Myburg. As my co-supervisor, your motivational words and belief in my potential were sufficient for me to stay for MSc, and I never regretted it. I would like to first thank you for always being open to all the ideas and suggestions that my supervisor and I would put forward. More importantly, when I started this work, I had absolutely no background in plant genetics, you managed to introduce me to an extremely important field of forest genetics and I now understand and appreciate it more. While in the learning process you never made me feel inadequate or not knowledgeable enough, which has helped me stay focused on learning and understanding several aspects of the field, which have contributed to the success of this study and I have no doubt it will be essential for my future endeavours.

Prof Eshchar Mizrachi and Dr Steven Hussey. Thank you all for your exceptional inputs during all the progress meetings, I know sometimes your feedback was painful, but I appreciate it and have always taken that as rather constructive. I thank you for all your suggestions and for tailoring my project to perfection. Esh, the small talks we had (probably only three) were significant in making me believe I got the strength and capability to stay in the field, I only wish we had more talks.

Associate Prof Vida Van Staden and **Prof Thulani Makhalanyane,** my honours lecturers. You did not only care about my academic life, but you showed concern about my personal life as well. You made it possible for me to get university accommodation, you gave me food when I had no hope of where I could find food. You were always available as my referees and saw a profound potential in me. Thank you for all your support, I would have most definitely dropped out very early in Honours year if not because of your assistance.

Thoriso K Sesing (University of Pretoria), and Tanyaradzwa C Kiwa (University of Cape Town), and. You guys are awesome, to my UP buddy, I will not forget the exceptional support you gave me when I got here even though you did not even know me. Your consistent confidence in me and support have been noted and are appreciated so much. Thank you for all the jokes, hanging out where possible, and cheering me up during the hardest times.

My siblings**, Gudani, Thalukanyo, Tovhowani,** and **Khwathelani Nangammbi**. Like our parents, I know you have all been waiting for this moment, and yes, it has happened. Thank you for all the financial and emotional support from the first day I started this journey. I shall not forget our never-ending laughs, the get-together times, and all the wonderful moments we shared throughout these hectic years.

The Forest Molecular Genetics (FMG) Programme, Forestry and Agricultural Biotechnology Institute (FABI), the Department of Physics at the University of Pretoria, and Sappi. I thank my research group and the research institute for providing a conducive environment and platform to perform all my experiments. From FMG, I specifically thank **Dr. Mark Maistry**, **Lungile Mabuza**, and **Sung Hyu Luke Kim** who have both made significant academic and personal support to the success of this project. The tissue culture platform and *Arabidopsis* growth room management team are also appreciated. Further thanking the department of physics, particularly **Erna Van Wilpe** who assisted me in my microscopic analyses. Finally, I thank **L'Zanne Jansen Van Rensburg** and her entire technical team for performing all my cell wall chemistry analyses.

National Research Foundation (NRF), the Department of Science and Technology (DST), Technology Innovation Agency (TIA) and Sappi. Finally, I am extending my acknowledgment to my main funding bodies for awarding me the NRF Master's Innovation Scholarship award which was more than enough for this study to be carried out.

"You stay at the top of your game by continuing what got you there in the first place"
Roger Federer: Tennislegende versus GLE Coupe (2015).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

4CL	4-COUMARATE:CoA
4-*O*-MeGlcA	a-1,2-linked 4-*O*-methylglucuronic acid
ABA	Abscisic acid
ABC	ATP-binding cassette
ABE	Adenine base editor
ACL	ATP-CITRATE LYASE
AdoMetS	ADENOSYLMETHIONINE SYNTHETASE
AG	AGAMOUS
amiRNA	Artificial microRNA
ARAF	ALPHA-L-ARABINOFURANOSIDASE
AXE	ACETYL XYLAN ESTERASE
AXS	UDP-Api/UDP-Xyl SYNTHASE
AXY	ALTERED XYLOGLUCAN
BC	Biotechnology cluster
BXL	β-XYLOSIDASE
Cas9	CRISPR-associated (Cas) protein 9
CBE	Cytodine base editor
CCoAOMT	CAFFEOYL CoA O-METHYLTRANSFERASE
CCR	CINNAMOYL-CoA REDUCTASE
CESA	CELLULOSE SYNTHASE
cINV	CYTOSOLIC INVERTASE
CMT	Cortical microtubule
CMP	Chromatin-modulating peptide
COMT	CAFFEIC ACID *O*-METHYLTRANSFERASE
CRISPR	Clustered regularly interspaced short palindromic repeat
CSC	Cellulose synthase complex
CSI	CELLULOSE SYNTHASE INTERACTING
CSLD5	CELLULOSE SYNTHASE-LIKE PROTEIN D5
DP	Dissolving pulp
DPP	Dissolving pulp processing
DSB	Double-stranded breaks
DUF	Domain of unknown function
EM	Expression module
EMP70	ENDOMEMBRANE PROTEIN 70
ESK	ESKIMO
GA	Gibberellic acid
GFP	Green-fluorescent protein
GlcA	D-glucuronic acid
GO	Gene ontology
GT	Glycosyltransferase
GUX	GLUCURONIC ACID SUBSTITUTION OF XYLAN

GX	Glucuronoxylan
GXM	GLUCURONOXYLAN METHYLTRANSFERASE
HDR	Homology directed repair
HNH	Homing endonuclease
HXK	HEXOKINASE
IRX	IRREGULAR XYLEM
iTOL	Interactive Tree of Life
LB	Lysogeny broth
LFY	LEAFY
LM	Light microscope
LR	London resin
LRC	Lignin-related compound
MAN	ENDO-BETA-MANNASE
MAP	Microtubule-associated protein
MS	Murashige and Skoog
NDP	Nucleotide diphosphate
NHEJ	Non-homologous end joining
NMR	Nuclear magnetic resonance
P717	*Populus alba* x *Populus tremula* hybrid aspen
PAM	Protospacer adjacent motif
PCA	Principal component analysis
PDS	*PHYTOENE DESATURASE*
PE	Prime editor
pegRNA	Prime editing guide RNA
PM	Plasma membrane
RT	REVERSE TRANSCRIPTASE
RuvC	RESOLVASE protein
RWA	REDUCE WALL ACETYLATION
SAM-e	*S*-Adenosyl methionine
SCW	Secondary cell wall
sgRNA	Single guide RNA
SNP	Single nucleotide polymorphism
SSB	Single-stranded breaks
SSF	Simultaneous saccharification and fermentation
SUSY	SUCROSE SYNTHASE
TALEN	Transcription-like effector nuclease
TBL	Trichome-birefringence-like protein family
T-DNA	Transfer DNA
TM9/TMN	TRANSMEMBRANE 9
UBQ10	Ubiquitin 10
UDP	Uridine diphosphate
UDP-Glc	UDP-glucose
UDP-GlcA	UDP-D-glucuronate
UDP-Xyl	UDP-D-Xylose

UGDH	UDP-GLC 6 DEHYDROGENASE
UTR	Untranslated region
UXS	UDP-XYL SYNTHASE
UXT	UDP-XYL TRANSPORTER
Ws	Wassilewskija ecotype
XSC	Xylan synthase complex
XYL	D-Xylosyl
XYN	XYLANASE
ZFN	Zinc finger nuclease

Chapter 1

Literature Review:

Secondary cell wall biosynthesis and the impact of lignocellulose ultrastructure on industrial bioprocessing

1.1 Introduction

Lignocellulosic biomass is predominantly comprised of the secondary cell wall (SCW) biopolymers; cellulose synthesized at the plasma membrane, hemicellulose synthesized in the Golgi, and lignin produced in the cell wall (Brown *et al.* 2005; Zhong *et al.* 2005; Mohnen 2008). These biopolymers are widely exploited for the production of a wide range of value-added products such as timber, paper, textiles, and bioplastics (Plomion *et al.* 2001; Carroll and Somerville 2008). Furthermore, significant efforts have been made to utilize woody biomass generated from energy crops such as *Populus* as a raw material for second-generation biofuel production (Pauly and Keegstra 2008; Pawar *et al.* 2013). Such reliance on woody biomass for a wide range of applications significantly contributes to the generation and maintenance of the bioeconomy. A healthy and thriving bioeconomy helps provide solutions to challenges involving climate change, sustainable development, food, and energy security (Golberg *et al.* 2016; Schütte 2018).

Owing to their fast growth rate and increased vegetative propagation capacity, *Eucalyptus* and *Populus* have been widely used as the major hardwood lignocellulose feedstocks for the production of various bioproducts (Park *et al.* 2004). Generating bioproducts from these feedstocks requires efficient industrial bioprocessing approaches that can completely deconstruct cell wall biopolymers into simple constituents. Two approaches, dissolving pulp processing (DPP) and simultaneous saccharification and fermentation (SSF) are currently the most widely used bioprocessing approaches in the biorefinery industry (Hahn-Hägerdal and Palmqvist 2000; Stenberg *et al.* 2000; Galbe and Zacchi 2002; Mosier *et al.* 2005; den Haan *et al.* 2015). Besides these fermentation approaches, the value from woody biomass can also be extracted through non-fermentation methods such pyrolysis, synthetic gasification, and the traditional wood burning to generate electrical energy (Mohan *et al.* 2006; Kawamoto 2017; Kingsley Ogemdi 2019).

While some of these approaches are indeed useful in generating value-added bioproducts from woody biomass, they are expensive, inefficient, time-consuming, and often achieve low product yields (discussed in Langan *et al.* 2011). Therefore, a substantial improvement in industrial processing is necessary to reduce costs and maximize biomass yield. Such improvements require broad knowledge about the biosynthesis of secondary cell wall biopolymers and the structure thereof. Fortunately, advancements in functional genomics and biochemical analyses have provided useful information about the biosynthesis of the cell wall and how the structure thereof contribute to biomass recalcitrance.

The biosynthesis of a plant secondary cell wall involves the production and interconversion of sugar-nucleotide precursors which are then polymerized into larger secondary cell wall biopolymers and transported to the cell wall for final assembly into the secondary cell wall ultrastructure. The secondary cell wall thus represents a major metabolically demanding carbon sink in plants. Cellulose, the most abundant secondary cell wall biopolymer, is synthesized from uridine diphosphate (UDP)-Glucose (UDP-Glc) by plasma membrane-localized cellulose synthase (CESA) complexes (CSCs). In addition to CSCs, cortical microtubules (CMTs), MICROTUBULE-ASSOCIATED PROTEINS (MAPs), and several non-CESA proteins are involved in the cellulose biosynthetic process, some of these are covered in detail in this review.

To form a rigid and stable secondary cell wall surrounding phloem fibers and xylem vessels, cellulose microfibrils tightly interact with both hemicellulose and lignin in a very specific pattern (Kang *et al.* 2019). There are several different types of hemicellulose in and amongst different plant species. However, glucuronoxylan (or xylan) is the most common and abundant hemicellulose in hardwood trees (McCartney *et al.* 2006). As such, there have been numerous elegant studies focusing on mechanistically understanding xylan structure, biosynthesis, and the nature of its interaction with other secondary cell wall biopolymers (Fatland *et al.* 2002; Pena *et al.* 2007; Kotake *et al.* 2009; Lee *et al.*

2009, 2010, 2012a; c, 2014; Mortimer *et al.* 2010; Brown *et al.* 2011; Urbanowicz *et al.* 2012a; Bromley *et al.* 2013; Kang *et al.* 2019). Findings from these and other studies have shown that xylan backbone biosynthesis involves the activity of various glycosyltransferase family 43 (GT43) and GT47 members. Furthermore, membrane transporters, backbone modification enzymes, and enzymes involved in the biosynthesis of nucleotide-sugar precursors and donor substrates are required for complete xylan biosynthesis. As opposed to cellulose, xylan backbone side chains are heavily methylated and acetylated. These modifications influence the interaction of xylan with other secondary cell wall biopolymers, which ultimately contribute significantly to biomass recalcitrance, making bioprocessing very challenging (Ebringerová and Heinze 2000; Urbanowicz *et al.* 2012a; Bromley *et al.* 2013).

The concluding step in SCW biosynthesis is lignin deposition into the polysaccharide matrix (Pawar *et al.* 2013). Contrary to cellulose and xylan structure, which are comprised of simple sugars joined by identical chemical bonds, lignin is a complex three-dimensional polyphenolic polymer composed of monolignols joined with variable bonds (Boerjan *et al.* 2003). The presence of these different chemical bonds requires the use of several enzymes and different optimal conditions during bioprocessing (Zhong and Ye 2015; Kumar *et al.* 2016; Meents *et al.* 2018). As a result, the complicated deconstruction of these bonds leads to very inefficient and expensive bioprocessing options. Moreover, similar to xylan, lignin is acetylated and methylated at various positions (Hanson and Roje 2001; Amthor 2003; Boerjan *et al.* 2003; Del Río *et al.* 2007), these modifications and lignin ultrastructure impede hydrolytic enzymes from accessing other biopolymers, preventing efficient deconstruction of the secondary cell wall (Helle *et al.* 2003; Qing and Wyman 2011).

To counteract the abovementioned bioprocessing challenges, several chemical and enzymatic pre-treatment strategies are employed during bioprocessing. Unfortunately, these pre-treatments are costly

and the by-products from the treatments can contribute to additional forms of recalcitrance (Zanuttini et al. 2003; Busse-Wicher et al. 2016). Ultimately, these limitations necessitate the application of novel biotechnology approaches such as genetic engineering to improve bioprocessing efficiency. It has been shown that genetic manipulation of endogenous biosynthetic genes can yield plants with traits that are beneficial for bioprocessing (Urbanowicz et al. 2012a; Kuang et al. 2016; Macaya-Sanz et al. 2017; Yang et al. 2017). Moreover, modifying metabolic flux to secondary cell wall biosynthetic pathways promises yet another avenue that can be pursued to generate plants with favorable bioprocessing characteristics (Fan et al. 2017). Furthermore, recent advancements in systems genetics have helped identify putative genes potentially involved in biosynthesis and modification of secondary cell wall biopolymers (Mizrachi et al. 2015; Wierzbicki et al. 2019a). Using this systems genetics data, five candidate genes were selected for functional characterization in the current study: *XYLANASE 1* (*XYN1*), *ALPHA-L-ARABINOFURANOSIDASE 1* (*ARAF1*), *MANNANASE 2* (*MAN2*), *UDP-XYL SYNTHASE 6* (*UXS6*), and the *ENDOMEMBRANE PROTEIN 70* (*EMP70*). *XYN1*, *ARAF1* and *UXS6* are all predominately expressed in the xylem vessels, and both *XYN1* and *ARAF1* encode a group of glycoside hydrolase enzymes which degrade xylan chain into simple sugars (Chavez Montes et al. 2008; Zhong et al. 2017; Endo et al. 2019). While these genes are largely functionally uncharacterized, overexpression of *AtXYN1* result in enhanced xylem transport in mutants (Endo *et al.* 2019), while overexpression of *AtARAF1* display altered stem architecture (Chavez Montes *et al.* 2008). *AtUXS6* and *AtMAN2* T-DNA insertional mutants do not show any consistent phenotype, but simultaneous mutation of *USX6* and related genes; *UXS3* and *UXS5,* yields the typical irregular xylem phenotype (Kuang *et al.* 2016), while mutations in *AtMAN2*-related genes lead to late seed germination (Iglesias-Fernández *et al.* 2011b). Amongst these genes, there has not been any attempts at functionally testing *EMP70* in plants. More details on how these genes were selected will be discussed in subsequent sections.

Knowing which genes to target for functional characterization and the tool to employ for such characterization is the first step in successful gene functional testing study. Amongst other tools, the (CRISPR)/CRISPR-associated Protein9 (Cas9) (CRISPR/Cas9) has been shown to be effective and consistent in generating both mono- and biallelic mutants in woody plants (Fan *et al.* 2015; Zhou *et al.* 2015; Elorriaga *et al.* 2018), which can be essential in functional characterization of the candidate genes. Furthermore, CRISPR/Cas system has been extensively modified to generate sophisticated technologies that are capable of multiple applications. Amongst others, CRISPR has been repurposed for manipulation of gene expression in plants by coupling CRISPR machinery with deactivated Cas9 (dCas9) or dCas9 fusion proteins (Bikard *et al.* 2013; Qi *et al.* 2013a; Gilbert *et al.* 2015; Liu *et al.* 2016a). Recently, second-generation CRISPR-based gene editing approaches, namely, base editing and prime editing, have been developed. These approaches extend the CRISPR editing scope by a massive margin and thus, present a promise for the rapid and efficient genetic engineering of cell walls aimed at improvement of bioprocessing efficiency.

The focus of this dissertation will be on xylan biosynthesis; however, this literature review will introduce the other major secondary cell wall components to first set the stage for secondary cell wall biosynthesis. In this review, we investigate the long-standing and current understanding regarding the biosynthesis of SCW biopolymers, focusing on both the well-characterized and emerging biosynthetic genes and the products thereof. Next, the fundamental aspects of lignocellulose bioprocessing are described, concentrating on how secondary cell ultrastructure and modifications potentially affect bioprocessing efficiency. Some of the studies that have provided significant evidence that modification of cell wall biosynthetic genes could improve bioprocessing efficiency and biomass yield are also discussed. We further describe how CRISPR/Cas9 technology could play an essential role in engineering plants with beneficial biorefinery traits, providing a promise for efficient improvement of bioprocessing. Furthermore, new, and emerging CRISPR-based technologies such as base editing and

prime editing are briefly discussed. Finally, we briefly touch on CRISPR reagent delivery mechanisms as they are critical for the success of gene editing through CRISPR-based approaches.

1.2 The biosynthesis and structure of the secondary cell wall

Secondary cell wall biosynthesis is a complex process requiring the coordination of different biosynthetic and metabolic precursor transport pathways. Nucleotide diphosphate (NDP)-sugar substrates are synthesized and interconverted, then polymerized into larger biopolymers, which are ultimately transported to the cell wall for assembly into the final secondary cell wall structure (Bar-Peled and O'Neill 2011; Yin *et al.* 2011). Importantly, unlike primary cell walls, which are deposited in virtually all cell types, secondary cell walls are restricted to some specialized cell types, predominantly the fibers and treachery elements of the primary and secondary xylem (Turner *et al.* 2001; Ko *et al.* 2006). Their primary role in woody plants is to provide mechanical support, allowing efficient transport of water and nutrients necessary for normal plant growth and development (Bashline *et al.* 2014; Kumar *et al.* 2016; Meents *et al.* 2018). Furthermore, wood is evolutionary 'designed' to withstand decomposition, thereby enabling longevity of the woody plants. As the principal component of wood, the secondary cell wall represents a major source of lignocellulosic biomass. The main constituents of this biomass, namely cellulose, hemicellulose, and lignin, are broadly exploited for industrial production of many bioproducts including timber, paper, bioplastics, and biofuels. Secondary cell walls are therefore of industrial and bioeconomic significance, accordingly, significant efforts have been invested in understanding the structural properties of the secondary cell wall and the molecular basis of its biosynthesis, these aspects will be concisely discussed in the following sections.

1.2.1 Cellulose biosynthesis through cellulose synthase complexes

Amongst the three main secondary cell wall biopolymers, cellulose is the most abundant, comprising up to 50% of the secondary cell walls (Pauly and Keegstra 2008; Scheller and Ulvskov 2010; McNamara *et al.* 2015; Kumar *et al.* 2016). It is characterized by linear chains of para-crystalline β-1-4-linked glucan residues (Ha *et al.* 1998; Cosgrove 2005; Thomas *et al.* 2013), which coalesce tightly into nascent microfibrils that are held together by hydrogen bonds and Van der Waals forces (Barros *et al.* 2015; McNamara *et al.* 2015). The ultimate cellulose microfibrils possess high crystallinity and an extensive degree of polymerization, allowing cellulose to act as the main load-bearing unit in the cell walls of terrestrial plants (Turner *et al.* 2001; Ko *et al.* 2006).

The biosynthesis of cellulose chains and their assembly into microfibrils is mediated by a large dynamic plasma membrane-localized cellulose synthase (CESA) complex (CSC) (Delmer 1999; Doblin *et al.* 2002; Saxena and Brown 2005; **Figure 1.1)**. In plants, the functional CSC is a six-lobbed rosette structure composed of different CESA proteins (Giddings *et al.* 1980; Mueller and Malcolm Brown 1980). In *Arabidopsis,* CESA4, CESA7, and CESA8, all belonging to the GT2 family, are necessary for the formation of functional CSC in the Golgi, and subsequent secondary cell wall cellulose biosynthesis at the plasma membrane (Turner and Somerville 1997; Taylor 2002; Timmers *et al.* 2009). In the primary cell wall, the CSC heterotrimer involved in the production of cellulose is composed by CESA1, CESA3, and CESA6-like (Gardiner *et al.* 2003; Persson *et al.* 2007). In addition to localization in the PM and the Golgi like SCW CESAs, primary cell wall CESAs additionally localize to the small CESA compartments (SmaCCs), which assist in the delivery of CSCs along the microtubules (Crowell *et al.* 2009; Gutierrez *et al.* 2009). Co-immunoprecipitation (CoIP) studies of CESAs have shown that these proteins interact with each other within the same CSC, and when one of them is lost the CSC is no longer detected (Gardiner *et al.* 2003; Perochon *et al.* 2011). This suggests

that all three secondary cell wall CESAs are essential for the formation of a functional CSC, although their activities within the CSC are variable, with CESA4 being least active and CESA8 most active (Kumar *et al.* 2018). More importantly, homologs of these proteins have been characterized in other commercially significant plants such as *Populus,* cotton and rice (Arioli *et al.* 1998; Pear *et al.* 2002), emphasizing their necessity in cell wall biosynthesis in vascular plants.

Earlier phenotypic analyses of *CESA* mutants have greatly contributed to the understanding of the biological roles and significance of CESAs. Turner and Somerville (2002) reported that mutation in any of the secondary cell wall *CESAs* causes a 70% reduction in cellulose content in *Arabidopsis* inflorescence stems. Furthermore, mutants had significantly reduced secondary cell wall thickening subsequently leading to collapsed xylem vessels (Turner and Somerville 1997; Turner *et al.* 2001; Taylor 2002). Similarly, CRISPR/Cas9-induced knock-outs of *Populus trichocarpa* secondary cell wall CesAs; *PtrCesA4, 7A, 7B*, *8A* and *8B* have recently been generated and analyzed (Xu *et al.* 2021). Complete loss of these genes leads to morphological defects similar to those observed in *Arabidopsis* knock-outs (Xu *et al.* 2021), indicating that these *PtCesAs* play the same non-redundant roles as *Arabidopsis* homologs. Alongside phenotypic analyses, other studies have employed sophisticated cell imaging techniques to visualize the activity of CSCs during xylem development. For example, following the artificial induction of xylem cell fate in *Arabidopsis* epidermal cells, the substantial accumulation of CSCs into narrow plasma membrane domains was demonstrated (Watanabe *et al.* 2015). Moreover, CSC clusters have been observed to rapidly move in a highly coordinated manner along cortical microtubules (CMTs) adjacent to the forming secondary cell walls (Gardiner *et al.* 2003; Sugiyama *et al.* 2017).

The observed microtubule-guided CSC movement during cellulose biosynthesis requires the functional interaction between the underlying CMTs, CSC, and the PM (Gu *et al.* 2010). In both

primary and secondary cell walls, the cellulose synthase interacting 1 (CSI1)/POM2 protein has been shown to mediate this functional interaction (Gu *et al.* 2010; Bringmann *et al.* 2012; Li *et al.* 2012; Schneider *et al.* 2017). CSI1/POM2 binds with CMTs and regulate the CSC-CMT co-alignment (**Figure 1.1**). Additionally, CSI1-like protein, CSI3, plays an auxiliary but essential scaffolding role, possibly regulating the velocity of CSCs (Lei *et al.* 2013). However, CSC movement along CMTs exert an extensive force on the underlying CMTs, thereby altering microtubule network and dynamics (Fisher and Cyr 1998; Paredez *et al.* 2008). A MICROTUBULE-ASSOCIATED PROTEIN (MAP), CELLULOSE SYNTHASE-MICROTUBULE UNCOUPLING 1 (CMU1) has been shown to play an essential role in stabilizing CMTs and controlling the movement of CSC (Liu *et al.* 2016a). Thus, generating secondary cell wall cellulose requires functional PM-localized CSCs to rapidly move along CMU1-stabilized CMTs, but more importantly, CSCs require a sugar nucleotide precursor, uridine diphosphate (UDP)-Glucose (UDP-Glc), from which glycosyl residues are obtained to generate cellulose microfibrils.

1.2.2 Sugar-nucleotide precursor requirements for cellulose biosynthesis

As CSCs move along CMTs, the catalytic CSC subunits use the nucleotide sugar UDP-Glc as a substrate to transfer glycosyl residues to the nascent β−1-4 glucan chains which ultimately aggregate into cellulose microfibrils, forming part of the secondary cell wall (Amor *et al.* 1995; Fujii *et al.* 2010; Brill *et al.* 2011; Endler and Persson 2011). The biosynthesis of UDP-Glc required for cellulose biosynthesis is suggested to be facilitated by the glycosyltransferase enzyme SUCROSE SYNTHASE (SUSY) (Amor *et al.* 1995; **Figure 1.1**). Plasma membrane and cell wall-localized forms of SUSY have been identified and shown to directly associate with CSCs at least in wheat and tobacco (Albrecht and Mustroph 2003; Persia *et al.* 2008). The association of SUSY with CSC may allow direct distribution and partitioning of UDP-Glc to CESAs, which may be a useful strategy to avoid substrate

competition with other UDP-Glc-requiring pathways, thereby, favoring cellulose biosynthesis (Ross *et al.* 1991). The significance of SUSY in cellulose biosynthesis has been partially demonstrated through gene overexpression and downregulation studies. Overexpression of *SUSY* from cotton in *Populus* resulted in increased cellulose content, wood density, and cell wall thickness (Coleman *et al.* 2009). However, significant reduction of SUSY activity in hybrid aspen (*Populus tremula L.* x *tremuloides Michx.*) did not specifically result in a reduction of cellulose content, but an overall reduction in wood density caused by decreased lignin, cellulose, and hemicellulose content per wood volume (Gerber *et al.* 2014). Furthermore, Barratt *et al.* 2009 showed that *Arabidopsis SUSY1, SUSY2, SUSY3* and *4* quadruple mutants with SUSY activity only in the phloem are indistinguishable from the wild-type plants with respect to growth, cellulose content, and the overall cell wall structure. Taken together, these observations suggested that SUSY is not a mandatory requirement for cellulose biosynthesis, it may, however, play a more significant role in regulating the total carbon incorporated into secondary cell walls.

Two enzymes; CYTOSOLIC INVERTASE (cINV) and UDP-GLUCOSE PYROPHOSPHORYLASE (UGPase), have been suggested to be SUSY alternatives, playing vital roles in the biosynthesis of UDP-Glc required for cellulose biosynthesis (**Figure 1.1**). Convincing evidence supporting the role of cINV has been presented, for example, plants deficient in 2 isoforms of invertase (*cinv1/cinv2)* had severely reduced growth rates (Barratt *et al.* 2009). Similarly, imaging and biochemical data have demonstrated that *Arabidopsis cinv1cinv2* seedlings have impaired UDP-Glc biosynthesis accompanied by abnormal cellulose biosynthesis and altered cellulose organization (Barnes and Anderson 2018a). Both studies provided indispensable evidence supporting cINV as the primary active enzyme required for UDP-Glc biosynthesis in *Arabidopsis.*

Finally, while this review does not aim to cover all aspects of secondary cell wall cellulose biosynthesis, there are some noteworthy aspects of cellulose biosynthesis not covered in this review that may be valuable to the reader. For example, there are numerous other proteins, both well-characterized and emerging, which are implicated in cellulose biosynthesis. These include STELLO, PATROL1, and EXOCYST proteins all involved in CSC trafficking to the plasma membrane, and KORRIGAN, COBRA, and CTL which have established roles in the regulation of cellulose microfibril assembly and crystallinity. The precise roles of these and other proteins have been recently reviewed in detail in Zhong *et al.* 2019. Furthermore, it is important to note that the role of microtubules in cellulose biosynthesis is not limited to CSC tracking, but they also play a crucial role in both oriented cellulose deposition and patterned secondary cell wall deposition (reviewed in Derbyshire *et al.* 2015; Zhong and Ye 2015; Meents *et al.* 2018).

1.3 Overview of lignin biosynthesis and modifications

While this study focuses specifically on genes associated with xylan, the possibility that some of these genes' functions may affect lignin synthesis cannot be ruled out, it is therefore necessary to discuss some of the details pertaining to lignin structure and biosynthesis. The final step in secondary cell wall biosynthesis is lignification, which entails the secretion and deposition of the lignin polymer into the nascent cell wall to establish a functionally mature secondary cell wall that provides support for the plant even after the programmed cell death. Moreover, lignin facilitates water and solute transport in vascular tissues owing to its mechanical resistance and hydrophobic nature (Donaldson 2001; Boerjan *et al.* 2003). While both primary and secondary cell walls generally consist of cellulose and hemicellulose, lignin is largely restricted to the secondary cell walls, and alongside xylan, it is considered the primary biomass material responsible for biomass recalcitrance and associated lignocellulose deconstruction challenges (Martone *et al.* 2009; Espiñeira *et al.* 2011). Consequently,

there are limited lignin uses reported, however, secondary cell wall deconstruction processes produce a significant amount of lignin-rich waste products which can be converted into value-added products (Mentzen and Wurtele 2008). Furthermore, the significance and uses of lignin are growing as seen with its use as an essential additive in a wide array of industrial and agricultural applications. Therefore, a thorough understanding of lignin biosynthesis will facilitate the potential genetic manipulation of secondary cell wall ultrastructure for efficient lignocellulosic bioprocessing.

1.3.1 Biosynthesis of lignin monomers

Unlike cellulose and hemicellulose, which are both simple polysaccharides and structurally similar, lignin is a complex three-dimensional polyphenolic polymer produced through oxidative polymerization of the primary lignin precursors or monolignols: *p*-coumaryl, coniferyl and sinapyl alcohols (**Figure 1.1**). These primary lignin precursors then give rise to three lignin monomers, *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units, respectively (**Figure 1.1**). The proportion of these polymerized monolignols determines the overall lignin composition and differs significantly between different species and among different cells within the same species. For example, secondary cell walls in gymnosperms are predominantly composed of G lignin and undetectable levels of S and H, while eudicot secondary cell walls contain both G and S with small traces of H units (Vogel 2008; Bonawitz and Chapple 2010). Importantly, it has become clear that not only these classical monolignols are incorporated into lignin, small amounts of modified components of monolignol biosynthetic pathway and other naturally occurring non-traditional monolignols can also be incorporated into the lignin biopolymer (Chen *et al.* 2013a; Zhao *et al.* 2013).

Lignin biosynthesis commences with monolignol biosynthesis in the cytosol using L-phenylalanine generated from the Shikimate pathway as a starting material (**Figure 1.1**). Sequential hydroxylation,

methoxylation and side-chain activation/reduction of monolignols in the common phenylpropanoid pathway then give rise to the lignin monomers (Bonawitz and Chapple 2010; Vanholme *et al.* 2012a, 2013; Hao and Mohnen 2014). After biosynthesis in the cytosol, monolignols are transported to the developing secondary cell wall where they are incorporated into lignin polymer as S, G and H units (**Figure 1.1**), however, there is currently no consensus on the exact monolignol transport mechanism across the plasma membrane. Nevertheless, studies conducted in *Arabidopsis* and *Populus* have suggested both passive and active transport mechanisms for monolignols (Miao and Liu 2010; Alejandro *et al.* 2012; Sibout and Höfte 2012; **Figure 1.1**).

1.3.2 The transport and polymerization of lignin monomers

The active transport process that has received significant attention is the ATP-binding cassette (ABC)-like transport, which is suggested to involve ATP-dependent membrane transporters for export of monolignols across the plasma membrane (Miao and Liu 2010; Kaneda *et al.* 2011; Alejandro *et al.* 2012). Supporting the role of this process, heterologous expression of one *Arabidopsis* ABC-like gene, *AtABCG29*, in yeast lead to increased *p*-coumaryl transport, furthermore, *Arabidopsis* mutants lacking *AtABCG29* exhibited growth weakness on *p*-coumaryl with a modest decrease in lignin content (Alejandro *et al.* 2012). The modest decrease in lignin content observed has been attributed to the functional redundancy amongst several identified *ABCG* genes (Takeuchi *et al.* 2018). To further test whether ABC transporters were necessary for monolignol transport, membrane vesicles from *Arabidopsis* PM were treated by vanadate, which inhibits ATPase, this led to inhibition of coniferyl transport, similarly, coniferyl transport was inhibited when ATP was excluded (Miao and Liu 2010). These observations provided an insight into the possible role of *Arabidopsis* ABC transporters at least for *p*-coumaryl and coniferyl alcohol transport. Contrary to these observations, biochemical analysis of developing xylem from poplar (*Populus sieboldii*) and hybrid poplar (*Populus sieboldii* x *Populus*

grandidentata) did not find any evidence for the involvement of ABC transporters, instead, a vacuole-localized transporter with high affinity to monolignols was identified (Tsuyama *et al.* 2013), suggesting that different transport processes may be operating in different species and possibly different cell types within species (Barros *et al.* 2015).

Since the role of active transport was only supported in *Arabidopsis* and limited to *p*-coumaryl and coniferyl transport thus far, a recent study has presented data proposing that most lignin-related compounds (LRCs) may alternatively cross the lipid bilayer via passive transport (Vermaas *et al.* 2019). Analyses of membrane partitioning and permeability of 69 LRCs using molecular dynamics simulations have suggested that only carboxylated or glycosylated LRCs may require membrane transporters, while most LRCs, including monolignols, can passively permeate the plasma membrane (Vermaas *et al.* 2019). Convincingly, prior studies have demonstrated that the abundance of immobile lignin-polymerizing peroxidases and laccases at the secondary cell wall may create an energetically favorable gradient for monolignols to passively cross the membrane (Kaneda *et al.* 2008; Smith *et al.* 2013; Jaini *et al.* 2017). Furthermore, studies using liquid chromatography-tandem mass spectrometry and radiolabeling demonstrated that monolignols are rapidly incorporated into the cell wall without significant accumulation in the cytosol, indicating that monolignols may indeed rapidly diffuse passively across the plasma membrane (Boija *et al.* 2007; Dima *et al.* 2015; Yi Chou *et al.* 2018).

Once at the secondary cell wall, monolignols are oxidized by laccases and class III peroxidases to form the ultimate polymeric lignin structure which interacts with hemicellulose and cellulose in a very specific pattern: contrary to the previous models which suggested that xylan tightly interacts with cellulose, and the resulting xylan-cellulose structure is coated with lignin (Himmel *et al.* 2007; Kabel *et al.* 2007; Köhnke *et al.* 2011; Terashima *et al.* 2012; Martínez-Abad *et al.* 2017), Kang *et al.* 2019 have shown that there are limited xylan-cellulose and cellulose-lignin interactions. Instead, lignin

mainly binds xylan most likely through electrostatic interactions, furthermore, lignin and cellulose are largely spaced and joined together by xylan (Kang *et al.* 2019). The overall impact of these secondary cell wall biopolymer interactions on lignocellulose bioprocessing will be discussed in the next sections.

1.3.3 Lignin methylation and acetylation

Lignin monomers (G, S and H units) are derived from a variety of methylated or non-methylated monolignols. G is derived from coniferyl with one methyl-group at the C3 position, S is derived from synapyl consisting of methoxy-group at the C3 and C5 position, while H is produced from non-methylated *p*-coumaryl (Boerjan *et al.* 2003; **Figure 1.1**). Methylation of lignin precursors is catalyzed by CAFFEOYL COA *O*-METHYLTRANSFERASE (CCoAOMT) and CAFFEIC ACID *O*-METHYLTRANSFERASE (COMT) at C3 and C5, respectively (Boerjan *et al.* 2003; **Figure 1.1**). Both these methyltransferases use the methyl group from SAM for transmethylation of lignin, consequently, lignin biosynthesis in secondary cell wall tissues is associated with a large supply and efficient recycling of SAM (Hanson and Roje 2001; Amthor 2003). Several mutation studies have provided some compelling evidence supporting the importance of SAM in lignin biosynthesis. For example, the mutation in one of the *Arabidopsis* SAM synthetases, *S*-ADENOSYL METHIONINE SYNTHETASE 3 (AdoMetS3), results in a reduction in SAM synthetase activity, SAM pools, and lignin content (Shen *et al.* 2002). Moreover, a mutation in genes involved in the production of 5-methyltetrahydrofolate (a methyl donor for the production of methionine) from homocysteine, leads to reduced lignin levels in *Arabidopsis* and maize (Tang *et al.* 2014; Li *et al.* 2015; Srivastava *et al.* 2015). More importantly, it has been demonstrated in *Arabidopsis* stems that targeted expression of SAM hydrolases in secondary cell wall-forming tissues reduces the SAM pool, and the relative content of the dimethylated S lignin units, resulting in the enrichment of glucose content and reduction of cell wall recalcitrance (Eudes *et al.* 2016). The latter is particularly important as it showed that targeting

the biosynthetic pathways of precursor metabolites may reduce recalcitrance while enriching the production of monosaccharides, an important aspect of efficient bioprocessing.

In addition to methylation, some lignin is acetylated at the γ-Carbon position of their S and G monomers (Del Río *et al.* 2007). This acetylation possibly affects the lignin composition and its interaction with other secondary cell wall polymers (reviewed in Pawar *et al.* 2013). However, only a limited number of hardwood species have been reported to be acetylated in the xylem with lignin acetylation varying between 1 and 50%, while there have not been any reports for lignin acetylation in softwoods (Pawar *et al.* 2013). The exact function and consequences of lignin acetylation and variation thereof are thus unclear; though it is possible that lignin acetylation may hamper lignocellulose bioprocessing, like the presence of *O*-acetyl groups in xylan backbone, this xylan modification, and others, will be discussed in detail in the next section.

1.4 Xylan backbone biosynthesis and decorations

Hemicellulose represents 10-40% of the secondary cell walls, with xylan being the most abundant hemicellulose component in hardwood trees (Darvill *et al.* 1980; McCartney *et al.* 2006). Like cellulose, xylan contributes significantly to normal plant growth and development, however, in contrast to cellulose, the structure and biosynthesis of xylan is significantly complex. Commercially, xylan-containing secondary cell walls are the most abundant biomass repository on earth. Therefore, this abundant biomass provides a considerable potential to be used as a feedstock to produce large-scale biofuels and other value-added biobased products essential for a sustainable bioeconomy. Nonetheless, xylan biosynthesis is complex, requiring a plethora of metabolic precursors and enzymes. Moreover, the xylan backbone can consist of several chemical modifications, such as methylation and acetylation, and side-chain substitutions. These modifications impact xylan's solubility and its

interaction with other secondary cell wall biopolymers, making it a challenge to extract xylan from the lignocellulosic biomass (reviewed in Pawar *et al.* 2013; Hao and Mohnen 2014; Smith *et al.* 2017). Additionally, xylan modifications negatively impact the conversion of other secondary cell wall biopolymers into fermentable sugars, which is an important and expensive step in industrial processing (York and O'Neill 2008). Therefore, understanding xylan backbone biosynthesis, the overall xylan ultrastructure, and its associated chemical modifications will open doors for the application of novel approaches aimed at improving bioprocessing efficiency.

1.4.1 Xylan backbone biosynthesis

Xylan is composed of β–(1-4)-linked D-xylosyl (Xyl) residues which can be appended with different chemical modifications depending on the species (Teleman *et al.* 2002; Pena *et al.* 2007). Unlike cellulose, which is synthesized at the plasma membrane by a well-defined protein complex, xylan is synthesized by a variety of enzymes within the Golgi lumen (**Figure 1.1**). Like cellulose biosynthesis, UDP-Glc acts as a substrate for all nucleotide-sugar precursors required for xylan biosynthesis. UDP-Glc is oxidized to UDP-D-glucuronate (UDP-GlcA) in the cytosol by UDP-GLC 6-DEHYDROGENASE (UGDH) ((Raimund and Oliver 1996). UDP-GlcA is then interconverted into UDP-D-xylose (UDP-Xyl) through the action of UDP-XYL SYNTHASES (UXSs; (Kotake *et al.* 2009; Bar-Peled and O'Neill 2011; Rautengarten *et al.* 2011). Interestingly, genetic analyses of *Arabidopsis uxs* mutants have demonstrated that a reduction in xylan content only occurs when genes encoding the cytosolic UXSs, and not Golgi-localized UXSs, are mutated, indicating that cytosolic UXSs are required for UDP-Xyl biosynthesis (Kuang *et al.* 2016; Zhong *et al.* 2017). Following biosynthesis at the cytosol, UDP-Xyl is transported into the Golgi primarily by UDP-XYL TRANSPORTER 1 (UXT1; (Ebert *et al.* 2015; Zhao *et al.* 2018). Subsequent xylan backbone biosynthesis involves the transfer of xylosyl residues from UDP-Xyl into the growing xylan chain.

The overall biosynthesis of the xylan backbone involves the activity of two GT43 members, IRREGULAR XYLEM (IRX) 9 AND IRX14 (Brown *et al.* 2005, 2007; Pena *et al.* 2007; Urbanowicz *et al.* 2014) and a GT47 member, IRX10 (Jensen *et al.* 2014; Urbanowicz *et al.* 2014). In *Asparagus*, where these enzymes are highly active, IRX9, IRX10 and IRX14 are shown to form a xylan synthase complex (XSC) within the Golgi apparatus, this complex is responsible for xylan xylosyltransferase activity in planta (Zeng *et al.* 2016). A similar complex has been shown to exist in *Arabidopsis thaliana* (Reviewed in Smith *et al.* 2017; **Figure 1.1**), within the XSC, IRX10 is a xylosyltransferase necessary for the extension of the xylan biopolymer, while IRX9 and IRX14 play structural roles and are involved in the elongation of the xylan backbone. IRX9, IRX14, and IRX10 all have partial functional redundant paralogs, IRX9-L, IRX14-L, and IRX10-L, respectively. Single mutants in either one of the paralog pair do not all display a severe phenotype (Brown *et al.* 2009; Chen *et al.* 2013b). However, a mutation in both homologs of either *IRX9*, *IRX14*, or *IRX10* results in a shorter xylan chain and severe growth defects (Brown *et al.* 2009; Wu *et al.* 2009; Keppler and Showalter 2010; Lee *et al.* 2010). Importantly, mutation of *IRX9* or *IRX14* alone is sufficient to cause a significant reduction of xylosyltransferase activity, xylan chain length, and the overall xylan content (Lee *et al.* 2012a). Mutants also have severely reduced secondary cell wall thickening and collapsed xylem vessels (Lee *et al.* 2012c). Moreover, *IRX9* could not rescue the *irx14* mutant phenotype and likewise, *IRX14* failed to rescue the *irx9* phenotype (Brown *et al.* 2007; Pena *et al.* 2007; Lee *et al.* 2010; Wu *et al.* 2010). Interestingly, a dominant negative approach targeting IRX10 in *Arabidopsis* yielded some of the typical phenotypes observed when xylan biosynthetic genes are altered (Brandon *et al.* 2020). Here, overexpression of mutated forms of IRX10 was successful in suppressing xylan biosynthesis. Transgenic plants showed significantly reduced height indicating suppression of xylan production (Brandon *et al.* 2020). Furthermore, xylose levels were reduced in the overexpressing lines coupled with thinner secondary cell walls (Brandon *et al.* 2020). Altogether, these studies suggested that *IRX10, IRX9* and *IRX14* are functionally non-redundant and play major independent roles in xylan backbone biosynthesis.

Two functionally non-redundant GT43 members have also been identified in *Populus* and rice (Lee *et al.* 2012d, 2014; Chiniquy *et al.* 2013). Furthermore, a recent study in *Brachypodium distachyon* generated CRISPR knock-out and artificial microRNA (amiRNA) knockdown lines of the *IRX14* ortholog, *BdGT43B2* (Petrik *et al.* 2020). Amongst other phenotypes, a significant reduction in xylan and a small reduction in xylose was observed in the mutant lines, respectively, suggesting that *BdGT43B2* may be playing a similar role to *AtIRX14* in *Brachypodium* (Petrik *et al.* 2020). Similarly, an abundantly expressed GT43 wheat gene, *TaIRX9b*, was capable of complementing the *Arabidopsis IRX9* knock-out mutant, further, *TaIRX9b* knock-out mutant has reduced arabinoxylan amount and chain length (Pellny *et al.* 2020). Thus, *TaIRX9b* may have the same function in xylan biosynthesis as *AtIRX9* in wheat. So, the existence of GT43 xylosyltransferase orthologs in poplar, rice, grass, and wheat may indicate that the requirement of non-redundant glycosyl transferase-encoding genes for xylan biosynthesis is conserved in vascular plants.

1.4.2 Xylan side-chain substitution

In the secondary cell wall of eudicots, the xylan backbone is often substituted with D-glucuronic acid (GlcA), giving rise to glucuronoxylan (GX; Ebringerová and Heinze 2000). This modification involves the transfer of GlcA residues from UDP glucuronic acid (UDP-GlcA) to the xylosyl backbone. GlcA side chains may further be methylated at *O*-4 to yield α-1,2-linked 4–*O*–methyl glucuronic acid (4-*O*-MeGlcA) (Ebringerová and Heinze 2000). In *Arabidopsis* secondary cell walls, the addition of GlcA to xylan backbone involves three Golgi-localized glucanosyltransferases; GLUCURONIC ACID SUBSTITUTION OF XYLAN 1 (GUX1), GUX2 and GUX3, all belonging to the GT8 family (Mortimer *et al.* 2010; Lee *et al.* 2012a; Bromley *et al.* 2013). While mutation of a single gene encoding any one of these enzymes does not affect plant growth or secondary cell wall thickness, simultaneous mutation of all three genes leads to a complete loss of GlcA substitutions, reduced secondary cell wall

thickness and deformed vessels (Lee *et al.* 2012a), suggesting that xylan GlcA modification is essential for normal secondary cell wall development and integrity. Moreover, subsequent detailed analysis of several *gux* mutant lines indicated that different GUX enzymes contribute to specific GlcA substitution patterns. GUX1 adds GlcA side chains to evenly spaced xylosyl residues between 6-26 residues apart, while GUX2 adds GlcA to both evenly and oddly spaced xylose residues (Bromley *et al.* 2013). Importantly, the differences in frequency of GlcA addition leads to the biosynthesis of two distinct xylan domains, which may be spatially separated but within the same xylan molecule (Bromley *et al.* 2013). While the significance of these different domains as they relate to functionality is still unclear, Bromley *et al.* 2013 suggested that the domains may affect xylan's interaction with cellulose to form a mature secondary cell wall and therefore potentially affect biomass recalcitrance.

1.4.3 Metabolic requirements and synthesis of xylan backbone modifications

As already introduced, a significant proportion of xylan's GlcA side chains are methylated at *O*-4 (Urbanowicz *et al.* 2012a; Yuan *et al.* 2014), the proportion of these 4-*O*-methylated GlcA moieties varies with species. For example, in *Arabidopsis* stems about 60-80% of total GlcA substitutions are methylated while all GlcA chains are methylated in *Populus tremula* x *P. alba* (Urbanowicz *et al.* 2012a). A reduction in GX's 4-*O*-MeGlcA content leads to increased xylan release during saccharification, suggesting that GlcA methylation contribute to the overall structure and integrity of the secondary cell wall (Urbanowicz *et al.* 2012a). Three Golgi-localized glucuronoxylan methyltransferases; GLUCURONOXYLAN METHYLTRANSFERASE 1 (GXM1), GXM2, and GXM3 are necessary for the methylation of the GlcA side chains (Lee *et al.* 2012b; Urbanowicz *et al.* 2012a; **Figure 1.1**). These enzymes are characterized by the presence of Domain of Unknown Function 579 (DUF579) which has been suggested to possess cation-dependent polysaccharide specific *O*-methyltransferase activity (Kopycki *et al.* 2008; Urbanowicz *et al.* 2012a; Smith *et al.* 2020). Studies

on *Arabidopsis* T-DNA mutants show that defects in this domain lead to a 75% reduction in *O*-methylated GlcA side chains (Urbanowicz *et al.* 2012a). This reduction in *O*-methylated GlcA side chains signifies the significance of the domain in GlcA methylation. Furthermore, simultaneous mutation of *GXM1, GXM2,* and *GXM3* lead to complete loss of GlcA methylation while only 13% of GlcA residues are methylated in *gxm2 gxm3* double mutants (Yuan *et al.* 2014). Predictably, overexpression of *GXM1* and *GXM3* in *Arabidopsis* stems resulted in an increase in the proportion of 4-*O*-methylated residues from 60% observed in wild-type to 90% (Yuan *et al.* 2014), further indicating the significance of GXMs in GlcA methylation.

The mechanism of GlcA methylation involves the transfer of a methyl group from *S*-adenosylmethionine (SAM) onto the GlcA residues on the xylan backbone (Shen *et al.* 2002). It is known that SAM is an essential intermediate in one-carbon metabolism, not only serving as a universal methyl group donor for the methylation of the xylan backbone but also for monolignols (Lee *et al.* 2012b; Urbanowicz *et al.* 2012a; Sauter *et al.* 2013; Wang *et al.* 2016). Moreover, SAM acts as a precursor in the biosynthesis of various metabolites such as phytosiderophores and polyamines (Roje 2006). While this is a very specific example of the methylation from SAM, methylations are common in all the secondary metabolic pathways. SAM is synthesized from methionine in the cytosol and evidence indicates that SAM transporters are required for its transport into the Golgi Apparatus for the methylation of homogalacturonan, and possibly xylan (Ibar and Orellana 2007; Sauter *et al.* 2013; **Figure 1.1**). Notably, SAM biosynthesis relies on multiple essential metabolites, which are products of several biosynthetic pathways, therefore, tight regulation of SAM biosynthesis and transport is important as it potentially affects the biosynthesis of xylan and other cell wall polymers (reviewed in Wierzbicki *et al.* 2019b). In addition to methylation, a large proportion of xylosyl residues on the xylan backbone are acetylated at either oxygen 2 (*O*-2) and/or *O*-3 (Capek *et al.* 2002; Teleman *et al.* 2002; Gonçalves *et al.* 2008). *O*-Acetyl moieties on GX largely influence the polymer-polymer interactions,

which impacts secondary cell architecture and lignocellulose bioprocessing (Busse-Wicher *et al.* 2014). More importantly, acetylation alone greatly contributes to biomass recalcitrance, which demands the application of de-acetylation steps during industrial processing to improve precipitation of biopolymers (Zanuttini *et al.* 2005; Konn *et al.* 2006; Busse-Wicher *et al.* 2016). More details on how the overall secondary cell wall structure specifically contributes to biomass recalcitrance and how genetic engineering can be utilized to alleviate these challenges will be discussed in the last section of the review.

Acetyl-CoA acts as the donor substrate for acetyl groups required for xylan acetylation (Pauly and Scheller 2000; Fatland *et al.* 2005; Oliver *et al.* 2009), it is also a vital metabolite required for fatty acid biosynthesis and acetylation of proteins (Fatland *et al.* 2005). Since Acetyl-CoA plays numerous essential roles, it is compartmentalized between plastids and cytosol to ensure sufficient supply between different competing pathways. Amongst other sources, a significant proportion of acetyl-CoA is generated from citrate by cytosolic ATP-CITRATE LYASE (ACL; Fatland *et al.* 2005). Plants with just moderately reduced ACL activity have miniaturized organs, smaller cells, and aberrant plastid morphology (Fatland *et al.* 2002, 2005), suggesting that ACL is required for normal plant growth and that no other source of acetyl-CoA can compensate for ACL-derived acetyl-CoA (Fatland *et al.* 2002, 2005). Several potential transporters are then required to channel the cytosolic acetyl-CoA into the Golgi for xylan acetylation (**Figure 1.1**), however, such transporters have not yet been elucidated.

Two protein families, namely, REDUCE WALL ACETYLATION (RWA) and the TRICHOME-BIREFRINGENCE-like (TBL) family, are involved in the acetylation of xylan in plants (Gille *et al.* 2011; Manabe *et al.* 2011). Three *Arabidopsis* RWA members, RWA1, RWA3 and RWA4 redundantly regulate secondary cell wall acetylation (Lee *et al.* 2011), while RWA2 is involved in acetylation of pectin and xyloglucan (Manabe *et al.* 2011). Concurrent mutation in all four *Arabidopsis*

RWA genes leads to a 42% reduction in xylan *O*-acetylation, a 40% reduction in acetyl content in the stem, reduced secondary cell wall thickening, and collapsed xylem vessels (Lee *et al.* 2011). These characteristics represent a common phenotype often observed when essential secondary cell wall genes are mutated. Therefore, it is easy to assume that the *RWA* genes play an essential role in xylan acetylation and that acetylation is required for normal secondary cell wall development. Similar to observations in *Arabidopsis*, RNAi knockdown of *RWA* orthologs in *Populus* resulted in reduced xylan acetylation and increased glucose and xylose yields following enzymatic hydrolysis (Pawar *et al.* 2017b).

A biochemical and homology study has shown that although RWAs possess several transmembrane domains, they lack the putative XYLAN *O*-ACETYLTRANSFERASE (XOAT) domain required for the transfer of acetyl groups to the xylosyl residues (Gille *et al.* 2011). It has been consequently proposed that RWAs may be putative acetyl donor transporters, channeling acetyl-CoA from the cytosol to the Golgi lumen (Gu *et al.* 2010; Lee *et al.* 2011; Rennie and Scheller 2014). In addition to RWAs, TBL29/ESKIMO 1 (ESK1) is an essential component in xylan acetylation in *Arabidopsis* (Xiong *et al.* 2013a). TBL29 belongs to the plant-specific DUF231 protein family, which contains several other members involved in the *O*-acetylation of other plant cell wall polymers (Gille and Pauly 2012). The *Arabidopsis tbl29* mutant causes the typical deformed xylem vessel phenotype and a reduction in secondary cell wall thickening, similar to the *RWA* quadruple mutant phenotype, however, no reduction in xylan content is observed in the *tbl29* mutant (Lee *et al.* 2011; Gille and Pauly 2012; Xiong *et al.* 2013a). Thus, the reduced xylan acetylation in the *tbl29* mutant is sufficient to cause the irregular xylem phenotype, which emphasizes the significant of xylan acetylation in xylem development.

Contrary to the *RWA* quadruple mutant, biochemical analysis of *tbl29* xylan acetyltransferase activity has shown an approximate 60% reduction in 2-*O* and 3-*O* xylan monoacetylation, and a significant decrease in xylan acetyltransferase activity (Xiong *et al.* 2013a; Yuan *et al.* 2013). These data indicated that TBL29 is a putative acetyltransferase, specifically involved 2-*O* and 3-*O* xylan monoacetylation. Indeed, through a combination of computational simulations, mutagenesis, and enzyme activity assays, it has been shown that xylan acetylation is accomplished through TBL29-catalyzed 2-*O*-acetylation followed by migration of acetyl to *O*-3 (Urbanowicz *et al.* 2014; Lunin *et al.* 2020). Thus, supporting TBL29 role as an *O*-acetyltransferase specifically catalyzing 2-*O* and 3-*O* xylan monoacetylation. Interestingly, genetic studies in *Arabidopsis* have proposed that acetyl may be transferred from acetyl-CoA to *O*-acetyltransferases by an intermediate protein called ALTERED XYLOGLUCAN 9 (AXY9; Schultink *et al.* 2015; Zhong *et al.* 2018). Although mutation of *AXY9* leads to reduced xyloglucan and xylan acetylation, it remains unknown what the exact biochemical function of AXY9 is (Schultink *et al.* 2015).

Clearly, xylan biosynthesis is complex, involving various enzymes for backbone elongation, substitutions, and side-chain modifications. Furthermore, xylan biosynthesis requires a sufficient supply of nucleotide sugar precursors and metabolite (acetyl and methyl group) donors for backbone biosynthesis and modifications which is achieved by the action of many uncharacterized substrate biosynthetic enzymes and membrane transporters. Moreover, it remains unknown how xylan is targeted to the specific plasma membrane domains to yield the patterned secondary cell wall observed in many cell types such as protoxylem in treachery elements. Therefore, future studies aimed at characterizing genes encoding precursor biosynthetic enzymes and membrane transporters will be highly valuable, and the understanding of precise mechanisms involved in xylan secretion from the Golgi during secondary cell wall deposition will benefit from current advancements in microscopy and imaging technologies, amongst other approaches.

1.5 Lignocellulose bioprocessing

As lignocellulosic biomass represents the most abundant source of renewable materials on earth, it is commonly used as a feedstock for pulp and papermaking as well as biorefinery applications that produce valuable bioproducts such as packaging materials and biofuels (Plomion *et al.* 2001; Carroll and Somerville 2008; Pauly and Keegstra 2008). Moreover, lignocellulose hydrolysates contain a variety of sugars, such as glucose, which can further serve as alternative sources for the production of bio-based products (Menon and Rao 2012). This wide variety of lignocellulose applications requires efficient separation and purification of lignocellulose into its constituents (Klemm *et al.* 2005; Restolho *et al.* 2009; Nasatto *et al.* 2015). However, as mentioned previously, the structure and modifications on the different secondary cell wall biopolymers make lignocellulose resistant to industrial bioprocessing. Subsequently, recalcitrance contributes to a substantially higher cost of biomass conversion coupled with low product quality and yields (Langan *et al.* 2011). Therefore, a greater understanding of secondary cell wall biosynthesis will allow the application of indispensable biotechnology approaches to reduce recalcitrance and improve biomass yield. Previous sections of the review provided information on the current understanding of how secondary cell biopolymers are synthesized and modified, in this section, the potential impact of this secondary cell wall composition, structure and modifications on current industrial processing approaches and how biotechnology could aid in generating favorable bioconversion characteristics will be discussed.

1.5.1 Overview of industrial bioprocessing techniques

Several industrial processing approaches are used for the deconstruction of wood in the biorefinery industry; these includes the non-fermentation approaches such as pyrolysis and synthetic gasification. In addition, value-added products from wood can be extracted through fermentation methods such as

the simultaneous saccharification and fermentation (SSF) and the dissolving pulp processing (DPP). Pyrolysis is the initial step in most processes that convert wood into liquid fuels, it involves slow heating at 480 °C (in the absence of oxygen) to thermally degrade the wood biomass into a liquid fraction, a gaseous fraction, and carbon-containing solid fraction (González-Vila et al. 1999; Del Río et al. 2005; Mohan et al. 2006). Some major products that are generated from pyrolysis are charcoal and bio-oils, the latter of which has several advantages to fossil fuels such as reduced Nitric oxide containing emissions and their ability to generate carbon dioxide credits (Kumar et al. 2010). Like pyrolysis, gasification involves combustion at high temperatures with limited oxygen, producing combustible gases (hydrogen and carbon dioxide) and hydrocarbons. The produced gases can be converted into syngas which can be converted into synthetic biodiesel through catalytic conversion (Senneca 2007; Yung *et al.* 2009).

Between these non-fermentation bioconversion approaches and the fermentation approaches, large-scale biorefinery industry often rely on fermentation strategies such as SSF and DP. In SSF, secondary cell wall biopolymers are separated through pre-treatment and enzymatic hydrolysis to yield monosaccharides (Spindler *et al.* 1990; Wyman *et al.* 1992; Wu and Lee 1998). These monosaccharides are simultaneously fermented by microbes to produce ethanol (Ragauskas *et al.* 2014; den Haan *et al.* 2015; Chen *et al.* 2017a). SSF is thus a two-step process conducted in a single vessel (Wingren *et al.* 2003). The simultaneous occurrence of hydrolysis and fermentation, as opposed to fermentation after hydrolysis, ensures attenuation of end-product enzyme inhibition, thus ensuring a high ethanol yield at the lowest possible cost (Wingren *et al.* 2003). There is, however, a principal drawback associated with the nature of SSF, namely, the need for optimal conditions (e.g. temperature and pH) for both enzymatic hydrolysis and fermentation, and the challenge to recycle fermenting microbes and enzymes (Hahn-Hägerdal and Palmqvist 2000; Stenberg *et al.* 2000; Galbe and Zacchi 2002; Mosier *et al.* 2005; den Haan *et al.* 2015).

In contrast to SSF, DPP primarily uses either highly acidic (Sulphite process) or basic liquors (Kraft process) to separate xylan and lignin from cellulose to yield highly pure cellulose (Henriksson *et al.* 2005; Grönqvist *et al.* 2014; Gao and Chen 2017). Furthermore, an acid pre-hydrolysis step is included to remove hemicelluloses. While the main aim of SSF is to produce ethanol, the resulting cellulose-rich (>90% cellulose) pulp from DPP is used to produce textiles, paper, aerogels, resin-impregnated fibers, and other innovative cellulose applications. To obtain these bioproducts from SSF and DPP, a stringent separation of lignocelluloses, inhibitors and contaminants into individual processing sub-pathways that could efficiently generate useful end bioproducts is required (Vardon *et al.* 2016). However, lignocellulose ultrastructure, specifically the presence of lignin, and xylan modifications, hinders this separation process, necessitating some form of lignocellulose pre-treatments in the biorefinery process.

1.5.2 The impact of secondary cell wall structure and modifications on bioprocessing efficiency

As already discussed, lignin and cell wall polysaccharides interact tightly to yield a rigid structurally and functionally mature secondary cell wall necessary for providing mechanical support to plants cells (Zhong and Ye 2015). However, these complex interactions and the structure of biopolymers at the mature secondary cell wall also make industrial bioprocessing significantly challenging. For example, xylan is composed mainly of simple pentose sugars, the breakdown thereof which requires complex metabolic pathways distinct from those required for bioconversion of hexoses that make up cellulose (Chandel *et al.* 2011). Amongst other challenges, the fermentative microbial strains necessary for pentose bioprocessing are often lacking in most biorefinery industries (Chandel *et al.* 2011). Furthermore, the complex interaction between the xylan-lignin complex and cellulose makes efficient xylan separation from lignocellulose a challenge for both SSF and DPP (Pereira *et al.* 2017). Similarly, the presence and the spacing between xylan's GlcA decorations have a profound impact on

recalcitrance and bioprocessing as they affect how cellulose interact with xylan (Bromley *et al.* 2013; Grantham *et al.* 2017; Pereira *et al.* 2017).

Alongside its side-chain modifications, xylan's largest contribution to recalcitrance is attributed to the presence of *O*-acetyl groups on its backbone. During bioprocessing, the release of *O*-acetyl contributes to a significant reduction in pH which may inhibit the activity of fermentative microbes as well as downstream enzymes in the bioconversion pathway (Basen *et al.* 2014). Furthermore, the presence of *O*-acetyl groups create steric hindrance for the binding of hydrolytic enzymes to the biopolymers, preventing efficient separation of polymers and xylan removal in the DPP process, this subsequently reduces the yield and purity of cellulose (Helle *et al.* 2003; Qing and Wyman 2011; Chong *et al.* 2015). Finally, xylan acetylation and methylation patterns alter lignin composition and solubility (Grantham *et al.* 2017), thus affecting the overall bioprocessing efficiency (Kang *et al.* 2019). Therefore, the ability to alter xylan structure perhaps through reduction of the extent of backbone substitutions and side-chain modifications may lessen lignocellulose deconstruction challenges and the associated costs.

Like xylan, lignin structure and modifications contribute significantly to biomass recalcitrance (Himmel *et al.* 2007; Zeng *et al.* 2017). As discussed before, lignin is a complex polymer composed of polyphenolic monomer units that are cross-linked through different chemical bonds. As a result, lignin is highly resistant to biodegradation, often requiring a range of different enzymes for its complete degradation. Lignin also tightly binds xylan, while lignin and cellulose are interconnected through xylan to yield a tightly packed secondary cell wall architecture (Kang *et al.* 2019). Taken together, the complexity of lignin structure, and the tight lignin-polysaccharide interactions impede industrial processing of cellulose and xylan by physically restricting the surface area available for enzymatic penetration and activity (Ming-Ju Chen, Kreuter 1996; Pan *et al.* 2005; Kang *et al.* 2019). Such resistance in bioprocessing due to lignin necessitates the initial removal of lignin during

bioprocessing, a fact that has been proven to be an important step for paper production and the conversion of biomass into biofuels (Qing and Wyman 2011). In addition to the presence of lignin, the specific composition of lignin in hardwood contributes to recalcitrance. As already discussed, the lignin present in hardwood tree species is mainly composed of syringyl (S) and guaiacyl (G) units. The ratio of these subunits (S/G ratio) correlates with the ease of de-lignification (Chen and Dixon 2007; Nunes *et al.* 2010; Studer *et al.* 2011); G lignin is highly resistant to hydrolysis, while S lignin can be easily hydrolyzed through DPP. Therefore, lignocellulose with a higher S/G ratio is easier to delignify and contributes to higher pulp yield (González-Vila *et al.* 1999; Del Río *et al.* 2005). Thus, total lignin content and the S/G ratio determine biomass recalcitrance with respect to lignin.

To bypass recalcitrance-associated bioprocessing limitations, biorefinery industries have employed specific enzymatic and chemical pre-treatments in both SSF and DPP. For example, hemicelluloses are deacetylated in the initial steps of DPP using large quantities of alkali which leads to an increase in yield and the tensile strength of paper (Laffend 1967). However, while several pre-treatment strategies exist to counteract biomass recalcitrance and ensure efficient bioprocessing, most of these are at an extra cost. For example, the de-acetylation of hemicellulose consumes large quantities of alkali during the Kraft process, which have repercussions on the overall cost of the process (Zanuttini *et al.* 2003). In addition to cost, most of the by-products of pre-treatment contribute to a separate form of recalcitrance. During the heating stage of pre-treatment, the GlcA modifications on the xylan backbone cause xylan to shift from the hydrophilic to the hydrophobic face of cellulose, rather than being released, requiring additional treatment for complete xylan removal (Busse-Wicher *et al.* 2016; Pereira *et al.* 2017). Furthermore, during deacetylation, acetate accumulates as a by-product in the spent liquor contributing to the inhibition of enzymatic activity, the alteration of the hydrolytic enzymes, and the pH of the medium (De Mey *et al.* 2007). Altogether, current bioprocessing approaches are expensive and time-consuming, while still suffering from low yields and compromised

product quality (Söderström *et al.* 2002; Nguyen *et al.* 2008; Sassner *et al.* 2008). These limitations demand the application of novel strategies, such as genetic engineering of plants to yield plants whose secondary cell walls can be easily deconstructed to help improve the bioprocessing efficiency, and successively counteract the limitations associated with current bioprocessing approaches.

1.5.3 Current achievements in the improvement of lignocellulose biorefinery

Applying genetic engineering approaches to improve bioprocessing requires a far-reaching understanding of the traits that potentially affect bioconversion and the precise means through which these traits alter the output of bioprocessing. In this regard, molecular and cell biology have uncovered important aspects regarding the biosynthesis of lignocellulose and how the structure and modifications thereof affect industrial processing. This knowledge has helped identify important lignocellulose traits that could potentially be affecting the secondary cell wall's resistance to bioprocessing. As previously discussed, these traits include the cell wall's total lignin content, the composition of lignin, and the extent of xylan acetylation. Variation in these traits affects the cell wall's recalcitrance to bioprocessing and the overall bioprocessing outputs, such as glucose and pulp yield. With this understanding, several studies have manipulated endogenous secondary cell wall biosynthetic genes to engineer plants with desired lignocellulose traits, for example, reduced lignin content, subsequent reduction of recalcitrance, and associated improvement in bioprocessing efficiency.

Upon stem saccharification treatment, *AtUXS* (involved in the biosynthesis of UDP-Xyl) *uxs3 uxs5 uxs6* T-DNA triple mutants exhibited a slight improvement in the monosaccharide release efficiency compared to wild-type (Kuang *et al.* 2016). Similarly, simultaneous downregulation of *IRX9* and *IRX14* in hybrid aspen (*Populus tremula* x *P. tremuloides*) using wood-specific promoter leads to reduced xylose content accompanied by a higher lignocellulose saccharification efficiency (Ratke *et al.* 2018). Genetic editing of lignin biosynthetic genes also presents some improvements in

saccharification efficiency. For example, *Populus tremula* x *P. alba* transgenics with downregulated CINNAMOYL-COA REDUCTASE (CCR, an enzyme catalyzing the first step of monolignol biosynthesis) activity, showed up to a 161% increase in ethanol yield upon saccharification assays (Acker *et al.* 2014). Furthermore, a large scale omics study aimed at identifying wood traits impacted by altering lignin biosynthetic pathway used RNAi to perturb the expression of at least 21 lignin biosynthetic pathway genes in *Populus trichocarpa* (Wang *et al.* 2018). Subsequent multi-omics quantitative analyses of transgenics showed that engineering the expression of lignin biosynthetic genes could alter tree growth and saccharification efficiency (Wang *et al.* 2018).Altogether, these studies showed that genetic alteration of secondary cell wall biosynthetic genes can yield plants with secondary cell walls that are easier to deconstruct.

In addition to targeting the backbone biosynthetic genes, an alternative attempt to improve lignocellulose bioprocessing efficiency may involve alteration of biopolymer backbone modification and side-chain substitution genes. *Arabidopsis* plants lacking *GXM1* have a severely reduced degree of 4-*O*-MeGlcA content and an increased glucuronoxylan release following saccharification (Urbanowicz *et al.* 2012a). While this result showed that GlcA 4-O-methylation potentially affects bioprocessing, it is the extensive *O*-acetylation of xylan that significantly hinders lignocellulose saccharification, consequently, several studies have focused on modifying xylan acetylation to potentially improve saccharification efficiency. Simultaneous downregulation of *RWA* gene clades *A* and *B* in hybrid aspen (*Populus tremula* x *tremuloides*) lead to 25% reduction in xylan acetylation with no impact on plant growth (Pawar *et al.* 2017b). Furthermore, transgenic plants showed increased glucose and xylose yield following enzymatic hydrolysis (Pawar *et al.* 2017b). Similarly, expression of *Aspergillus niger* ACETYL XYLAN ESTERASE (*An*AXE1, which deacetylate xylans) in aspen (*Populus tremula L.* x *tremuloides Michx*.) lead to normally growing transgenic plants with reduced xylan and over 25% increase in glucose yielded per unit dry weight compared to wild-type (Pawar *et*

al. 2017a). Similar improvements in lignocellulose saccharification were observed when fungal acetyl xylan esterase was expressed in *Arabidopsis thaliana* (Pawar *et al.* 2016). Thus, these observations reinforce the idea that reducing xylan acetylation may reduce steric hindrance for hydrolytic enzymes, and potentially increase the efficiency of lignocellulose bioprocessing. Finally, although most of the studies described here focused on the modification of biosynthetic genes, it has also been demonstrated that altering metabolic flux to secondary cell wall biosynthetic pathways could potentially yield secondary cell walls with traits beneficial for bioprocessing (Fan *et al.* 2017). More importantly, targeting the regulators (e.g., transcription factors) of the secondary cell wall biosynthetic pathways is also an important avenue that can be taken to alter the overall cell wall biosynthesis and possibly yield desirable phenotypes (Hori *et al.* 2020).

While the above-mentioned studies indeed showed that targeting secondary cell wall biosynthetic and modification genes may be a feasible way of improving lignocellulose bioprocessing, such improvements may be accompanied by plants with compromised growth or other unfavorable outcomes. Some of the undesired traits may be attributed to the use of constitutive promoters instead of cell type-specific promoters to drive the expression of gene modifications. Moreover, the extent of genetic modification, for example, the level of gene silencing, may affect the overall effect of genetic modification on plant development. For example, whilst CRISPR/Cas9-mediated *BdGT43B2* knock-out indeed suggested that this gene may be important in xylan biosynthesis, mutants were stunted and died at the seedling stage (Petrik *et al.* 2020). Similarly, downregulation of *Populus* CCR enzyme leads to increased ethanol yield as previously discussed, however, a stronger CCR down regulation leads to reduction of biomass yield (Acker *et al.* 2014). Therefore, utilizing gene modification approach for rapid gene functional characterization may be a challenge towards improvement of industrial bioprocessing since extra work could be required to corroborate whether the observed phenotype is due to the indispensable role of knocked-out gene or the method and extent of gene silencing. This

further leads to the limited functional characterization of novel secondary cell wall biosynthetic genes, which in turn makes it difficult to predict the impact that genetic modification of a specific gene of interest may have on the desired traits and overall plant growth. Such limitations have also been observed in the previously described *Arabidopsis tbl29* mutants, which yield the desirable 60% reduction in *O*-acetylation, but accompanied by the unwanted reduction in plant growth, collapsed xylem, and reduced biomass production (Xiong *et al.* 2013a). Interestingly, at least for this mutant, overexpression of *At*GUX enzyme in *tbl29* background restores plant growth while maintaining low acetylation level (Xiong *et al.* 2015). Thus, improvement of bioprocessing through gene modification may not just involve modifying a single specific gene in a specific way, sometimes gene stacking, or complementation of mutants may be necessary. Interestingly, sophisticated cell-specific approaches which avoid modification of xylem vessels (but alter other cell types) have also been employed in cell wall engineering. For example, the promoter of *C4H*, a key lignin gene, was replaced by the vessel-specific promoter of *VND6*, limiting lignin biosynthesis to vessels from the fibers (Yang *et al.* 2013). This resulted in reduced lignin and increased polysaccharide deposition in the fiber cells, accompanied with higher sugar yields following enzymatic hydrolysis (Yang *et al.* 2013)

Clearly, successful application of genetic engineering strategy for improvement of industrial bioprocessing requires sufficient knowledge about cellular metabolism, and genes involved in different metabolic pathways. This requires advanced methods of gene identification together with precise and efficient genetic engineering approaches.

1.6 The potential improvement of bioprocessing efficiency through genetic engineering

Recent advancements in systems genetics, which is a computational and mathematical modelling approach that elucidates the behavior of complex biological organization and pathways (Kirschner 2005), have helped identify several candidate secondary cell wall biosynthetic and modification genes. Using this approach, known xylan modification genes were used as bait to pull-out a suite of xylan-associated genes in *Eucalyptus*, subsequent co-expression, and correlation analyses was able to correlate these genes with lignocellulose traits that may be potentially affecting bioprocessing efficiency (Wierzbicki et al. 2019a). Correlation of genes with bioprocessing traits provides important information enabling prediction of the impact of gene manipulation on lignocellulose traits and bioprocessing efficiency. While the xylan-associated genes were primarily identified in Eucalyptus, identifying gene homologs in other commercial plant species will enable the application of gene editing to attain desired traits in those species. Thus, advancement in systems genetics coupled with improvements in genetic engineering approaches such as the CRISPR, may allow precisely targeted engineering of plant secondary cell walls, a possible one step towards potential improvement of bioprocessing efficiency.

1.6.1 Understanding the fundamental principles and limitations of CRISPR/Cas9

Numerous functional genetics studies have extensively utilized the CRISPR system to study many significant traits in a wide range of important crop and model plants such as *Arabidopsis,* soybean, wheat, poplar, rice, and cotton. Amongst other CRISPR/Cas systems, the type II CRISPR/Cas system repurposed from *Streptococcus pyogenes* Cas9 (SpCas9) has been widely used for a broad range of applications due to its simplicity, versatility, and efficiency. Under the guidance of a user-designed 20

bp sequence-specific single guide RNA (sgRNA) and the conserved protospacer adjacent motif (PAM), Cas9 endonuclease cleaves both DNA strands at the target region (**Figure 1.2**). The two DNA strands are cut by a pair of conserved Cas9 nuclease subdomains; HNH and RuvC, HNH cuts the strand forming DNA-RNA duplex with the sgRNA while RuvC cuts the opposite strand (Garneau *et al.* 2010; Chapman *et al.* 2012; Jinek *et al.* 2012; Ran *et al.* 2013b). This strand-specific DNA cleavage introduces site-specific double-stranded DNA breaks (DSBs) into the genomic locus, the resulting DSBs can then either be repaired via the error-prone nonhomologous end-joining (NHEJ) or the homology-directed repair (HDR) (Jinek *et al.* 2012; **Figure 1.2**). When repaired via the NHEJ, gene knock-outs often result from frameshift mutations and stop codons generated through an induced mixture of random insertions and deletions (Jinek *et al.* 2012; Cong *et al.* 2013; Danner *et al.* 2017). Conversely, in the presence of an exogeneous DNA repair template encoding a specific trait, HDR can introduce precise DNA modifications as directed by the template (**Figure 1.2**), however, HDR-mediated repair is inefficient and NHEJ remains the predominant repair mechanism even in the presence of a repair template (Chapman *et al.* 2012; Cox *et al.* 2015). The inefficiency of HDR repair makes precise genome editing very challenging, this and other CRISPR limitations are discussed next.

The limited precise genomic insertion capability offered by CRISPR/Cas9 is attributed to the blunt ends generated by Cas9's DNA cleavage coupled with the inefficiency of HDR repair. This is partly addressed by the type V CRISPR system which uses Cas12a (or Cpf1) endonuclease instead of Cas9. Cas12a cuts DNA 3' downstream of PAM in a staggered fashion, producing sticky ends with 4 or 5 nucleotides overhangs instead of blunt ends generated by Cas9 (Zetsche *et al.* 2015; Fonfara *et al.* 2016). These sticky ends increase the efficiency of the template DNA insertions into a complementary site through HDR-mediated repair, furthermore, the enhancement of DNA insertions may reduce the frequency of NHEJ-mediated repair, reducing the possibility of unwanted gene effects (reviewed in Zaidi *et al.* 2017; Chaudhary *et al.* 2018). Moreover, Cas12a recognizes a thymidine-rich (5'-TTN-3')

PAM sequences instead of to Guanine-rich (5'-NGG-3') PAM required for Cas9 DNA cleavage, this allows editing of AT-rich regions such as the promoter and the untranslated regions (UTRs), increasing the editing scope (Zetsche *et al.* 2015). More importantly, the overall size of gRNA required for Cas12a editing is ~43 bp, which is approximately 50% smaller than Cas9's gRNA, this allows an efficient packaging and delivery of CRISPR reagents into plant cells (Zetsche *et al.* 2015, 2017). Finally, one of the major shortcomings of CRISPR/Cas9 is its high off-target activities, which is of course a big concern in mammalian cells but can also yield lethal phenotypes in plants. Fortunately, CRISPR/Cas12a has been shown to offer limited to no off-target effects in both mammalian and plant cells (Kleinstiver *et al.* 2016; Xu *et al.* 2017; Li *et al.* 2018a). Indeed, Cas12a is superior to Cas9 in most aspects, and to this end, CRISPR/Cas12a has been used extensively for efficient and precise gene editing at least in tobacco and rice (Endo *et al.* 2016; Begemann *et al.* 2017; Hu *et al.* 2017; Kim *et al.* 2017; Wang *et al.* 2017; Yin *et al.* 2017). However, there are several other identified new generation endonucleases such as those targeting single-stranded RNA and others offering several advantages to the classic Cas9 or Cas12a (see Chaudhary *et al.* 2018 and Chen *et al.* 2019 for detailed review).

A key step in CRISPR-based gene editing approaches described above is the successful delivery of CRISPR reagents into plant cells. Amongst many different delivery options, *Agrobacterium*-mediated transformation is the most robust and simple to use, it is consequently the most commonly used delivery method in most plant species such as tobacco, rice, *Arabidopsis*, and poplar (Feng *et al.* 2013; Mao *et al.* 2013; Jiang *et al.* 2013; Gao *et al.* 2015; Char *et al.* 2017). Both Cas9 and sgRNA expression cassettes are carried by the T-DNA binary vector within *Agrobacterium,* plants are then transformed with the T-DNA-carrying *Agrobacteria* mostly through the floral dipping method, this allows the direct delivery of expression cassettes into the plant genome. More importantly, the expression cassettes also carry selectable markers used for screening seedlings, after which, putative transformants are screened for genetic modification at the target site. To achieve the desirable gene

editing outcomes with this approach, the expression cassettes need to completely integrate into the genome, which poses some limitations. For example, integration of constructs and gene markers into the genome may cause increased off-target effects which could hinder the success of this approach in commercial applications (Gao *et al.* 2016). Furthermore, depending on the species, the successful and stable integration of the expression cassettes into the genome may not always occur, this presents a major limiting factor in the adoption of the platform for gene editing in several species. Fortunately, gene mutations can also be generated without the need for stable integration of the CRISPR reagents into the plant genome. This transgene-free gene editing is achieved through transient expression of CRISPR/Cas9 reagents, which can be mediated through diverse alternative delivery approaches such as viral delivery, RNA delivery, ribonucleotide protein complex delivery, and plasmid delivery. A comprehensive review of these mechanisms is provided elsewhere: Lowder *et al.* 2016; Liu *et al.* 2017; Chen *et al.* 2019; Mao *et al.* 2019.

Besides its shortcomings, several studies have reported an exceptional efficiency and consistency of the CRISPR/Cas9 system for genome modifications in plants. Amongst others, targeted mutations were observed in 95% of 88 transgenic events following CRISPR/Cas9 modification of soybean genes (Jacobs *et al.* 2015). Similarly, to generate reliably contained transgenic plants, Elorriaga *et al.* 2018 achieved high mutation rates following CRISPR/Cas9-mediated alteration of the floral meristem identity gene, leafy *(LFY),* and two floral organ identity gene (agamous; *AG)* orthologs in poplar. Furthermore, Fan *et al.* 2015 targeted the phytoene desaturase gene 8 (*PtoPDS*) in poplar, where an obvious mutant albino phenotype was observed in all transgenic plants. Moreover, the primary role of *4-COUMARATE:CA* (*4CL*) genes in lignin biosynthesis were confirmed through CRISPR/Cas9 editing of *4CL1* and *4CL2* in *Populus* (Tsai and Xue 2015). Recently, CRISPR/Cas9 knock-out of *BdGT43B2* in *Brachypodium distachyon* led to reduction of xylan and cell size in knock-outs, suggesting that *BdGT43B2* may be playing a similar role to the *Arabidopsis* GT43 protein family

(Petrik *et al.* 2020). These studies showed that CRISPR/Cas9 technology is indeed effective in inducing targeted mutations and can thus be useful in precisely engineering plants with desired bioprocessing traits. However, CRISPR/Cas9 system have many limitations which has subsequently led to the development of a wide range of more flexible CRISPR-based gene editing approaches. Some of these second-generation approaches and the CRISPR/Cas9 limitations they aim to address are discussed next.

1.6.2 Repurposing CRISPR/Cas9 for second-generation gene-editing techniques

A different approach in improving CRISPR-based gene editing is modifying the Cas9 protein, this approach has been exploited to generate several second-generation novel gene-editing technologies. For example, the RuvC catalytic domain of Cas9 can be mutated through the aspartate-to-alanine (D10A) mutation to generate the Cas9n nickase capable of nicking DNA to yield single-stranded breaks (SSBs) (Sapranauskas *et al.* 2011; Gasiunas *et al.* 2012; Jinek *et al.* 2012). Different from DSBs, SSBs are preferentially repaired through HDR-mediated repair mechanism, such preferred HDR-mediated repair could potentially decrease the frequency of random NHEJ-mediated indels while simultaneously allowing the introduction of precise gene modifications (Cong *et al.* 2013). Additionally, double nicking with Cas9n to induce DSBs at the target locus can be achieved. This is important because compared to the blunt-end DSBs generated by Cas9, the DSBs generated by Cas9 nickases are staggered with overhangs (Jinek *et al.* 2012; Cong *et al.* 2013; Mali *et al.* 2013b), this ensures an increased target specificity and efficient gene modification (Ran *et al.* 2013a).

A more sophisticated alternative CRISPR application is achieved when both HNH and RuvC Cas9 domains are mutated to generate the catalytically inactive or deactivated Cas9 (dCas9). Under the guidance of sgRNA, dCas can tightly bind to the specified target sequence and interfere with the

binding of regulatory proteins such as transcription activators to the DNA, thereby repressing gene expression (Qi *et al.* 2013b). Gene repression can also be achieved by fusing dCas with a strong trans repression domain (CRISPR mediated interference or CRISPRi); while fusing dCas with an activator allows transcriptional activation (CRISPRa; (Maeder *et al.* 2013; Qi *et al.* 2013b; Gilbert *et al.* 2015). Thus, a fusion of dCas with effector domains enables efficient transcriptional regulation, an application that can be important in modulating metabolite pathways during cell wall biosynthesis and modification. This may be one way of affecting the amount and flux of metabolites during cell wall synthesis. Although modification of gene expression through CRISPR/dCas9 has not been widely reported in woody plants, successful transcriptional activation through dCas9 fusion with EDLL and TAL motifs have been reported in *Nicotiana benthamiana* (Piatek *et al.* 2015). Moreover, dCas9 fusion with the SRDX repression domain has been successful in inducing repression in *N. benthamiana* (Piatek *et al.* 2015). In parallel to these approaches, CRISPR target specificity has been further improved through the inducible CRISPR/Cas system. Here, the activity of Cas9 or derivatives thereof is controlled spatially and temporally through light, specific chemicals, or other forms of ligands (Hemphill *et al.* 2015; Nihongaki *et al.* 2015; Polstein and Gersbach 2015; Oakes *et al.* 2016). The inducible Cas system can be beneficial for improved gene editing programmability and reducing off-target effects or toxicity to the cells, these aspects and few attempts at applying inducible CRISPR/Cas systems are recently reviewed in McCarty *et al.* 2020.

In addition to the modified CRISPR/Cas approaches discussed above, base editing and prime editing are the two newly developed second-generation genome editing technologies based on CRISPR. Base editing is aimed at addressing the limitations associated with HDR repair inefficiency by directly introducing precise single base substitutions into the genome in a template- and DSBs-independent manner. Two classes of base editors exist: (i) cytidine base editors (CBEs) are comprised of Cas9n or dCas9 or dCas12a fused with a cytidine deaminase which converts CG into TA pairs, (ii) in adenine

base editors (ABEs) Cas9n or dCas9 is fused with adenine deaminase which converts AT to GC (Komor *et al.* 2016; Nishida *et al.* 2016; Li *et al.* 2018b). The precise mechanism of these systems are reviewed in Bharat *et al.* 2019; Mao *et al.* 2019, and several studies have used these systems in many applications including functional annotation in *Arabidopsis* and rice (Chen *et al.* 2017b; Li *et al.* 2017; Ren *et al.* 2017; Yan *et al.* 2018).

While base editors have certainly managed to outperform HDR-mediated substitutions with respect to efficiency and the purity of the edited product, prime editors (PEs) go several steps further. Far from being able to generate indels like CRISPR/Cas9, creating single base substitutions like base editors, prime editors (PEs) can search and precisely replace target DNA with the desired DNA sequence (Anzalone *et al.* 2019). Moreover, prime editing is superior to both the classic CRISPR/Cas9 and base editing owing to its ability to introduce indels and all 12 possible base-base transitions and transversions without inducing DSBs (Anzalone *et al.* 2019). To achieve its multifaceted roles, the prime editing machinery consists of the (i) Cas9n (H840A) to nick the target DNA strand, (ii) prime-editing RNA (pegRNA) that serve as both a guide to the target and a template specifying the desired trait and (iii) a reverse transcriptase (RT) which copies information from the template into the target to encode the specified trait (Anzalone *et al.* 2019; **Figure 1.3**). This is followed by a DNA repair mechanism that converts the DNA strand into the desired sequence instead of reverting to the original sequence (reviewed in Marzec *et al.* 2020 and Urnov 2020). The abovementioned components constitute the prime editor 1 (PE1) gene editing system which is not very efficient. To optimize its genetic editing efficiency, the reverse transcriptase can be linked to Cas9 (H840A) nickases and the triple mutant Moloney murine leukemia virus reverse transcriptase (M-MLV-RT) to yield the second-generation prime editing system 2 (PE2) (Anzalone *et al.* 2019). This was shown to be effective in introducing single nucleotide insertions into the *OsPDS* gene, and mediating G-C transversions in *OsACC* gene in rice with a 1.6 – 5.1-fold improvement in efficiency compared to PE1 (Xu *et al.* 2020).

In both cases, precisely edited mutants were obtained, although the mutation rates were low (7.3% and 14.6% respectively). Consequently, to further improve editing efficiency, in addition to engineered RT (PE2), the PE3 system employs nickases that also nick the non-target strand, which leads to up to a 10% increase in indel formation at least in human cells (Anzalone *et al.* 2019; Xu *et al.* 2020). Importantly, the editing efficiency of any of these three prime editing systems depends on species, the cell type, and possibly the specific target site (Flotte and Gao 2019). Such inefficiencies are to be expected as this technology is still in its infancy, however, prime editing presents the most versatile gene editing technology thus far, allowing installations of all possible point mutations, smalls indels and precise combinations thereof. More importantly, prime editing enables precise targeted gene editing with increased target specificity and minimal unwanted genomic changes.

1.7 Conclusions

The main secondary cell wall biopolymers (cellulose, hemicellulose, and lignin) present major renewable raw materials to produce many bioproducts and alternative forms of bioenergy. Generating such value-added bioproducts requires a complete separation of these biopolymers into their individual constituents through biorefinery methods. However, the tight interaction amongst the cell wall biopolymers generally creates steric hindrance for the hydrolytic enzymes during bioprocessing, thereby presenting a challenge for the complete deconstruction of the biopolymers (Helle *et al.* 2003; Qing and Wyman 2011; Chong *et al.* 2015). Counteracting such challenges has generally relied on alteration of the biorefinery techniques; this mostly involved the inclusion of additional chemical and enzymatic pre-treatment steps in the biorefinery pathways. Nevertheless, the additional costs and accompanying compromises on biomass yield associated with additional biorefinery steps necessitates the application of novel genetic engineering approaches for the improvement of bioprocessing efficiency. The application of such novel approaches requires an extensive understanding of the

secondary cell wall structure and how alteration thereof can lead to bioprocessing improvement. Accordingly, substantial progress has been made towards understanding the biosynthesis of SCWs including the elucidation of genes and gene products involved in the biosynthesis of the basic cell wall structure and modifications thereof. Furthermore, the importance of sugar-nucleotide precursors as a substrate for biosynthesis of biopolymer backbones has been emphasized, similarly, the requirement of several metabolic precursors for cell wall modifications has been demonstrated. However, some unanswered questions exist about the overall cell wall biosynthetic pathway. For example, conclusive biochemical evidence for the involvement of IRX9 and IRX14 in xylan backbone formation is still lacking. This further question the exact constituents of the XSC, i.e., whether only IRX9, IRX14, and IRX10 are necessary and sufficient to generate xylan backbone, or whether there are other essential players in this process. More importantly, the implication of secondary cell wall biopolymers' interaction in biomass recalcitrance has resulted in interrogations of the nature of such interactions. It has previously been suggested that there are extensive xylan-cellulose interactions determined by the presence of minor and major xylan domains; which are attributed to the MeGlcA spacing (Bromley *et al.* 2013). Conversely, recent evidence showed that there are very limited xylan-cellulose interactions in the secondary cell wall, instead, xylan mainly binds lignin through electrostatic forces, while further acting as a connection between lignin and cellulose (Kang *et al.* 2019). In the prior case, the extent of biopolymer interaction can be reduced through alteration of the expression of methylation and GlcA modification genes (perhaps through gene stacking), thereby allowing easier hydrolysis of the cell wall. A different approach can be metabolic engineering, which generally involves alteration of the metabolic flux of specific metabolites in a tissue-specific manner. For secondary cell wall modifications, the application of such approach may suffer from the limited knowledge about the exact mechanisms and transporters involved in the transport of SAM (methyl donor for MeGlcA modification) and Acetyl-CoA necessary for xylan and lignin modifications (Fan *et al.* 2017).

Therefore, studies aimed at understanding the metabolic flux and the regulation thereof are required; similarly, the elucidation of metabolite transport mechanisms will be equally beneficial.

Nevertheless, with the current understanding of cell wall biosynthetic genes and metabolic precursors, genetic engineering holds an enormous promise for an effective way to alter cell wall biosynthesis and yield plants with desired traits for efficient bioprocessing. For example, downregulation of cinnamoyl-CoA reductase (CCR), a key enzyme in lignin biosynthesis, leads to a significant increase in ethanol yield following saccharification in hybrid poplar (Acker *et al.* 2014). Similarly, simultaneous downregulation of putative xylan backbone biosynthetic genes leads to reduced xylose and improved wood saccharification efficiency (Ratke *et al.* 2018). Several other studies have also demonstrated such improvements following downregulation or knock-out of modification genes such as *GXM* involved in methylation and *RWA* necessary for xylan acetylation (Urbanowicz *et al.* 2012a; Kuang *et al.* 2016; Pawar *et al.* 2017a; b). Some of the notable limitations of these prior studies, however, are the inability to predict the expected impact that targeting a specific gene may have on cell wall digestibility and subsequent bioconversion efficiency, furthermore, a gene downregulation approach suffers from the residual gene function that may make it a challenge to achieve biotechnologically relevant and desired traits. To this end, systems genetics have helped establish correlation relationships between gene expression and some bioprocessing traits (Wierzbicki *et al.* 2019a), which is useful in predicting the impact that targeting a gene may have on bioprocessing efficiency.

To stably fine-tune residual gene expression and effectively target genes, few studies have employed CRISPR/Cas9 in several plant species (Fan *et al.* 2015; Jacobs *et al.* 2015; Elorriaga *et al.* 2018; Petrik *et al.* 2020; De Meester *et al.* 2020). CRISPR further increases the scope of editing by its highly adaptable nature, allowing other useful CRISPR-based technologies to be applied for a wide range of gene modifications, these include CRISPR/dCas9, which uses deactivated Cas9 to specifically regulate

gene expression; base editing, which precisely edits specific DNA bases and most recently; prime editing has allowed precise edits to be introduced into the genome without the need for cutting both DNA strands (Anzalone *et al.* 2019; Xu *et al.* 2020; Lin *et al.* 2020).

Altogether, genetic engineering through CRISPR-based tools promises a far more efficient and reliable means to alter cell wall biosynthetic genes and generate plants with desirable traits. Moreover, the extensive understanding of secondary cell wall biosynthesis from the necessary metabolic precursors, to how the biopolymers interact to form mature SCW is important in establishing how we can alter the cell wall to achieve desired bioprocessing traits. Here, I aimed to combine the knowledge about cell wall biosynthetic genes and genetic engineering to determine how altering novel xylan-associated genes could impact cell wall development, plant growth, and physiology in *Arabidopsis* and *Populus.*

1.8 References

Acker R. Van, J. C. Leple, D. Aerts, V. Storme, G. Goeminne, *et al.*, 2014 Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. Proc. Natl. Acad. Sci. 111: 845–850. https://doi.org/10.1073/pnas.1321673111

Albrecht G., and A. Mustroph, 2003 Localization of sucrose synthase in wheat roots: Increased *in situ* activity of sucrose synthase correlates with cell wall thickening by cellulose deposition under hypoxia. Planta 217: 252–260. https://doi.org/10.1007/s00425-003-0995-6

Alejandro S., Y. Lee, T. Tohge, D. Sudre, S. Osorio, *et al.*, 2012 AtABCG29 is a monolignol transporter involved in lignin biosynthesis. Curr. Biol. 22: 1207–1212. https://doi.org/10.1016/j.cub.2012.04.064

Amor Y., C. H. Haigler, S. Johnson, M. Wainscorr, and D. P. Delmer, 1995 A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. Proc. Natl. Acad. Sci. 92: 9353–9357. https://doi.org/10.1088/0256-307X/33/7/072901

Amthor J. S., 2003 Efficiency of lignin biosynthesis: A quantitative analysis. Ann. Bot. 91: 673–695. https://doi.org/10.1093/aob/mcg073

Anzalone A. V., P. B. Randolph, J. R. Davis, A. A. Sousa, L. W. Koblan, *et al.*, 2019 Search-and-replace genome editing without double-strand breaks or donor DNA. Nature 576: 149–157. https://doi.org/10.1038/s41586-019-1711-4

Arioli T., L. Peng, A. S. Betzner, J. Burn, W. Wittke, *et al.*, 1998 Molecular analysis of cellulose biosynthesis in *Arabidopsis*. Science 279: 717–720. https://doi.org/10.1126/science.279.5351.717

Bar-Peled M., and M. A. O'Neill, 2011 Plant nucleotide sugar formation, interconversion, and salvage by sugar recycling. Annu. Rev. Plant Biol. 62: 127–155. https://doi.org/10.1146/annurev-arplant-042110-103918

Barnes W. J., and C. T. Anderson, 2018 Cytosolic invertases contribute to cellulose biosynthesis and influence carbon partitioning in seedlings of *Arabidopsis thaliana*. Plant J. 94: 956–974. https://doi.org/10.1111/tpj.13909

Barratt D. H. P., P. Derbyshire, K. Findlay, M. Pike, N. Wellner, *et al.*, 2009 Normal growth of *Arabidopsis* requires cytosolic invertase but not sucrose synthase. Proc. Natl. Acad. Sci. 106: 13124–13129. https://doi.org/10.1073/pnas.0900689106

Barros J., H. Serk, I. Granlund, and E. Pesquet, 2015 The cell biology of lignification in higher plants. Ann. Bot. 115: 1053–1074. https://doi.org/10.1093/aob/mcv046

Basen M., A. M. Rhaesa, I. Kataeva, C. J. Prybol, I. M. Scott, *et al.*, 2014 Degradation of high loads of crystalline cellulose and of unpretreated plant biomass by the thermophilic bacterium *Caldicellulosiruptor bescii*. Bioresour. Technol. 152: 384–392. https://doi.org/10.1016/j.biortech.2013.11.024

Bashline L., L. Lei, S. Li, and Y. Gu, 2014 Cell wall, cytoskeleton, and cell expansion in higher plants. Mol. Plant 7: 586–600. https://doi.org/10.1093/mp/ssu018

Begemann M. B., B. N. Gray, E. January, G. C. Gordon, Y. He, *et al.*, 2017 Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-11760-6

Bharat S. S., S. Li, J. Li, L. Yan, and L. Xia, 2019 Base editing in plants: Current status and challenges. Crop J. 1–12. https://doi.org/10.1016/j.cj.2019.10.002

Bikard D., W. Jiang, P. Samai, A. Hochschild, F. Zhang, *et al.*, 2013 Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. 41: 7429–7437. https://doi.org/10.1093/nar/gkt520

Boerjan W., J. Ralph, and M. Baucher, 2003 Lignin biosynthesis. Annu. Rev. Plant Biol. 54: 519–546. https://doi.org/10.1146/annurev.arplant.54.031902.134938

Boija E., A. Lundquist, K. Edwards, and G. Johansson, 2007 Evaluation of bilayer disks as plant cell membrane models in partition studies. Anal. Biochem. 364: 145–152. https://doi.org/10.1016/j.ab.2007.02.012

Bonawitz N. D., and C. Chapple, 2010 The genetics of lignin biosynthesis: Connecting genotype to phenotype. Annu. Rev. Genet. 44: 337–363. https://doi.org/10.1146/annurev-genet-102209-163508

Brandon A. G., D. S. Birdseye, and H. V. Scheller, 2020 A dominant negative approach to reduce xylan in plants. Plant Biotechnol. J. 18: 5–7. https://doi.org/10.1111/pbi.13198

Brill E., M. van Thournout, R. G. White, D. Llewellyn, P. M. Campbell, *et al.*, 2011 A novel isoform of sucrose synthase is targeted to the cell wall during secondary cell wall synthesis in cotton fiber. Plant Physiol. 157: 40–54. https://doi.org/10.1104/pp.111.178574

Bringmann M., E. Li, A. Sampathkumar, T. Kocabek, M.-T. Hauser, *et al.*, 2012 POM-POM2/CELLULOSE SYNTHASE INTERACTING1 is essential for the functional association of cellulose synthase and microtubules in *Arabidopsis*. Plant Cell 24: 163–177. https://doi.org/10.1105/tpc.111.093575

Bromley J. R., M. Busse-Wicher, T. Tryfona, J. C. Mortimer, Z. Zhang, *et al.*, 2013 GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. Plant J. 74: 423–434. https://doi.org/10.1111/tpj.12135

Brown D. M., L. A. H. Zeef, J. Ellis, R. Goodacre, and S. R. Turner, 2005 Identification of novel genes in *Arabidopsis* involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell 17: 2281–2295. https://doi.org/10.1105/tpc.105.031542

Brown D. M., F. Goubet, V. W. Wong, R. Goodacre, E. Stephens, *et al.*, 2007 Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. Plant J. 52: 1154–1168. https://doi.org/10.1111/j.1365-313X.2007.03307.x

Brown D. M., Z. Zhang, E. Stephens, P. Dupree, and S. R. Turner, 2009 Characterization of *IRX10* and I*RX10-like* reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. Plant J. 57: 732–746. https://doi.org/10.1111/j.1365-313X.2008.03729.x

Brown D., R. Wightman, Z. Zhang, L. D. Gomez, I. Atanassov, *et al.*, 2011 *Arabidopsis* genes *IRREGULAR XYLEM* (*IRX15*) and *IRX15L* encode DUF579-containing proteins that are essential for normal xylan deposition in the secondary cell wall. Plant J. 66: 401–413. https://doi.org/10.1111/j.1365-313X.2011.04501.x

Busse-Wicher M., T. C. F. Gomes, T. Tryfona, N. Nikolovski, K. Stott, *et al.*, 2014 The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of *Arabidopsis thaliana*. Plant J. 79: 492–506. https://doi.org/10.1111/tpj.12575

Busse-Wicher M., A. Li, R. L. Silveira, C. S. Pereira, T. Tryfona, *et al.*, 2016 Evolution of xylan substitution patterns in gymnosperms and angiosperms: implications for xylan interaction with cellulose. Plant Physiol. 171: pp.00539.2016. https://doi.org/10.1104/pp.16.00539

Capek P., J. Alfo, and L. Desana, 2002 An acetylated galactoglucomannan from *Picea abies* L. Karst. Carbohydr. Res. 337: 1033–1037. https://doi.org/10.1016/s0008-6215(02)00090-3

Carroll A., and C. Somerville, 2008 Cellulosic biofuels. Annu. Rev. Plant Biol. 60: 165–182. https://doi.org/10.1146/annurev.arplant.043008.092125

Chandel A. K., C. Gajula, R. Konakalla, R. Rudravaram, and R. Pogaku, 2011 Bioconversion of pentose sugars into ethanol: A review and future directions. Biotechnol. Mol. Biol. Rev. 6: 8–20. https://doi.org/10.1079/9781845936662.0101

Chapman J. R., M. R. G. Taylor, and S. J. Boulton, 2012 Playing the end game: DNA double-strand break repair pathway choice. Mol. Cell 47: 497–510. https://doi.org/10.1016/j.molcel.2012.07.029

Char S. N., A. K. Neelakandan, H. Nahampun, B. Frame, M. Main, *et al.*, 2017 An *Agrobacterium*-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. Plant Biotechnol. J. 15: 257–268. https://doi.org/10.1111/pbi.12611

Chaudhary K., A. Chattopadhyay, and D. Pratap, 2018 The evolution of CRISPR/Cas9 and their cousins: hope or hype? Biotechnol. Lett. 40: 465–477. https://doi.org/10.1007/s10529-018-2506-7

Chen F., and R. A. Dixon, 2007 Lignin modification improves fermentable sugar yields for biofuel production. Nat. Biotechnol. 25: 759–761. https://doi.org/10.1038/nbt1316

Chen F., Y. Tobimatsu, L. Jackson, J. Nakashima, J. Ralph, *et al.*, 2013a Novel seed coat lignins in the Cactaceae: Structure, distribution and implications for the evolution of lignin diversity. Plant J. 73: 201–211. https://doi.org/10.1111/tpj.12012

Chen X., M. E. Vega-Sánchez, Y. Verhertbruggen, D. Chiniquy, P. E. Canlas, *et al.*, 2013b Inactivation of *OsIRX10* leads to decreased xylan content in rice culm cell walls and improved biomass saccharification. Mol. Plant 6: 570–573. https://doi.org/10.1093/mp/sss135

Chen J., Q. Jiang, G. Yang, Q. Wang, and P. Fatehi, 2017a Ultrasonic-assisted ionic liquid treatment of chemithermomechanical pulp fibers. Cellulose 24: 1483–1491. https://doi.org/10.1007/s10570-016-1180-y

Chen Y., Z. Wang, H. Ni, Y. Xu, Q. Chen, *et al.*, 2017b CRISPR/Cas9-mediated base-editing system efficiently generates gain-of-function mutations in *Arabidopsis*. Sci. China Life Sci. 60: 520–523. https://doi.org/10.1007/s11427-017-9021-5

Chen K., Y. Wang, R. Zhang, H. Zhang, and C. Gao, 2019 CRISPR/Cas Genome editing and precision plant breeding in agriculture. Annu. Rev. Plant Biol. 70: 667–697. https://doi.org/10.1146/annurev-arplant-050718-100049

Chiniquy D., P. Varanasi, T. Oh, J. Harholt, J. Katnelson, *et al.*, 2013 Three novel rice genes closely related to the *Arabidopsis IRX9*, *IRX9L*, and *IRX14* genes and their roles in xylan biosynthesis. Front. Plant Sci. 4: 1–13. https://doi.org/10.3389/fpls.2013.00083

Chong S. L., M. Derba-Maceluch, S. Koutaniemi, L. D. Gómez, S. J. McQueen-Mason, *et al.*, 2015 Active fungal GH115 aα-glucuronidase produced in *Arabidopsis thaliana* affects only the UX1-reactive glucuronate decorations on native glucuronoxylans. BMC Biotechnol. 15: 1–13. https://doi.org/10.1186/s12896-015-0154-8

Coleman H. D., J. Yan, and S. D. Mansfield, 2009 Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. Proc. Natl. Acad. Sci. 106: 13118–13123. https://doi.org/10.1073/pnas.0900188106

Cong L., F. A. Ran, D. Cox, S. Lin, R. Barretto, *et al.*, 2013 Multiplex genome engineering using CRISPR/Cas systems. Science. 339: 819–823. https://doi.org/10.1126/science.1231143

Cosgrove D. J., 2005 Growth of the plant cell wall. Nat. Rev. Mol. Cell Biol. 6: 850–861. https://doi.org/10.1038/nrm1746

Cox D. B. T., R. J. Platt, and F. Zhang, 2015 Therapeutic Genome editing: Prospects and challenges. Nat. Med. 21: 121–131. https://doi.org/10.1038/nm.3793

Crowell E. F., V. Bischoff, T. Desprez, A. Rolland, Y. D. Stierhof, *et al.*, 2009 Pausing of golgi bodies on microtubules regulates secretion of cellulose synthase complexes in *Arabidopsis*. Plant Cell 21: 1141–1154. https://doi.org/10.1105/tpc.108.065334

Danner E., S. Bashir, S. Yumlu, W. Wurst, B. Wefers, *et al.*, 2017 Control of gene editing by manipulation of DNA repair mechanisms. Mamm. Genome 28: 262–274. https://doi.org/10.1007/s00335-017-9688-5

Darvill J. E., M. Mcneil, A. G. Darvill, and P. Albersheim, 1980 Structure of Plant Cell Walls. Phytochemistry 66: 1135–1139. https://dx.doi.org/10.1104%2Fpp.66.6.1135

Delmer D. P., 1999 CELLULOSE BIOSYNTHESIS: Exciting times for a difficult field of study. Annu. Rev. Plant Physiol. Plant Mol. Biol. 50: 245–276. https://doi.org/10.1146/annurev.arplant.50.1.245

Derbyshire P., D. Ménard, P. Green, G. Saalbach, H. Buschmann, *et al.*, 2015 Proteomic analysis of microtubule interacting proteins over the course of xylem tracheary element formation in *Arabidopsis*. Plant Cell 27: 2709–2726. https://doi.org/10.1105/tpc.15.00314

Dima O., K. Morreel, B. Vanholme, H. Kim, J. Ralph, *et al.*, 2015 Small glycosylated lignin oligomers are stored in *Arabidopsis* leaf vacuoles. Plant Cell 27: 695–710. https://doi.org/10.1105/tpc.114.134643

Doblin M. S., I. Kurek, D. Jacob-Wilk, and D. P. Delmer, 2002 Cellulose biosynthesis in plants: From genes to rosettes. Plant Cell Physiol. 43: 1407–1420. https://doi.org/10.1093/pcp/pcf164

Donaldson L. A., 2001 Lignification and lignin topochemistry - An ultrastructural view. Phytochemistry 57: 859–873. https://doi.org/10.1016/S0031-9422(01)00049-8

Ebert B., C. Rautengarten, X. Guo, G. Xiong, S. Stonebloom, *et al.*, 2015 Identification and characterization of a golgi-localized UDP-xylose transporter family from *Arabidopsis*. Plant Cell 27: 1218–1227. https://doi.org/10.1105/tpc.114.133827

Ebringerová A., and T. Heinze, 2000b Xylan and xylan derivatives - Biopolymers with valuable properties, 1: Naturally occurring xylans structures, isolation procedures and properties. Macromol. Rapid Commun. 21: 542–556. https://doi.org/10.1002/1521-3927(20000601)21:9<542::AID-MARC542>3.0.CO;2-7

Elorriaga E., A. L. Klocko, C. Ma, and S. H. Strauss, 2018a Variation in mutation spectra among CRISPR/Cas9 mutagenized poplars. Front. Plant Sci. 9: 594. https://doi.org/10.3389/fpls.2018.00594

Endler A., and S. Persson, 2011 Cellulose synthases and synthesis in *Arabidopsis*. Mol. Plant 4: 199–211. https://doi.org/10.1093/mp/ssq079

Endo A., M. Masafumi, H. Kaya, and S. Toki, 2016 Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. Sci. Rep. 6: 1–9. https://doi.org/10.1038/srep38169

Espiñeira J. M., E. Novo Uzal, L. V. Gómez Ros, J. S. Carrión, F. Merino, *et al.*, 2011 Distribution of lignin monomers and the evolution of lignification among lower plants. Plant Biol. 13: 59–68. https://doi.org/10.1111/j.1438-8677.2010.00345.x

Eudes A., N. Zhao, N. Sathitsuksanoh, E. E. K. Baidoo, J. Lao, *et al.*, 2016 Expression of S-adenosylmethionine hydrolase in tissues synthesizing secondary cell walls alters specific methylated cell wall fractions and improves biomass digestibility. Front. Bioeng. Biotechnol. 4: 1–10. https://doi.org/10.3389/fbioe.2016.00058

Fan D., T. Liu, C. Li, B. Jiao, S. Li, *et al.*, 2015 Efficient CRISPR/Cas9-mediated targeted mutagenesis in *Populus* in the first generation. Sci. Rep. 5: 1–7. https://doi.org/10.1038/srep12217

Fan C., S. Feng, J. Huang, Y. Wang, L. Wu, *et al.*, 2017 *AtCesA8*-driven *OsSUS3* expression leads to largely enhanced biomass saccharification and lodging resistance by distinctively altering lignocellulose features in rice. Biotechnol. Biofuels 10: 1–12. https://doi.org/10.1186/s13068-017-0911-0

Fatland B. L., J. Ke, M. D. Anderson, W. I. Mentzen, W. C. Li, *et al.*, 2002 Molecular characterization of a heteromeric ATP-citrate lyase that generates cytosolic acetyl-coenzyme A in *Arabidopsis*. Plant Physiol. 130: 740–756. https://doi.org/10.1104/pp.008110

Fatland B. L., B. J. Nikolau, and E. S. Wurtele, 2005 Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP-citrate lyase in *Arabidopsis*. Plant Cell 17: 182–203. https://doi.org/10.1105/tpc.104.026211

Feng Z., B. Zhang, W. Ding, X. Liu, D. L. Yang, *et al.*, 2013 Efficient genome editing in plants using a CRISPR/Cas system. Cell Res. 23: 1229–1232. https://doi.org/10.1038/cr.2013.114

Fisher D., and R. Cyr, 1998 Extending the Microtubule/Microfibril paradigm. Cellulose synthesis is required for normal cortical microtubule alignment in elongating cells. Plant Physiol. 116: 1043–51. https://doi.org/10.1104/pp.116.3.1043

Flotte T. R., and G. Gao, 2019 Prime Editing: A Novel Cas9-reverse transcriptase fusion may revolutionize genome editing. Hum. Gene Ther. 30: 1445–1446. https://doi.org/10.1089/hum.2019.29098.trf

Fonfara I., H. Richter, M. BratoviÄ, A. Le Rhun, and E. Charpentier, 2016 The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature 532: 517–521. https://doi.org/10.1038/nature17945

Fujii S., T. Hayashi, and K. Mizuno, 2010 Sucrose synthase is an integral component of the cellulose synthesis machinery. Plant Cell Physiol. 51: 294–301. https://doi.org/10.1093/pcp/pcp190

Galbe M., and G. Zacchi, 2002 A review of the production of ethanol from softwood. Appl. Microbiol. Biotechnol. 59: 618–628. https://doi.org/10.1007/s00253-002-1058-9

Gao J., G. Wang, S. Ma, X. Xie, X. Wu, *et al.*, 2015 CRISPR/Cas9-mediated targeted mutagenesis in *Nicotiana tabacum*. Plant Mol. Biol. 87: 99–110. https://doi.org/10.1007/s11103-014-0263-0

Gao X., J. Chen, X. Dai, D. Zhang, and Y. Zhao, 2016 An effective strategy for reliably isolating heritable and Cas9-free *Arabidopsis* mutants generated by CRISPR/Cas9-mediated genome editing. Plant Physiol. 171: 1794–1800. https://doi.org/10.1104/pp.16.00663

Gao W., and K. Chen, 2017 Physical properties and thermal behavior of reconstituted tobacco sheet with precipitated calcium carbonate added in the coating process. Cellulose 24: 2581–2590. https://doi.org/10.1007/s10570-017-1270-5

Gardiner J. C., N. G. Taylor, and S. R. Turner, 2003 Control of cellulose synthase complex localization in developing xylem. Plant Cell 15: 1740–1748. https://doi.org/10.1105/tpc.012815

Garneau J. E., M. È. Dupuis, M. Villion, D. A. Romero, R. Barrangou, *et al.*, 2010 The CRISPR/cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468: 67–71. https://doi.org/10.1038/nature09523

Gasiunas G., R. Barrangou, P. Horvath, and V. Siksnys, 2012 Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc. Natl. Acad. Sci. 109: 2579–2586. https://doi.org/10.1073/pnas.1208507109

Gerber L., B. Zhang, M. Roach, U. Rende, A. Gorzsás, *et al.*, 2014 Deficient sucrose synthase activity in developing wood does not specifically affect cellulose biosynthesis, but causes an overall decrease in cell wall polymers. New Phytol. 203: 1220–1230. https://doi.org/10.1111/nph.12888

Giddings T. H., D. L. Brower, and L. Andrew Staehelin, 1980 Visualization of particle complexes in the plasma membrane of *Micrasterias Denticulata* associated with the formation of cellulose fibrils in primary and secondary cell walls. J. Cell Biol. 84: 327–339. https://doi.org/10.1083/jcb.84.2.327

Gilbert L. A., M. A. Horlbeck, B. Adamson, E. Jacqueline, Y. Chen, *et al.*, 2015 Genome-scale CRISPR-mediated control of gene repression and activation. Cell 159: 647–661. https://doi.org/10.1016/j.cell.2014.09.029

Gille S., A. de Souza, G. Xiong, M. Benz, K. Cheng, *et al.*, 2011 O-acetylation of *Arabidopsis* hemicellulose xyloglucan requires AXY4 or AXY4L proteins with a TBL an DUF231 domain. Plant Cell 23: 4041–4053. https://doi.org/10.1105/tpc.111.091728

Gille S., and M. Pauly, 2012 O-Acetylation of plant cell wall polysaccharides. Front. Plant Sci. 3: 1–7. https://doi.org/10.3389/fpls.2012.00012

Golberg A., M. Sack, J. Teissie, G. Pataro, U. Pliquett, *et al.*, 2016 Energy-efficient biomass processing with pulsed electric fields for bioeconomy and sustainable development. Biotechnol. Biofuels 9: 1–22. https://doi.org/10.1186/s13068-016-0508-z

Gonçalves V. M. F., D. V. Evtuguin, and M. R. M. Domingues, 2008 Structural characterization of the acetylated heteroxylan from the natural hybrid *Paulownia elongata/Paulownia fortunei*. Carbohydr. Res. 343: 256–266. https://doi.org/10.1016/j.carres.2007.11.002

González-Vila F. J., G. Almendros, J. C. Del Rio, F. Martín, A. Gutiérrez, *et al.*, 1999 Ease of delignification assessment of wood from different *Eucalyptus* species by pyrolysis (TMAH)-GC/MS and CP/MAS 13C-NMR spectrometry. J. Anal. Appl. Pyrolysis 49: 295–305. https://doi.org/10.1016/S0165-2370(98)00097-7

Grantham N. J., J. Wurman-Rodrich, O. M. Terrett, J. J. Lyczakowski, K. Stott, *et al.*, 2017 An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. Nat. Plants 3: 859–865. https://doi.org/10.1038/s41477-017-0030-8

Grönqvist S., T. K. Hakala, T. Kamppuri, M. Vehviläinen, T. Hänninen, *et al.*, 2014 Fibre porosity development of dissolving pulp during mechanical and enzymatic processing. Cellulose 21: 3667–3676. https://doi.org/10.1007/s10570-014-0352-x

Gu Y., N. Kaplinsky, M. Bringmann, A. Cobb, A. Carroll, *et al.*, 2010 Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. Proc. Natl. Acad. Sci. 107: 12866–12871. https://doi.org/10.1073/pnas.1007092107

Gutierrez R., J. J. Lindeboom, A. R. Paredez, A. M. C. Emons, and D. W. Ehrhardt, 2009 *Arabidopsis* cortical microtubules position cellulose synthase delivery to the plasma membrane and interact with cellulose synthase trafficking compartments. Nat. Cell Biol. 11: 797–806. https://doi.org/10.1038/ncb1886

Ha M. A., D. C. Apperley, B. W. Evans, I. Max Huxham, W. Gordon Jardine, *et al.*, 1998 Fine structure in cellulose microfibrils: NMR evidence from onion and quince. Plant J. 16: 183–190. https://doi.org/10.1046/j.1365-313X.1998.00291.x

Haan R. den, E. van Rensburg, S. H. Rose, J. F. Görgens, and W. H. van Zyl, 2015 Progress and challenges in the engineering of non-cellulolytic microorganisms for consolidated bioprocessing. Curr. Opin. Biotechnol. 33: 32–38. https://doi.org/10.1016/j.copbio.2014.10.003

Hahn-Hägerdal B., and E. Palmqvist, 2000 Fermentation of lignocellulosic hydrolysates. II: Inhibitors and mechanisms of inhibition. Bioresour. Technol. 74: 25–33. https://doi.org/10.1016/S0960-8524(99)00161-3

Hanson A. D., and S. Roje, 2001 One-carbon metabolism in higher plants. Annu. Rev. Plant Biol. 52: 119–137. https://doi.org/10.1146/annurev.arplant.53.100201.160729

Hao Z., and D. Mohnen, 2014 A review of xylan and lignin biosynthesis: Foundation for studying *Arabidopsis* irregular xylem mutants with pleiotropic phenotypes. Crit. Rev. Biochem. Mol. Biol. 49: 212–241. https://doi.org/10.3109/10409238.2014.889651

Helle S., D. Cameron, J. Lam, B. White, and S. Duff, 2003 Effect of inhibitory compounds found in biomass hydrolysates on growth and xylose fermentation by a genetically engineered strain of *S. cerevisiae*. Enzyme Microb. Technol. 33: 786–792. https://doi.org/10.1016/S0141-0229(03)00214-X

Hemphill J., E. K. Borchardt, K. Brown, A. Asokan, and A. Deiters, 2015 Optical control of CRISPR/Cas9 gene editing. J. Am. Chem. Soc. 137: 5642–5645. https://doi.org/10.1021/ja512664v

Henriksson G., M. Christiernin, and R. Agnemo, 2005 Monocomponent endoglucanase treatment increases the reactivity of softwood sulphite dissolving pulp. J. Ind. Microbiol. Biotechnol. 32: 211–214. https://doi.org/10.1007/s10295-005-0220-7

Himmel M. E., S. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, *et al.*, 2007 Biomass recalcitrance: Engineering plants and enzymes for biofuels production. Nature 454: 804–807. https://doi.org/10.1126/science.1137016

Hori C., N. Takata, P. Y. Lam, Y. Tobimatsu, S. Nagano, et al., 2020 Identifying transcription factors that reduce wood recalcitrance and improve enzymatic degradation of xylem cell wall in Populus. Sci. Rep. 10: 1–13. https://doi.org/10.1038/s41598-020-78781-6

Hu X., C. Wang, Q. Liu, Y. Fu, and K. Wang, 2017 Targeted mutagenesis in rice using CRISPR-Cpf1 system. J. Genet. Genomics 44: 71–73. https://doi.org/10.1016/j.jgg.2016.12.001

Ibar C., and A. Orellana, 2007 The import of S-Adenosylmethionine into the Golgi apparatus is required for the methylation of homogalacturonan. Plant Physiol. 145: 504–512. https://doi.org/10.1104/pp.107.104679

Jacobs T. B., P. R. LaFayette, R. J. Schmitz, and W. A. Parrott, 2015 Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnol. 15: 1–10. https://doi.org/10.1186/s12896-015-0131-2

Jaini R., P. Wang, N. Dudareva, C. Chapple, and J. A. Morgan, 2017 Targeted Metabolomics of the Phenylpropanoid Pathway in Arabidopsis thaliana using Reversed Phase Liquid Chromatography Coupled with Tandem Mass Spectrometry. Phytochem. Anal. 28: 267–276. https://doi.org/10.1002/pca.2672

Jensen J. K., N. R. Johnson, and C. G. Wilkerson, 2014 *Arabidopsis thaliana* IRX10 and two related proteins from psyllium and *Physcomitrella patens* are xylan xylosyltransferases. Plant J. 80: 207–215. https://doi.org/10.1111/tpj.12641

Jiang W., H. Zhou, H. Bi, M. Fromm, B. Yang, *et al.*, 2013 Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. Nucleic Acids Res. 41: e188–e188. https://doi.org/10.1093/nar/gkt780

Jinek M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, *et al.*, 2012 A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 337: 816–821. https://doi.org/10.1126/science.1225829

Johnston S. L., R. Prakash, N. J. Chen, M. H. Kumagai, H. M. Turano, *et al.*, 2013 An enzyme activity capable of endotransglycosylation of heteroxylan polysaccharides is present in plant primary cell walls. Planta 237: 173–187. https://doi.org/10.1007/s00425-012-1766-z

Kabel M. A., H. van den Borne, J. P. Vincken, A. G. J. Voragen, and H. A. Schols, 2007 Structural differences of xylans affect their interaction with cellulose. Carbohydr. Polym. 69: 94–105. https://doi.org/10.1016/j.carbpol.2006.09.006

Kaneda M., K. H. Rensing, J. C. T. Wong, B. Banno, S. D. Mansfield, *et al.*, 2008 Tracking monolignols during wood development in lodgepole pine. Plant Physiol. 147: 1750–1760. https://doi.org/10.1104/pp.108.121533

Kaneda M., M. Schuetz, B. S. P. Lin, C. Chanis, B. Hamberger, *et al.*, 2011 ABC transporters coordinately expressed during lignification of *Arabidopsis* stems include a set of ABCBs associated with auxin transport. J. Exp. Bot. 62: 2063–2077. https://doi.org/10.1093/jxb/erq416

Kang X., A. Kirui, M. C. Dickwella Widanage, F. Mentink-Vigier, D. J. Cosgrove, *et al.*, 2019 Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. Nat. Commun. 10: 1–9. https://doi.org/10.1038/s41467-018-08252-0

Kawamoto H., 2017 Lignin pyrolysis reactions. J. Wood Sci. 63: 117–132. https://doi.org/10.1007/s10086-016-1606-z

Keppler B. D., and A. M. Showalter, 2010 IRX14 and IRX14-LIKE, two glycosyl transferases involved in glucuronoxylan biosynthesis and drought tolerance in *Arabidopsis*. Mol. Plant 3: 834–841. https://doi.org/10.1093/mp/ssq028

Kim H., S. T. Kim, J. Ryu, B. C. Kang, J. S. Kim, *et al.*, 2017 CRISPR/Cpf1-mediated DNA-free plant genome editing. Nat. Commun. 8: 1–7. https://doi.org/10.1038/ncomms14406

Kingsley Ogemdi I., 2019 Combustion of Wood by Pyrolysis: A Review. Int. J. Atmos. Ocean. Sci. 3: 1. https://doi.org/10.11648/j.ijaos.20190301.11

Kirschner M. W., 2005 The meaning of systems biology. Cell 121: 503–504. https://doi.org/10.1016/j.cell.2005.05.005

Kleinstiver B. P., S. Q. Tsai, M. S. Prew, N. T. Nguyen, M. M. Welch, *et al.*, 2016 Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. Nat. Biotechnol. 34: 869–874. https://doi.org/10.1038/nbt.3620

Klemm D., B. Heublein, H. P. Fink, and A. Bohn, 2005 Cellulose: Fascinating biopolymer and sustainable raw material. Angew. Chemie - Int. Ed. 44: 3358–3393. https://doi.org/10.1002/anie.200460587

Ko J. H., E. P. Beers, and K. H. Han, 2006 Global comparative transcriptome analysis identifies gene network regulating secondary xylem development in *Arabidopsis thaliana*. Mol. Genet. Genomics 276: 517–531. https://doi.org/10.1007/s00438-006-0157-1

Köhnke T., Å. Östlund, and H. Brelid, 2011 Adsorption of arabinoxylan on cellulosic surfaces: Influence of degree of substitution and substitution pattern on adsorption characteristics. Biomacromolecules 12: 2633–2641. https://doi.org/10.1021/bm200437m

Komor A. C., Y. B. Kim, M. S. Packer, J. A. Zuris, and D. R. Liu, 2016 Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533: 420–424. https://doi.org/10.1038/nature17946

Konn J., L. Vähäsalo, A. Pranovich, and B. Holmbom, 2006 De-esterification and sulfonation in spruce CTMP: Effects on pulp and paper properties. Holzforschung 60. http://dx.doi.org/10.1515/HF.2006.056

Kopycki J. G., D. Rauh, A. A. Chumanevich, P. Neumann, T. Vogt, *et al.*, 2008 Biochemical and structural analysis of substrate promiscuity in plant Mg^{2+}-dependent O-methyltransferases. J. Mol. Biol. 378: 154–164. https://doi.org/10.1016/j.jmb.2008.02.019

Kotake T., R. Takata, R. Verma, M. Takaba, D. Yamaguchi, *et al.*, 2009 Bifunctional cytosolic UDP-glucose 4-epimerases catalyse the interconversion between UDP-D-xylose and UDP-L-arabinose in plants. Biochem. J. 424: 169–177. https://doi.org/10.1042/BJ20091025

Kuang B., X. Zhao, C. Zhou, W. Zeng, J. Ren, *et al.*, 2016 Role of UDP-Glucuronic acid decarboxylase in xylan biosynthesis in *Arabidopsis*. Mol. Plant 9: 1119–1131. https://doi.org/10.1016/j.molp.2016.04.013

Kumar M., L. Campbell, and S. Turner, 2016 Secondary cell walls: Biosynthesis and manipulation. J. Exp. Bot. 67: 515–531. https://doi.org/10.1093/jxb/erv533

Kumar M., L. Mishra, P. Carr, M. Pilling, P. Gardner, *et al.*, 2018 Exploiting CELLULOSE SYNTHASE (CESA) class specificity to probe cellulose Microfibril Biosynthesis. Plant Physiol. 177: 151–167. https://doi.org/10.1104/pp.18.00263

Laffend K., 1967 The effect of acetyl content of glucomannan on its sorption onto cellulose and on its beater additive properties. Tappi 51: 141–143. http://hdl.handle.net/1853/5606

Langan P., S. Gnanakaran, K. D. Rector, N. Pawley, D. T. Fox, *et al.*, 2011 Exploring new strategies for cellulosic biofuels production. Energy Environ. Sci. 4: 3820–3833. https://doi.org/10.1039/c1ee01268a

Lee C., Q. Teng, W. Huang, R. Zhong, and Z. H. Ye, 2009 The F8H glycosyltransferase is a functional paralog of FRA8 involved in glucuronoxylan biosynthesis in *Arabidopsis*. Plant Cell Physiol. 50: 812–827. https://doi.org/10.1093/pcp/pcp025

Lee C., Q. Teng, W. Huang, R. Zhong, and Z.-H. Ye, 2010 The *Arabidopsis* Family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. Plant Physiol. 153: 526–541. https://doi.org/10.1104/pp.110.155309

Lee C., Q. Teng, R. Zhong, and Z. H. Ye, 2011 The four *Arabidopsis REDUCED WALL ACETYLATION* genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. Plant Cell Physiol. 52: 1289–1301. https://doi.org/10.1093/pcp/pcr075

Lee C., Q. Teng, R. Zhong, and Z. H. Ye, 2012a *Arabidopsis* GUX proteins are glucuronyltransferases responsible for the addition of glucuronic acid side chains onto xylan. Plant Cell Physiol. 53: 1204–1216. https://doi.org/10.1093/pcp/pcs064

Lee C., Q. Teng, R. Zhong, Y. Yuan, M. Haghighat, *et al.*, 2012b Three *Arabidopsis* DUF579 domain-containing GXM proteins are methyltransferases catalyzing 4-O-methylation of glucuronic acid on xylan. Plant Cell Physiol. 53: 1934–1949. https://doi.org/10.1093/pcp/pcs138

Lee C., R. Zhong, and Z. H. Ye, 2012c *Arabidopsis* family GT43 members are xylan xylosyltransferases required for the elongation of the xylan backbone. Plant Cell Physiol. 53: 135–143. https://doi.org/10.1093/pcp/pcr158

Lee C., R. Zhong, and Z. H. Ye, 2012d Biochemical characterization of xylan xylosyltransferases involved in wood formation in poplar. Plant Signal. Behav. 7. https://doi.org/10.4161/psb.19269

Lee C., Q. Teng, R. Zhong, Y. Yuan, and Z. Ye, 2014 Functional roles of rice glycosyltransferase family GT43 in xylan biosynthesis. Plant Signal. Behav. e27809: 1–13. https://doi.org/10.4161/psb.27809

Lei L., S. Li, J. Du, L. Bashline, and Y. Gu, 2013 CELLULOSE SYNTHASE INTERACTIVE3 regulates cellulose biosynthesis in both a microtubule-dependent and microtubule-independent manner in *Arabidopsis*. Plant Cell 25: 4912–4923. https://doi.org/10.1105/tpc.113.116715

Li X., and C. Chapple, 2010 Understanding lignification: Challenges beyond monolignol biosynthesis. Plant Physiol. 154: 449–452. https://doi.org/10.1104/pp.110.162842

Li S., L. Lei, C. R. Somerville, and Y. Gu, 2012 Cellulose synthase interactive protein 1 (CSI1) links microtubules and cellulose synthase complexes. Proc. Natl. Acad. Sci. 109: 185–190. https://doi.org/10.1073/pnas.1118560109

Li L., S. Hill-Skinner, S. Liu, D. Beuchle, H. M. Tang, *et al.*, 2015 The maize brown *midrib4* (*bm4*) gene encodes a functional folylpolyglutamate synthase. Plant J. 81: 493–504. https://doi.org/10.1111/tpj.12745

Li J., Y. Sun, J. Du, Y. Zhao, and L. Xia, 2017 Generation of targeted point mutations in rice by a modified CRISPR/Cas9 system. Mol. Plant 10: 526–529. https://doi.org/10.1016/j.molp.2016.12.001

Li L., K. Wei, G. Zheng, X. Liu, S. Chen, *et al.*, 2018a CRISPR-Cpf1-assisted multiplex genome editing and transcriptional repression in *Streptomyces*. Appl. Environ. Microbiol. 84: e00827-18. https://doi.org/10.1128/AEM.00827-18

Li X., Y. Wang, Y. Liu, B. Yang, X. Wang, *et al.*, 2018b Base editing with a Cpf1-cytidine deaminase fusion. Nat. Biotechnol. 36: 324–327. https://doi.org/10.1038/nbt.4102

Lin Q., Y. Zong, C. Xue, S. Wang, S. Jin, *et al.*, 2020 Prime genome editing in rice and wheat. Nat. Biotechnol. 38: 582–585. https://doi.org/10.1038/s41587-020-0455-x

Liu D., R. Hu, K. J. Palla, G. A. Tuskan, and X. Yang, 2016a Advances and perspectives on the use of CRISPR/Cas9 systems in plant genomics research. Curr. Opin. Plant Biol. 30: 70–77. https://doi.org/10.1016/j.pbi.2016.01.007

Liu Z., R. Schneider, C. Kesten, Y. Zhang, M. Somssich, *et al.*, 2016b Cellulose-microtubule uncoupling proteins prevent lateral displacement of microtubules during cellulose synthesis in *Arabidopsis*. Dev. Cell 38: 305–315. https://doi.org/10.1016/j.devcel.2016.06.032

Liu X., C. Xie, H. Si, and J. Yang, 2017 CRISPR/Cas9-mediated genome editing in plants. Methods 121–122: 94–102. https://doi.org/10.1016/j.ymeth.2017.03.009

Lowder L., A. Malzahn, and Y. Qi, 2016 Rapid evolution of manifold CRISPR systems for plant genome editing. Front. Plant Sci. 7: 1–12. https://doi.org/10.3389/fpls.2016.01683

Lunin V., H.-T. Wang, V. Bharadwaj, P. M. Alahuhta, M. J. Pena, *et al.*, 2020 Molecular mechanism of polysaccharide acetylation by the *Arabidopsis* xylan O-acetyltransferase XOAT1. Plant Cell. https://doi.org/10.1105/tpc.20.00028

Macaya-Sanz D., J. G. Chen, U. C. Kalluri, W. Muchero, T. J. Tschaplinski, *et al.*, 2017 Agronomic performance of *Populus deltoides* trees engineered for biofuel production. Biotechnol. Biofuels 10: 1–13. https://doi.org/10.1186/s13068-017-0934-6

Maeder M. L., S. J. Linder, V. M. Cascio, Y. Fu, Q. H. Ho, *et al.*, 2013 CRISPR RNA–guided activation of endogenous human genes. Nat. Methods 10: 977–979. https://doi.org/10.1038/nmeth.2598

Mali P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, *et al.*, 2013 RNA-guided human genome engineering via Cas9. Science. 339: 823–826. https://doi.org/10.1126/science.1232033

Manabe Y., M. Nafisi, Y. Verhertbruggen, C. Orfila, S. Gille, *et al.*, 2011 Loss-of-Function mutation of *REDUCED WALL ACETYLATION2* in *Arabidopsis* leads to reduced cell wall acetylation and increased resistance to *Botrytis cinerea*. Plant Physiol. 155: 1068–1078. https://doi.org/10.1104/pp.110.168989

Mao Y., H. Zhang, N. Xu, B. Zhang, F. Gou, *et al.*, 2013 Application of the CRISPR–Cas system for efficient genome engineering in plants. Mol. Plant 6: 2008–2011. https://doi.org/10.1093/mp/sst121

Mao Y., J. R. Botella, Y. Liu, and J. K. Zhu, 2019 Gene editing in plants: Progress and challenges. Natl. Sci. Rev. 6: 421–437. https://doi.org/10.1093/nsr/nwz005

Martínez-Abad A., J. Berglund, G. Toriz, P. Gatenholm, G. Henriksson, *et al.*, 2017 Regular motifs in xylan modulate molecular flexibility and interactions with cellulose surfaces. Plant Physiol. 175: 1579–1592. https://doi.org/10.1104/pp.17.01184

Martone P. T., J. M. Estevez, F. Lu, K. Ruel, M. W. Denny, *et al.*, 2009 Discovery of Lignin in Seaweed Reveals Convergent Evolution of Cell-Wall Architecture. Curr. Biol. 19: 169–175. https://doi.org/10.1016/j.cub.2008.12.031

Marzec M., A. Brąszewska-Zalewska, and G. Hensel, 2020 Prime editing: A new way for genome editing. Trends Cell Biol. 30: 257–259. https://doi.org/10.1016/j.tcb.2020.01.004

McCartney L., A. W. Blake, J. Flint, D. N. Bolam, A. B. Boraston, *et al.*, 2006 Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules. Proc. Natl. Acad. Sci. 103: 4765–4770. https://doi.org/10.1073/pnas.0508887103

McCarty N. S., A. E. Graham, L. Studená, and R. Ledesma-Amaro, 2020 Multiplexed CRISPR technologies for gene editing and transcriptional regulation. Nat. Commun. 11: 1–13. https://doi.org/10.1038/s41467-020-15053-x

McNamara J. T., J. L. W. Morgan, and J. Zimmer, 2015 A Molecular description of cellulose biosynthesis. Annu. Rev. Biochem. 84: 895–921. https://doi.org/10.1146/annurev-biochem-060614-033930

Meents M. J., Y. Watanabe, and A. L. Samuels, 2018 The cell biology of secondary cell wall biosynthesis. Ann. Bot. 1–19. https://doi.org/10.1093/aob/mcy005

Meester B. De, B. Madariaga Calderón, L. de Vries, J. Pollier, G. Goeminne, *et al.*, 2020 Tailoring poplar lignin without yield penalty by combining a null and haploinsufficient *CINNAMOYL-CoA REDUCTASE2* allele. Nat. Commun. 11: 5020. https://doi.org/10.1038/s41467-020-18822-w

Menon V., and M. Rao, 2012 Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept. Prog. Energy Combust. Sci. 38: 522–550. https://doi.org/10.1016/j.pecs.2012.02.002

Mentzen W. I., and E. S. Wurtele, 2008 Regulon organization of *Arabidopsis*. BMC Plant Biol. 8: 1–22. https://doi.org/10.1186/1471-2229-8-99

Mey M. De, S. De Maeseneire, W. Soetaert, and E. Vandamme, 2007 Minimizing acetate formation in *E. coli* fermentations. J. Ind. Microbiol. Biotechnol. 34: 689–700. https://doi.org/10.1007/s10295-007-0244-2

Miao Y.-C., and C.-J. Liu, 2010 ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. Proc. Natl. Acad. Sci. 107: 22728–22733. https://doi.org/10.1073/pnas.1007747108

Ming-Ju Chen, Kreuter J. Y.-T. K., 1996 Nanoparticles and microparticles for drug and vaccine delivery. J. Anat. 189 Pt 3: 503–505. https://doi.org/10.1002/bit

Mizrachi E., V. J. Maloney, J. Silberbauer, C. A. Hefer, D. K. Berger, *et al.*, 2015 Investigating the molecular underpinnings underlying morphology and changes in carbon partitioning during tension wood formation in *Eucalyptus*. New Phytol. 206: 1351–1363. https://doi.org/10.1111/nph.13152

Mohan D., C. U. Pittman, and P. H. Steele, 2006 Pyrolysis of wood/biomass for bio-oil: A critical review. Energy and Fuels 20: 848–889. https://doi.org/10.1021/ef0502397

Mohnen D., 2008 Pectin structure and biosynthesis. Curr. Opin. Plant Biol. 11: 266–277. https://doi.org/10.1016/j.pbi.2008.03.006

Mortimer J. C., G. P. Miles, D. M. Brown, Z. Zhang, M. P. Segura, *et al.*, 2010 Absence of branches from xylan in *Arabidopsis gux* mutants reveals potential for simplification of lignocellulosic biomass. Proc. Natl. Acad. Sci. 107: 17409–17414. https://doi.org/10.1073/pnas.1005456107

Mosier N., C. Wyman, B. Dale, R. Elander, Y. Y. Lee, *et al.*, 2005 Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresour. Technol. 96: 673–86. https://doi.org/10.1016/j.biortech.2004.06.025

Mueller S. C., and R. Malcolm Brown, 1980 Evidence for an intramembrane component associated with a cellulose microfibril-synthesizing complex in higher plants. J. Cell Biol. 84: 315–326. https://doi.org/10.1083/jcb.84.2.315

Nasatto P. L., F. Pignon, J. L. M. Silveira, M. E. R. Duarte, M. D. Noseda, *et al.*, 2015 Methylcellulose, a cellulose derivative with original physical properties and extended applications. Polymers (Basel). 7: 777–803. https://doi.org/10.3390/polym7050777

Nguyen Q. A., M. P. Tucker, B. L. Boynton, F. A. Keller, and D. J. Schell, 2008 Dilute acid pretreatment of softwoods. Appl. Biochem. Biotechnol. 70–72: 77–87. https://doi.org/10.1007/bf02920125

Nihongaki Y., F. Kawano, T. Nakajima, and M. Sato, 2015 Photoactivatable CRISPR-Cas9 for optogenetic genome editing. Nat. Biotechnol. 33: 755–760. https://doi.org/10.1038/nbt.3245

Nishida K., T. Arazoe, N. Yachie, S. Banno, M. Kakimoto, *et al.*, 2016 Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. Science. 353. https://doi.org/10.1126/science.aaf8729

Nunes C. A., C. F. Lima, L. C. A. Barbosa, J. L. Colodette, A. F. G. Gouveia, et al., 2010 Determination of *Eucalyptus* spp lignin S/G ratio: A comparison between methods. Bioresour. Technol. 101: 4056–4061. https://doi.org/10.1016/j.biortech.2010.01.012

Oakes B. L., D. C. Nadler, A. Flamholz, C. Fellmann, B. T. Staahl, *et al.*, 2016 Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. Nat. Biotechnol. 34: 646–651. https://doi.org/10.1038/nbt.3528

Oliver D. J., B. J. Nikolau, and E. S. Wurtele, 2009 Acetyl-CoA-Life at the metabolic nexus. Plant Sci. 176: 597–601. https://doi.org/10.1016/j.plantsci.2009.02.005

Pan X., D. Xie, N. Gilkes, D. J. Gregg, and J. N. Saddler, 2005 Strategies to enhance the enzymatic hydrolysis of pretreated softwood with high residual lignin content. Appl. Biochem. Biotechnol. 121: 1069–1079. https://doi.org/10.1385/abab:124:1-3:1069

Paredez A. R., S. Persson, D. W. Ehrhardt, and C. R. Somerville, 2008 Genetic evidence that cellulose synthase activity influences microtubule cortical array organization. Plant Physiol. 147: 1723–1734. https://doi.org/10.1104/pp.108.120196

Park S., S. Oh, and K.-H. Han, 2004 Large-scale computational analysis of poplar ESTs reveals the repertoire and unique features of expressed genes in the poplar genome. Mol. Breed. 14: 429–440. https://doi.org/10.1007/s11032-004-0603-x

Pauly M., and H. V. Scheller, 2000 O-acetylation of plant cell wall polysaccharides: Identification and partial characterization of a rhamnogalacturonan O-acetyl-transferase from potato suspension-cultured cells. Planta 210: 659–667. https://doi.org/10.1007/s004250050057

Pauly M., and K. Keegstra, 2008 Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J. 54: 559–568. https://doi.org/10.1111/j.1365-313X.2008.03463.x

Pawar P. M.-A., S. Koutaniemi, M. Tenkanen, and E. J. Mellerowicz, 2013 Acetylation of woody lignocellulose: significance and regulation. Front. Plant Sci. 4: 1–8. https://doi.org/10.3389/fpls.2013.00118

Pawar P. M. A., M. Derba-Maceluch, S. L. Chong, L. D. Gómez, E. Miedes, *et al.*, 2016 Expression of fungal acetyl xylan esterase in Arabidopsis thaliana improves saccharification of stem lignocellulose. Plant Biotechnol. J. 14: 387–397. https://doi.org/10.1111/pbi.12393

Pawar P. M. A., M. Derba-Maceluch, S. L. Chong, M. L. Gandla, S. S. Bashar, *et al.*, 2017a *In muro* deacetylation of xylan affects lignin properties and improves saccharification of aspen wood. Biotechnol. Biofuels 10: 1–11. https://doi.org/10.1186/s13068-017-0782-4

Pawar P. M. A., C. Ratke, V. K. Balasubramanian, S. L. Chong, M. L. Gandla, *et al.*, 2017b Downregulation of *RWA* genes in hybrid aspen affects xylan acetylation and wood saccharification. New Phytol. 214: 1491–1505. https://doi.org/10.1111/nph.14489

Pear J. R., Y. Kawagoe, W. E. Schreckengost, D. P. Delmer, and D. M. Stalker, 2002 Higher plants contain homologs of the bacterial *celA* genes encoding the catalytic subunit of cellulose synthase. Proc. Natl. Acad. Sci. 93: 12637–12642. https://doi.org/10.1073/pnas.93.22.12637

Pellny T. K., A. Patil, A. J. Wood, J. Freeman, K. Halsey, *et al.*, 2020 Loss of *TaIRX9b* gene function in wheat decreases chain length and amount of arabinoxylan in grain but increases cross-linking. Plant Biotechnol. J. https://doi.org/10.1111/pbi.13393

Pena M. J., R. Zhong, G.-K. Zhou, E. A. Richardson, M. A. O'Neill, *et al.*, 2007 *Arabidopsis irregular xylem8* and *irregular xylem9*: Implications for the Complexity of Glucuronoxylan Biosynthesis. Plant Cell 19: 549–563. https://doi.org/10.1105/tpc.106.049320

Pereira C. S., R. L. Silveira, P. Dupree, and M. S. Skaf, 2017 Effects of xylan side-chain substitutions on xylan-cellulose interactions and implications for thermal pretreatment of cellulosic biomass. Biomacromolecules 18: 1311–1321. https://doi.org/10.1021/acs.biomac.7b00067

Perochon A., D. Aldon, J. P. Galaud, and B. Ranty, 2011 Calmodulin and calmodulin-like proteins in plant calcium signaling. Biochimie 93: 2048–2053. https://doi.org/10.1016/j.biochi.2011.07.012

Persia D., G. Cai, C. Del Casino, C. Faleri, M. T. M. Willemse, *et al.*, 2008 Sucrose synthase is associated with the cell wall of tobacco pollen tubes. Plant Physiol. 147: 1603–1618. https://doi.org/10.1104/pp.108.115956

Persson S., A. Paredez, A. Carroll, H. Palsdottir, M. Doblin, *et al.*, 2007 Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis*. Proc. Natl. Acad. Sci. U. S. A. 104: 15566–15571. https://doi.org/10.1073/pnas.0706592104

Petrik D. L., T. Tryfona, P. Dupree, and C. T. Anderson, 2020 *BdGT43B2* functions in xylan biosynthesis and is essential for seedling survival in *Brachypodium distachyon*. Plant Direct 4: 1–16. https://doi.org/10.1002/pld3.216

Piatek A., Z. Ali, H. Baazim, L. Li, A. Abulfaraj, *et al.*, 2015 RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. Plant Biotechnol. J. 13: 578–589. https://doi.org/10.1111/pbi.12284

Plomion C., G. Leprovost, and A. Stokes, 2001 Wood formation in trees. Plant Physiol. 127: 1513–1523. https://doi.org/10.1104/pp.010816

Polstein L. R., and C. A. Gersbach, 2015 A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. Nat. Chem. Biol. 11: 198–200. https://doi.org/10.1038/nchembio.1753

Qi L. S., M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman, *et al.*, 2013 Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152: 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022

Qing Q., and C. E. Wyman, 2011 Supplementation with xylanase and β-xylosidase to reduce xylo-oligomer and xylan inhibition of enzymatic hydrolysis of cellulose and pretreated corn stover. Biotechnol. Biofuels 4: 1–12. https://doi.org/10.1186/1754-6834-4-18

Ragauskas A. J., G. T. Beckham, M. J. Biddy, R. Chandra, F. Chen, *et al.*, 2014 Lignin valorization: Improving lignin processing in the biorefinery. Science. 344. https://doi.org/10.1126/science.1246843

Raimund T., and T. Oliver, 1996 Cloning of an enzyme that synthesizes a key nucleotide-sugar precursor of hemicellulose biosynthesis from soybean: UDP-Glucose dehydrogenase. Plant Physiol. 112: 1127–1134. https://doi.org/10.1104/pp.112.3.1127

Ran F. A., P. D. Hsu, C. Y. Lin, J. S. Gootenberg, S. Konermann, *et al.*, 2013a Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. Cell 154: 1380–1389. https://doi.org/10.1016/j.cell.2013.08.021

Ran F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, *et al.*, 2013b Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8: 2281–2308. https://doi.org/10.1038/nprot.2013.143

Ratke C., B. K. Terebieniec, S. Winestrand, M. Derba-Maceluch, T. Grahn, *et al.*, 2018 Downregulating aspen xylan biosynthetic GT43 genes in developing wood stimulates growth via reprograming of the transcriptome. New Phytol. 219: 230–245. https://doi.org/10.1111/nph.15160

Rautengarten C., B. Ebert, T. Herter, C. J. Petzold, T. Ishii, *et al.*, 2011 The interconversion of UDP-arabinopyranose and UDP-arabinofuranose is indispensable for plant development in *Arabidopsis*. Plant Cell 23: 1373–1390. https://doi.org/10.1105/tpc.111.083931

Ren B., F. Yan, Y. Kuang, N. Li, D. Zhang, *et al.*, 2017 A CRISPR/Cas9 toolkit for efficient targeted base editing to induce genetic variations in rice. Sci. China Life Sci. 60: 516–519. https://doi.org/10.1007/s11427-016-0406-x

Rennie E. A., and H. V. Scheller, 2014 Xylan biosynthesis. Curr. Opin. Biotechnol. 26: 100–107. https://doi.org/10.1016/j.copbio.2013.11.013

Restolho J. A., A. Prates, M. N. de Pinho, and M. D. Afonso, 2009 Sugars and lignosulphonates recovery from *Eucalyptus* spent sulphite liquor by membrane processes. Biomass and Bioenergy 33: 1558–1566. https://doi.org/10.1016/j.biombioe.2009.07.022

Río J. C. Del, A. Gutiérrez, M. Hernando, P. Landín, J. Romero, *et al.*, 2005 Determining the influence of eucalypt lignin composition in paper pulp yield using Py-GC/MS. J. Anal. Appl. Pyrolysis 74: 110–115. https://doi.org/10.1016/j.jaap.2004.10.010

Río J. C. Del, G. Marques, J. Rencoret, Á. T. Martínez, and A. Gutiérrez, 2007 Occurrence of naturally acetylated lignin units. J. Agric. Food Chem. 55: 5461–5468. https://doi.org/10.1021/jf0705264

Roje S., 2006 S-Adenosyl-l-methionine: Beyond the universal methyl group donor. Phytochemistry 67: 1686–1698. https://doi.org/10.1016/j.phytochem.2006.04.019

Ross P., R. Mayer, and M. Benziman, 1991 Cellulose biosynthesis and function in bacteria. Microbiol. Rev. 55: 35–58. https://doi.org/10.1128/mmbr.55.1.35-58.1991

Sapranauskas R., G. Gasiunas, C. Fremaux, R. Barrangou, P. Horvath, *et al.*, 2011 The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in Escherichia coli. Nucleic Acids Res. 39: 9275–9282. https://doi.org/10.1093/nar/gkr606

Sassner P., C. G. Mårtensson, M. Galbe, and G. Zacchi, 2008 Steam pretreatment of H_2SO_4-impregnated *Salix* for the production of bioethanol. Bioresour. Technol. 99: 137–145. https://doi.org/10.1016/j.biortech.2006.11.039

Sauter M., B. Moffatt, M. C. Saechao, R. Hell, and M. Wirtz, 2013 Methionine salvage and S-adenosylmethionine: essential links between sulfur, ethylene and polyamine biosynthesis . Biochem. J. 451: 145–154. https://doi.org/10.1042/Bj20121744

Saxena I. M., and R. M. Brown, 2005 Cellulose biosynthesis: Current views and evolving concepts. Ann. Bot. 96: 9–21. https://doi.org/10.1093/aob/mci155

Scheller H. V., and P. Ulvskov, 2010 Hemicelluloses. Annu. Rev. Plant Biol. 61: 263–289. https://doi.org/10.1146/annurev-arplant-042809-112315

Schneider R., L. Tang, E. R. Lampugnani, S. Barkwill, R. Lathe, *et al.*, 2017 Two complementary mechanisms underpin cell wall patterning during xylem vessel development. Plant Cell 29: 2433–2449. https://doi.org/10.1105/tpc.17.00309

Schultink A., D. Naylor, M. Dama, and M. Pauly, 2015 The role of the plant-specific ALTERED XYLOGLUCAN9 protein in *Arabidopsis* cell wall polysaccharide o-acetylation. Plant Physiol. 167: 1271–1283. https://doi.org/10.1104/pp.114.256479

Schütte G., 2018 What kind of innovation policy does the bioeconomy need? N. Biotechnol. 40: 82–86. https://doi.org/10.1016/j.nbt.2017.04.003

Senneca O., 2007 Kinetics of pyrolysis, combustion and gasification of three biomass fuels. Fuel Process. Technol. 88: 87–97. https://doi.org/10.1016/j.fuproc.2006.09.002

Shen B., C. Li, and M. C. Tarczynski, 2002 High free-methionine and decreased lignin content result from a mutation in the Arabidopsis S-adenosyl-L-methionine synthetase 3 gene. Plant J. 29: 371–380. https://doi.org/10.1046/j.1365-313X.2002.01221.x

Sibout R., and H. Höfte, 2012 Plant cell biology: The ABC of monolignol transport. Curr. Biol. 22: 533–535. https://doi.org/10.1016/j.cub.2012.05.005

Smith R. A., M. Schuetz, M. Roach, S. D. Mansfield, B. Ellis, *et al.*, 2013 Neighboring parenchyma cells contribute to *Arabidopsis* xylem lignification, while lignification of interfascicular fibers is cell autonomous. Plant Cell 25: 3988–3999. https://doi.org/10.1105/tpc.113.117176

Smith P. J., H. T. Wang, W. S. York, M. J. Peña, and B. R. Urbanowicz, 2017 Designer biomass for next - generation biorefineries : leveraging recent insights into xylan structure and biosynthesis. Biotechnol. Biofuels 10: 1–14. https://doi.org/10.1186/s13068-017-0973-z

Smith P. J., M. A. O'Neill, J. Backe, W. S. York, M. J. Peña, *et al.*, 2020 Analytical techniques for determining the role of domain of unknown function 579 proteins in the synthesis of O-methylated plant polysaccharides. SLAS Technol. https://doi.org/10.1177/2472630320912692

Söderström J., L. Pilcher, M. Galbe, and G. Zacchi, 2002 Two-step steam pretreatment of softwood with SO_2 impregnation for ethanol production, pp. 5–21 in *Applied Biochemistry and Biotechnology*

Spindler D., C. Wyman, and K. Grohmann, 1990 Evaluation of pretreated herbaceous crops for the simultaneous saccharification and fermentation process. Appl. Biochem. Biotechnol. 24–25: 275–286. https://doi.org/10.1007/BF02920252

Srivastava A. C., F. Chen, T. Ray, S. Pattathil, M. J. Peña, *et al.*, 2015 Loss of function of folylpolyglutamate synthetase 1 reduces lignin content and improves cell wall digestibility in *Arabidopsis*. Biotechnol. Biofuels 8: 1–18. https://doi.org/10.1186/s13068-015-0403-z

Stenberg K., M. Galbe, and G. Zacchi, 2000 The influence of lactic acid formation on the simultaneous saccharification and fermentation (SSF) of softwood to ethanol. Enzyme Microb. Technol. 26: 71–79. https://doi.org/10.1016/S0141-0229(99)00127-1

Studer M. H., J. D. DeMartini, M. F. Davis, R. W. Sykes, B. Davison, *et al.*, 2011 Lignin content in natural *Populus* variants affects sugar release. Proc. Natl. Acad. Sci. 108: 6300–6305. https://doi.org/10.1073/pnas.1009252108

Sugiyama Y., M. Wakazaki, K. Toyooka, H. Fukuda, and Y. Oda, 2017 A novel plasma membrane-anchored protein regulates xylem cell-wall deposition through microtubule-dependent lateral inhibition of Rho GTPase Domains. Curr. Biol. 27: 2522-2528.e4. https://doi.org/10.1016/j.cub.2017.06.059

Takeuchi M., T. Kegasa, A. Watanabe, M. Tamura, and Y. Tsutsumi, 2018 Expression analysis of transporter genes for screening candidate monolignol transporters using *Arabidopsis thaliana* cell suspensions during tracheary element differentiation. J. Plant Res. 131: 297–305. https://doi.org/10.1007/s10265-017-0979-4

Tang H. M., S. Liu, S. Hill-Skinner, W. Wu, D. Reed, *et al.*, 2014 The maize brown *midrib2* (*bm2*) gene encodes a methylenetetrahydrofolate reductase that contributes to lignin accumulation. Plant J. 77: 380–392. https://doi.org/10.1111/tpj.12394

Taylor G., 2002 *Populus*: *Arabidopsis* for forestry. Do we need a model tree? Ann. Bot. 90: 681–689. https://doi.org/10.1093/aob/mcf255

Teleman A., J. Lundqvist, F. Tjerneld, H. Stålbrand, and O. Dahlman, 2002 Characterization of acetylated 4-O-methylglucuronoxylan isolated from aspen employing 1H and 13C NMR spectroscopy. Carbohydr. Res. 329: 807–815. https://doi.org/10.1016/s0008-6215(00)00249-4

Terashima N., M. Yoshida, J. Hafrén, K. Fukushima, and U. Westermark, 2012 Proposed supramolecular structure of lignin in softwood tracheid compound middle lamella regions. Holzforschung 66: 907–915. https://doi.org/10.1515/hf-2012-0021

Thomas L. H., V. T. Forsyth, A. Sturcova, C. J. Kennedy, R. P. May, *et al.*, 2013 Structure of cellulose microfibrils in primary cell walls from collenchyma. Plant Physiol. 161: 465–476. https://doi.org/10.1104/pp.112.206359

Timmers J., S. Vernhettes, T. Desprez, J. P. Vincken, R. G. F. Visser, *et al.*, 2009 Interactions between membrane-bound cellulose synthases involved in the synthesis of the secondary cell wall. FEBS Lett. 583: 978–982. https://doi.org/10.1016/j.febslet.2009.02.035

Tsai C. J., and L. J. Xue, 2015 CRISPRing into the woods. GM Crops Food 6: 206–215. https://doi.org/10.1080/21645698.2015.1091553

Tsuyama T., R. Kawai, N. Shitan, T. Matoh, J. Sugiyama, *et al.*, 2013 Proton-dependent coniferin transport, a common major transport event in differentiating xylem tissue of woody plants. Plant Physiol. 162: 918–926. https://doi.org/10.1104/pp.113.214957

Turner S. R., and C. R. Somerville, 1997 Collapsed xylem phenotype of arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. Plant Cell 9: 689–701. https://doi.org/10.1105/tpc.9.5.689

Turner S. R., N. G. Taylor, and L. Jones, 2001 Analysis of secondary cell wall formation in *Arabidopsis*. Prog. Biotechnol. 18: 85–92. https://doi.org/10.1016/S0921-0423(01)80059-9

Urbanowicz B. R., M. J. Pena, S. Ratnaparkhe, U. Avci, J. Backe, *et al.*, 2012 4-O-methylation of glucuronic acid in *Arabidopsis* glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. Proc. Natl. Acad. Sci. 109: 14253–14258. https://doi.org/10.1073/pnas.1208097109

Urbanowicz B. R., M. J. Peña, H. A. Moniz, K. W. Moremen, and W. S. York, 2014 Two *Arabidopsis* proteins synthesize acetylated xylan in vitro. Plant J. 80: 197–206. https://doi.org/10.1111/tpj.12643

Urnov F., 2020 Prime Time for Genome Editing? N. Engl. J. Med. 382: 481–484. https://doi.org/10.1056/nejmcibr1914271

Vanholme R., V. Storme, B. Vanholme, L. Sundin, J. H. Christensen, *et al.*, 2012 A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. Plant Cell 24: 3506–3529. https://doi.org/10.1105/tpc.112.102574

Vanholme R., I. Cesarino, K. Rataj, Y. Xiao, L. Sundin, *et al.*, 2013 CAFFEOYL SHIKIMATE ESTERASE (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. Science. 341: 1103–1106. https://doi.org/10.1126/science.1241602

Vardon D. R., N. A. Rorrer, D. Salvachúa, A. E. Settle, C. W. Johnson, *et al.*, 2016 Cis, cis-muconic acid: separation and catalysis to bio-adipic acid for nylon-6,6 polymerization. Green Chem. 18: 3397–3413. https://doi.org/10.1039/c5gc02844b

Vermaas J. V., R. A. Dixon, F. Chen, S. D. Mansfield, W. Boerjan, *et al.*, 2019 Passive membrane transport of lignin-related compounds. Proc. Natl. Acad. Sci. 116: 23117–23123. https://doi.org/10.1073/pnas.1904643116

Vogel J., 2008 Unique aspects of the grass cell wall. Curr. Opin. Plant Biol. 11: 301–307. https://doi.org/10.1016/j.pbi.2008.03.002

Wang Q., D. Ci, T. Li, P. Li, Y. Song, *et al.*, 2016 The role of DNA methylation in xylogenesis in different tissues of poplar. Front. Plant Sci. 7: 1–12. https://doi.org/10.3389/fpls.2016.01003

Wang M., Y. Mao, Y. Lu, X. Tao, and J. kang Zhu, 2017 Multiplex gene editing in rice using the CRISPR-Cpf1 system. Mol. Plant 10: 1011–1013. https://doi.org/10.1016/j.molp.2017.03.001

Wang J. P., M. L. Matthews, C. M. Williams, R. Shi, C. Yang, *et al.*, 2018 Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. Nat. Commun. 9: 1579. https://doi.org/10.1038/s41467-018-03863-z

Watanabe Y., M. J. Meents, L. M. McDonnell, S. Barkwill, A. Sampathkumar, *et al.*, 2015 Visualization of cellulose synthases in *Arabidopsis* secondary cell walls. Science. 350: 198–203. https://doi.org/10.1126/science.aac7427

Wierzbicki M. P., N. Christie, D. Pinard, S. D. Mansfield, E. Mizrachi, *et al.*, 2019a A systems genetics analysis in *Eucalyptus* reveals coordination of metabolic pathways associated with xylan modification in wood-forming tissues. New Phytol. 223: 1952–1972. https://doi.org/10.1111/nph.15972

Wierzbicki M. P., V. Maloney, E. Mizrachi, and A. A. Myburg, 2019b Xylan in the Middle: Understanding xylan biosynthesis and its metabolic dependencies toward improving wood fiber for industrial processing. Front. Plant Sci. 10: 1–29. https://doi.org/10.3389/fpls.2019.00176

Wightman R., and S. R. Turner, 2008 The roles of the cytoskeleton during cellulose deposition at the secondary cell wall. Plant J. 54: 794–805. https://doi.org/10.1111/j.1365-313X.2008.03444.x

Wingren A., M. Galbe, and G. Zacchi, 2003 Techno-economic evaluation of producing ethanol from softwood: Comparison of SSF and SHF and identification of bottlenecks. Biotechnol. Prog. 19: 1109–1117. https://doi.org/10.1021/bp0340180

Wu Z., and Y. Y. Lee, 1998 Nonisothermal simultaneous saccharification and fermentation for direct conversion of lignocellulosic biomass to ethanol. Appl. Biochem. Biotechnol. 70–72: 479–492. https://doi.org/10.1007/BF02920161

Wu A. M., C. Rihouey, M. Seveno, E. Hörnblad, S. K. Singh, *et al.*, 2009 The *Arabidopsis* IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. Plant J. 57: 718–731. https://doi.org/10.1111/j.1365-313X.2008.03724.x

Wu A. M., E. Hörnblad, A. Voxeur, L. Gerber, C. Rihouey, *et al.*, 2010 Analysis of the arabidopsis IRX9/IRX9-L and IRX14/IRX14-L pairs of glycosyltransferase genes reveals critical contributions to biosynthesis of the hemicellulose glucuronoxylan. Plant Physiol. 153: 542–554. https://doi.org/10.1104/pp.110.154971

Wyman C. E., D. D. Spindler, and K. Grohmann, 1992 Simultaneous saccharification and fermentation of several lignocellulosic feedstocks to fuel ethanol. Biomass and Bioenergy 3: 301–307. https://doi.org/10.1016/0961-9534(92)90001-7

Xiong G., K. Cheng, and M. Pauly, 2013 Xylan O-Acetylation Impacts xylem development and enzymatic recalcitrance as indicated by the *Arabidopsis* Mutant *tbl29*. Mol. Plant 6: 1373–1375. https://doi.org/10.1093/mp/sst014

Xiong G., M. Dama, and M. Pauly, 2015 Glucuronic acid moieties on xylan are functionally equivalent to O-Acetyl-substituents. Mol. Plant 8: 1119–1121. https://doi.org/10.1016/j.molp.2015.02.013

Xu R., R. Qin, H. Li, D. Li, L. Li, *et al.*, 2017 Generation of targeted mutant rice using a CRISPR-Cpf1 system. Plant Biotechnol. J. 15: 713–717. https://doi.org/10.1111/pbi.12669

Xu R., J. Li, X. Liu, T. Shan, R. Qin, *et al.*, 2020 Development of plant prime-editing systems for precise genome editing. Plant Commun. 100043. https://doi.org/10.1016/j.xplc.2020.100043

Xu W., H. Cheng, S. Zhu, J. Cheng, H. Ji, *et al.*, 2021 Functional understanding of secondary cell wall cellulose synthases in *Populus trichocarpa* via the Cas9/gRNA-induced gene knockouts. New Phytol. n/a. https://doi.org/https://doi.org/10.1111/nph.17338

Yan F., Y. Kuang, B. Ren, J. Wang, D. Zhang, *et al.*, 2018 Highly Efficient A·T to G·C base editing by Cas9n-guided tRNA adenosine deaminase in rice. Mol. Plant 11: 631–634. https://doi.org/10.1016/j.molp.2018.02.008

Yang F., P. Mitra, L. Zhang, L. Prak, Y. Verhertbruggen, *et al.*, 2013 Engineering secondary cell wall deposition in plants. Plant Biotechnol. J. 11: 325–335. https://doi.org/10.1111/pbi.12016

Yang Y., C. G. Yoo, K. A. Winkeler, C. M. Collins, M. A. W. Hinchee, *et al.*, 2017 Overexpression of a Domain of Unknown Function 231-containing protein increases O-xylan acetylation and cellulose biosynthesis in *Populus*. Biotechnol. Biofuels 10: 1–16. https://doi.org/10.1186/s13068-017-0998-3

Yi Chou E., M. Schuetz, N. Hoffmann, Y. Watanabe, R. Sibout, *et al.*, 2018 Distribution, mobility, and anchoring of lignin-related oxidative enzymes in *Arabidopsis* secondary cell walls. J. Exp. Bot. 69: 1849–1859. https://doi.org/10.1093/jxb/ery067

Yin Y., J. Huang, X. Gu, M. Bar-Peled, and Y. Xu, 2011 Evolution of plant nucleotide-sugar interconversion enzymes. PLoS One 6. https://doi.org/10.1371/journal.pone.0027995

Yin X., A. K. Biswal, J. Dionora, K. M. Perdigon, C. P. Balahadia, *et al.*, 2017 CRISPR-Cas9 and CRISPR-Cpf1 mediated targeting of a stomatal developmental gene EPFL9 in rice. Plant Cell Rep. 36: 745–757. https://doi.org/10.1007/s00299-017-2118-z

York W. S., and M. A. O'Neill, 2008 Biochemical control of xylan biosynthesis - which end is up? Curr. Opin. Plant Biol. 11: 258–265. https://doi.org/10.1016/j.pbi.2008.02.007

Yuan Y., Q. Teng, R. Zhong, and Z. H. Ye, 2013 The *Arabidopsis* DUF231 domain-containing protein ESK1 mediates 2-O- and 3-O-acetylation of xylosyl residues in xylan. Plant Cell Physiol. 54: 1186–1199. https://doi.org/10.1093/pcp/pct070

Yuan Y., Q. Teng, C. Lee, R. Zhong, and Z. H. Ye, 2014 Modification of the degree of 4-O-methylation of secondary wall glucuronoxylan. Plant Sci. 219–220: 42–50. https://doi.org/10.1016/j.plantsci.2014.01.005

Yung M. M., W. S. Jablonski, and K. A. Magrini-Bair, 2009 Review of catalytic conditioning of biomass-derived syngas. Energy and Fuels 23: 1874–1887. https://doi.org/10.1021/ef800830n

Zaidi S. S. e. A., M. M. Mahfouz, and S. Mansoor, 2017 CRISPR-Cpf1: A New Tool for Plant Genome Editing. Trends Plant Sci. 22: 550–553. https://doi.org/10.1016/j.tplants.2017.05.001

Zanuttini M., V. Marzocchi, M. Citroni, and R. Mocchiutti, 2003 Alkali impregnation of hardwoods. Part 1: Moderate treatment of poplar wood. J. Pulp Pap. Sci. 29: 313–317.

Zanuttini M., V. Marzocchi, P. Mocchiutti, and M. Inalbon, 2005 Deacetylation consequences in pulping processes. Holz als Roh- und Werkst. 63: 149–153. https://doi.org/10.1007/s00107-004-0557-z

Zeng Y., M. E. Himmel, and S. Y. Ding, 2017 Visualizing chemical functionality in plant cell walls. Biotechnol. Biofuels 10: 1–16. https://doi.org/10.1186/s13068-017-0953-3

Zeng W., E. R. Lampugnani, K. L. Picard, L. Song, A. M. Wu, *et al.*, 2016 Asparagus IRX9, IRX10, and IRX14A are components of an active xylan backbone synthase complex that forms in the Golgi apparatus. Plant Physiol. 171: 93–109. https://doi.org/10.1104/pp.15.01919

Zetsche B., J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, *et al.*, 2015 Cpf1 is a single RNA-Guided endonuclease of a class 2 CRISPR-Cas System. Cell 163: 759–771. https://doi.org/10.1016/j.cell.2015.09.038

Zetsche B., M. Heidenreich, P. Mohanraju, I. Fedorova, J. Kneppers, *et al.*, 2017 Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat. Biotechnol. 35: 31–34. https://doi.org/10.1038/nbt.3737

Zhao Q., Y. Tobimatsu, R. Zhou, S. Pattathil, L. Gallego-Giraldo, *et al.*, 2013 Loss of function of *CINNAMYL ALCOHOL DEHYDROGENASE 1* leads to unconventional lignin and a temperaturesensitive growth defect in Medicago truncatula. Proc. Natl. Acad. Sci. 110: 13660–13665. https://doi.org/10.1073/pnas.1312234110

Zhao X., N. Liu, N. Shang, W. Zeng, B. Ebert, *et al.*, 2018 Three UDP-xylose transporters participate in xylan biosynthesis by conveying cytosolic UDP-xylose into the Golgi lumen in *Arabidopsis*. J. Exp. Bot. 69: 1125–1134. https://doi.org/10.1093/jxb/erx448

Zhong R., M. J. Peña, G. K. Zhou, C. J. Nairn, A. Wood-Jones, *et al.*, 2005 *Arabidopsis FRAGILE FIBER8*, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. Plant Cell 17: 3390–3406. https://doi.org/10.1105/tpc.105.035501

Zhong R., and Z. H. Ye, 2015 Secondary cell walls: Biosynthesis, patterned deposition and transcriptional regulation. Plant Cell Physiol. 56: 195–214. https://doi.org/10.1093/pcp/pcu140

Zhong R., Q. Teng, M. Haghighat, Y. Yuan, S. T. Furey, *et al.*, 2017 Cytosol-localized UDP-xylose synthases provide the major source of UDP-xylose for the biosynthesis of xylan and xyloglucan. Plant Cell Physiol. 58: 156–174. https://doi.org/10.1093/pcp/pcw179

Zhong R., D. Cui, and Z.-H. Ye, 2018 Xyloglucan O-acetyltransferases from *Arabidopsis thaliana* and *Populus trichocarpa* catalyze acetylation of fucosylated galactose residues on xyloglucan side chains. Planta 248: 1159–1171. https://doi.org/10.1007/s00425-018-2972-0

Zhong R., D. Cui, and Z. H. Ye, 2019 Secondary cell wall biosynthesis. New Phytol. 221: 1703–1723. https://doi.org/10.1111/nph.15537

Zhou X., T. B. Jacobs, L. J. Xue, S. A. Harding, and C. J. Tsai, 2015 Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and redundancy. New Phytol. 208: 298–301. https://doi.org/10.1111/nph.13470

1.9 Figures

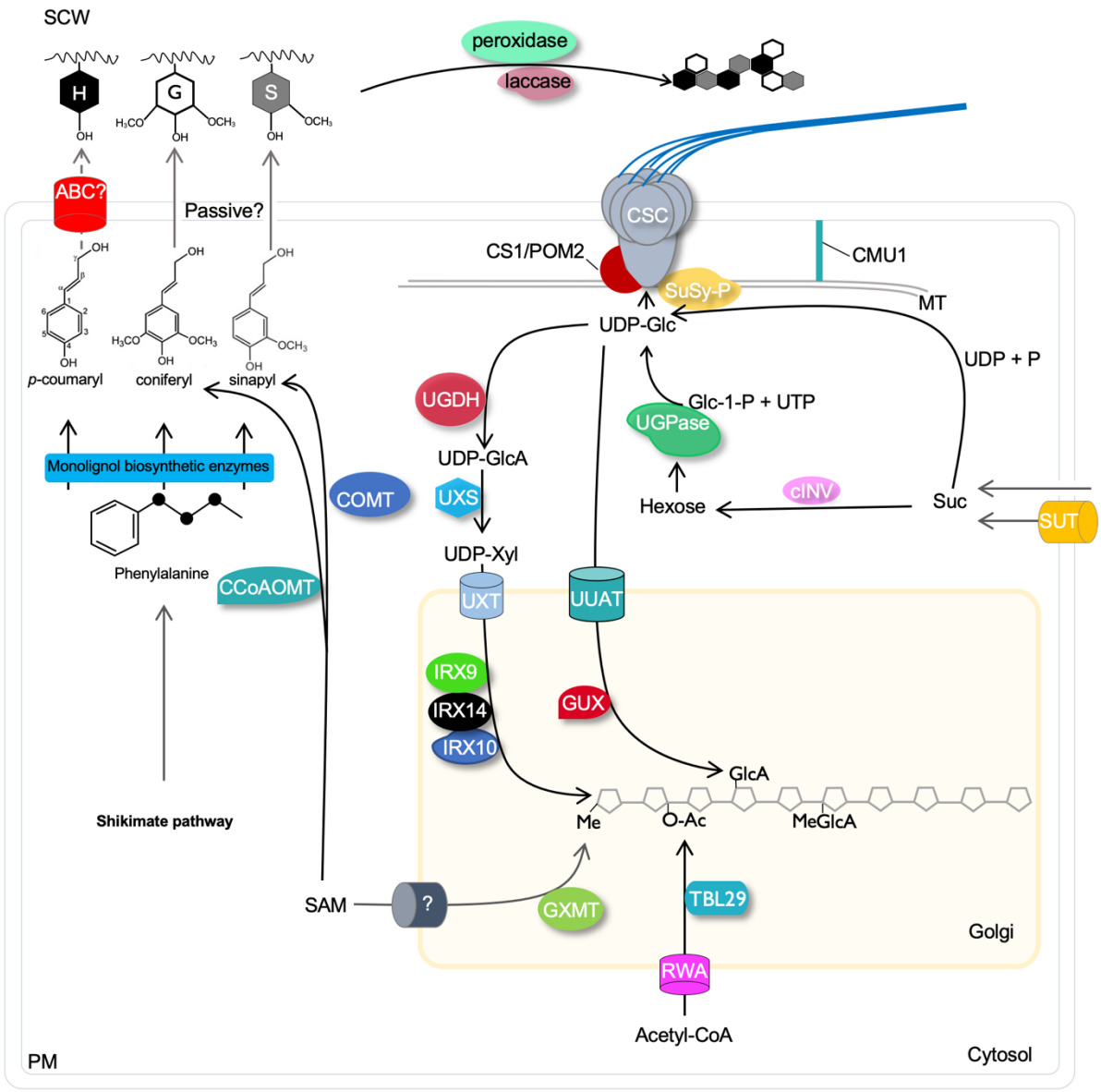


Figure 1.1 Biosynthesis of the main secondary cell wall biopolymers.

Sucrose is taken into the cells passively or through sucrose-H+ symporter (SUT). In the cytosol, sucrose is converted into UDP-Glc either through phosphorylated SUSY or cytosolic invertase (cINV) and UDP-Glc pyrophosphorylase (UGPase), membrane bound CSC uses UDP-Glc to synthase cellulose microfibrils which aggregate together at the SCW. UDP-Glc is used by UXS to generate UDP-Xyl which is transported into the Golgi through UXT; within the Golgi, xylan backbone is synthesized from UDP-Xyl by IRXs. Xylan backbone is substituted with GlcA which may be methylated by GXMTs using SAM precursor. Furthermore, acetyl-CoA is transported into the Golgi by RWA for acetylation of xylan backbone. Finally, monolignols are synthesized from phenylalanine by a suite of biosynthetic enzymes. After synthesis in the cytosol, monolignols are transported into the SCW via either transporters or passive transport mechanisms. At the cell wall, monolignols yield monomers which are polymerized into final lignin structure by peroxidase and laccase.

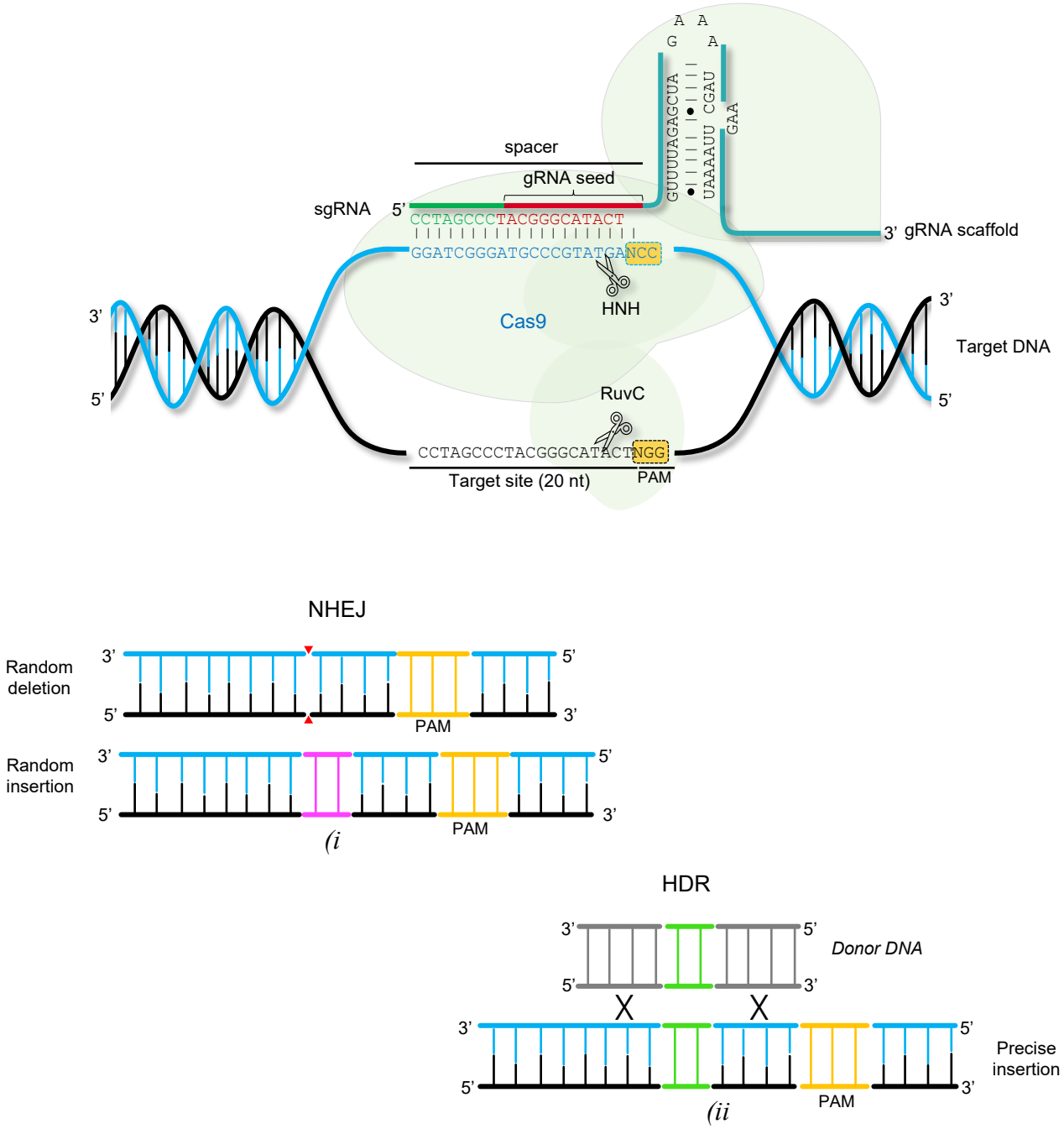


Figure 1.2 Overview of gene editing through the CRISPR/Cas9 system.

Cas9 is recruited to bind to the target site through the user-defined single guide RNA (sgRNA). sgRNA consists of the 20-bp protospacer which is complementary to the target site and the gRNA scaffold necessary for Cas9 binding. Once recruited to the target site, the protospacer adjacent motif (PAM) allows Cas9 protein to interact with DNA. Both DNA strands are then cleaved by Cas9's HNH and RuvC domains to generate double stranded breaks (DSBs). DSBs can be repaired through the (i) error-prone nonhomologous end joining (NHEJ) which leads to random insertions (pink) and deletions (red triangles), causing gene knock-outs by disruption. Alternatively, (ii) in the presence of an exogenous DNA template the DSBs can be repaired via the homology-directed repair (HDR) which can yield precise gene insertions (green).

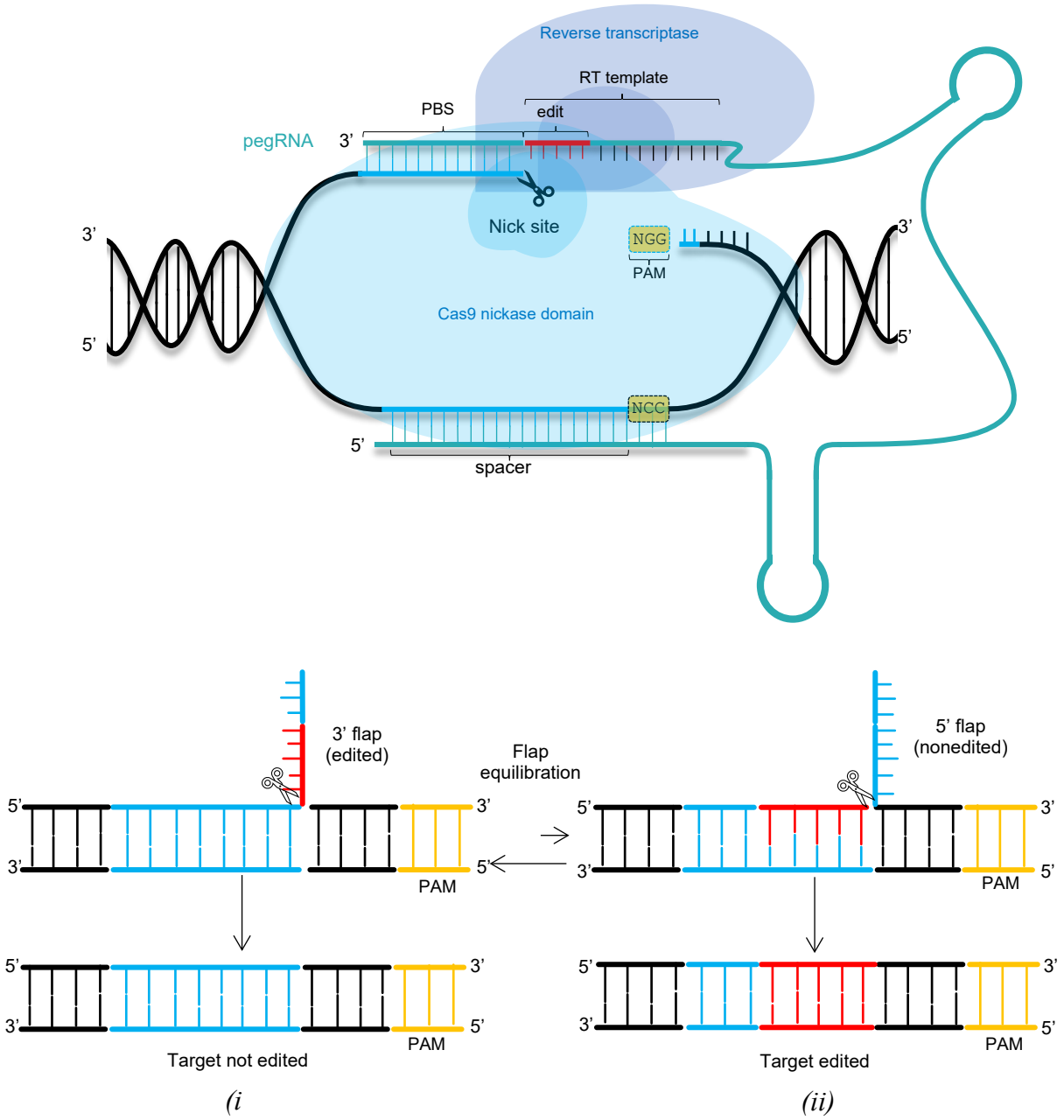


Figure 1.3 Overview of the DNA prime editing machinery.

The prime editing (PE) enzyme comprises the Cas9 nickase domain fused with reverse transcriptase domain complexed with the prime editing guide RNA (pegRNA). As CRISPR/Cas9 mechanism, the spacer of the pegRNA recognizes the target site. The cas9 nickase domain (H840A) within the PE then interacts with pegRNA and nicks the PAM strand. The subsequent 3' flap hybridizes with the primer binding site (PBS), the PBS allows reverse transcriptase to reverse transcribe the new desired sequence (red) using the RT template from pegRNA. After this, two possibilities exist: (i) in the absence of flap equilibration, 5' flap is excised, and the target sequence is not modified or (ii) flap equilibration allows the 5' flap to hybridize to the unmodified DNA strand and mutation may be introduced by DNA repair.

Chapter 2

Candidate gene selection and functional analyses of *Arabidopsis* T-DNA insertion mutants

Nganeavhutshilo N̂angammbi, Alexander A. Myburg[1], Eshchar Mizrachi[1] and Victoria J. Maloney[1]

[1]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), Forest Molecular Genetics (FMG), University of Pretoria, Private bag X20, Pretoria, 0028

This chapter has been prepared in the format of a manuscript for an ISI rated scientific journal (e.g., New Phytologist). I performed all the experimental work and data analyses and drafted the manuscript. Prof Myburg, Prof Mizrachi, and Dr Maloney conceived and supervised the study and helped edit the final manuscript.

2.1 Summary

Secondary cell walls provide mechanical support to plant cells allowing them to withstand the vast majority of biotic and abiotic stresses, this is essential for normal plant growth and development. However, SCWs also contribute significantly to biomass recalcitrance, making it a challenge to conveniently exploit woody biomass for their valuable applications. It is therefore necessary to engineer plants with cell walls that are easier to deconstruct; this requires identification of essential SCW biosynthetic genes that can be manipulated to potentially reduce biomass recalcitrance. In this study, systems genetics was used for identify xylan-associated genes as follows. Known xylan modification genes were used as bait to pull out genes that had significant co-expression with them in a *Eucalyptus* backcross population. The resulting number of genes were then used in a Pearson correlation with several biotechnologically-relevant traits such as xylan acetylation and pulp yield. Following a correlation analyses, few genes were either positively or negatively correlated with one or more traits. Genes that had a negative correlation with pulp yield and a positive correlation with xylan acetylation were then selected for this study, as these would be the best candidates for CRISPR editing/knockdown because knocking them out may reduce xylan acetylation and increase pulp yield.

Using this systems genetics data and support from literature, five *Arabidopsis* gene homologs; *UDP-XYL SYNTHASE 6* (*AtUXS6*), *XYLANASE 1* (*AtXYN1*), *ENDOMEMBRANE PROTEIN 70* (*AtEMP70*), *ENDO-BETA-MANNASE 2* (*AtMAN2*), and *ALPHA-L-ARABINOFURANOSIDASE 1* (*AtARAF1*), were selected for functional testing in *Arabidopsis thaliana*. *xyn1* T-DNA insertional mutants had significantly shorter roots and smaller rosette leaves, however, the delayed growth phenotype was restricted to early growth as mutants had significantly longer inflorescence stems when measured at the end of stem growth. In contrast, notably shorter inflorescence stems were observed in *araf1* mutants during early bolting. Finally, *emp70* mutants showed an interesting and consistent growth advantage;

mutants germinated significantly earlier, had larger and more rosettes leaves, and significantly taller stems compared to wild-type. Subsequent analyses of cell wall chemistry and morphology showed that *xyn1* mutants have significantly more lignin. Thus, while analyses of saccharification efficiency in these mutants are still pending, the preliminary mutant phenotypic data generated in this study provides a starting ground for eventual functional characterization of these candidate genes and understanding of their precise biological roles in plant growth and development.

2.2 Introduction

Plant secondary cell walls are the main constituent of wood and a major source of lignocellulosic biomass which is mainly composed of cellulose, hemicellulose, and lignin biopolymers (McDougall *et al.* 1993). The ability of trees to biosynthesize these carbohydrate-based secondary cell walls is of significant interest. Amongst many interesting characteristics, the SCWs are resistant to biodegradation for decades without any living cells aiding in their resistance. Moreover, the SCWs provide a lifelong mechanical stability to trees, allowing them to grow tall and transport water for very long distances over enormous pressure. Therefore, understanding secondary cell wall development and chemistry is of utmost importance both in its own right, and for the use of wood as biofuel feedstock.

As a biorefinery feedstock, lignocellulosic biomass is the most abundant source of renewable material; accordingly, it has been extensively exploited to produce several biobased products such as timber, textiles, and paper. Moreover, significant efforts have been made to utilize lignocellulosic biomass as a potential feedstock for the production of second-generation biofuels (Pauly and Keegstra 2008; Petersen *et al.* 2012; Gandla *et al.* 2018). Utilization of lignocellulosic biomass for its wide range of applications thus holds a promise for a sustainable bioeconomy. However, the overall structure of the secondary cell wall causes biomass recalcitrance during industrial bioprocessing. Amongst others, the

overall biomass structure, the lignin content, and xylan acetylation, are considered the major contributors to biomass recalcitrance (Himmel *et al.* 2007). Glucuronoxylan backbone is mainly composed of β-1-4-linked D-xylosyl residues substituted with D-Glucuronic acid (GlcA) and can be decorated with methyl and 2-O and/or 3-O acetyl moieties depending on species. The presence of *O*-Acetyl groups on xylan backbone create steric hindrance for the binding and subsequent activity of hydrolytic enzymes, preventing the removal of xylan and efficient extraction of high purity cellulose in dissolving pulp processing (Helle *et al.* 2003; Qing and Wyman 2011; Chong *et al.* 2015). Furthermore, *O*-acetyl groups may be released into the medium during bioprocessing, this leads to significant changes in pH, which could potentially inhibit the activity of fermentative microbes necessary for successful lignocellulose bioconversion (Basen *et al.* 2014). Besides xylan's significant contribution to biomass recalcitrance, lignin also has a notable contribution to biomass recalcitrance (Himmel *et al.* 2007). Lignin is highly resistant to biodegradation owing to its complex structure composed of cross-linked polyphenolic monomers (Himmel *et al.* 2007; Zeng *et al.* 2017). Furthermore, lignin tightly interacts and binds xylan while xylan further interconnects lignin to cellulose (Kang *et al.* 2019), this further impedes bioconversion. Altogether, xylan (and its modifications) and lignin structure largely contribute to the resistance of biomass to bioprocessing. A significant improvement of bioprocessing through reduction of biomass recalcitrance thus requires an in-depth understanding of the biosynthesis of these biopolymers and the overall secondary cell wall.

As previously mentioned, the most abundant hemicellulose in eudicots is xylan, and its significant contribution to biomass recalcitrance has stimulated significant interest in uncovering its overall biosynthetic pathway and an understanding of how its structure can be modified to improve industrial bioprocessing. The biosynthesis of the xylan backbone involves the activity of XYLAN SYNTHASE COMPLEX (XSC) within the Golgi (reviewed in Smith *et al.* 2017). In *Arabidopsis*, the XSC is consists of IRREGULAR XYLEM 9 (IRX9), IRX10, and IRX14 GT43 proteins, IRX9 and IRX14 are

functionally non-redundant structural proteins most likely involved in backbone elongation, while IRX10 possesses the xylosyltransferase activity (Peña *et al.* 2007; Brown *et al.* 2009; Lee *et al.* 2010, 2012d; Chen *et al.* 2013b). However, the activity of IRX9 and IRX10 has not been biochemically confirmed thus far. Nevertheless, IRX9 and IRX10 orthologs have been identified in poplar, rice, and wheat, suggesting that these enzymes may indeed be playing an indispensable role in xylan synthesis not only in eudicots but vascular plants in general (Chiniquy *et al.* 2013; Lee *et al.* 2014; Pellny *et al.* 2020; Petrik *et al.* 2020). During biosynthesis, the xylan backbone is substituted with D-glucuronic acid (GlcA) residues to generate glucuronoxylan, this involves GLUCURONIC ACID SUBSTITUTION OF XYLAN 1 (GUX1), GUX2, and GUX3 Golgi-localized glycosyltransferases (Ebringerová and Heinze 2000; Mortimer *et al.* 2010; Lee *et al.* 2012a). These substitutions are necessary for normal xylan function during secondary cell wall deposition, as simultaneous mutation of GUX1, GUX2, and GUX3 (*gux1/2/*3) in *Arabidopsis* lead to loss of GlcA decorations accompanied by reduced secondary cell wall thickness and collapsed xylem vessels (Lee *et al.* 2012a). The reduced cell wall thickening observed in these mutants is possibly due to the altered interaction between the unsubstituted xylan and cellulose, which affects the overall cell wall assembly, and consequently leads to defects in secondary cell wall thickening (Lee *et al.* 2012a).

It has been demonstrated that GUX1 and GUX2 generate different and specific GlcA substitution patterns and subsequent xylan domains, the latter of which may affect xylan's interaction with other biopolymers, thereby impacting bioconversion (Bromley *et al.* 2013). In *Arabidopsis*, about 80% of the GlcA side chains are O-4 methylated by GLUCURONIC ACID SUBSTITUTION OF XYLAN 1 (GXM1), GXM2, and GXM3 glucuronoxylan methyltransferases using *S*-Adenosyl methionine (SAM-e) as a methyl donor (Kopycki *et al.* 2008; Lee *et al.* 2012b; Urbanowicz *et al.* 2012b; Smith *et al.* 2020). Furthermore, as previously described, a large proportion of xylosyl residues are *O*-acetylated, this is achieved through the REDUCE WALL ACETYLATION (RWA) and TRICHOME-

BIREFRINGENCE-like (TBL) protein families. It has been proposed that RWAs may be involved in acetyl-CoA transport into the Golgi (Gu *et al.* 2010; Lee *et al.* 2011; Rennie and Scheller 2014); within the Golgi, TBL29 then transfers acetyl to *O*-2 and *O*-3 of the xylan backbone (Urbanowicz *et al.* 2014; Lunin *et al.* 2020). Finally, after biosynthesis in the Golgi, xylan is transported to the secondary cell wall through vesicles where it forms strong permanent interactions with cellulose microfibrils and lignin (Kang *et al.* 2011).

Once at the secondary cell wall, xylan is bound by lignin most likely through electrostatic forces, lignin polymer further interacts with cellulose through xylan "bridges" (Kang *et al.* 2019), thereby generating a strong and hydrophobic plant cell wall. Lignin polymer is mostly derived from the monolignols p-coumaryl, coniferyl, and sinapyl alcohols (Ralph *et al.* 2008). At the cell wall, these monomers give rise to hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units, respectively (Bonawitz and Chapple 2010; Vanholme *et al.* 2012b; Hao and Mohnen 2014), which get polymerized into final lignin structure through peroxidase and laccases (Boudet 2000). The proportion of the S and G monomers (referred to as the S/G ratio) determines the overall structure of lignin and correlates with how easy it is to delignify the secondary cell wall (González-Vila *et al.* 1999; Del Río *et al.* 2005; Chen and Dixon 2007; Nunes *et al.* 2010; Studer *et al.* 2011). Thus, the biosynthesis of secondary cell wall biopolymers is complex, requiring a repertoire of biosynthetic genes and a substantial supply of metabolic precursors for both backbone biosynthesis and modifications.

An important aspect in research aimed at improving bioprocessing efficiency through engineering plants with beneficial traits is an understanding of how the expression of these secondary cell wall biosynthetic genes correlates with bioprocessing traits. This subsequently allows the identification of a suite of genes that, when genetically altered, are more likely to present traits beneficial to industrial bioprocessing. To this end, a previous systems genetics study used an integrated population-wide

analysis of developing xylem gene co-expression and the expression quantitative trait loci (eQTLs) to identify potential xylan-modification genes in *Eucalyptus* (Wierzbicki *et al.* 2019a). This involved identifying few genes from the TBL, RWA, GUX and DUF579 families from the PLAZA 3.0 *E. grandis* genome. Twenty-nine genes from this set that were putative homologs of SCW xylan modification genes in *Eucalyptus* were determined through a combination of phylogenetic reconstruction, network and co-expression reconstruction (Pinard *et al.* 2019). These genes were then subjected to network and co-expression analyses to identify co-expressed genes in the developing xylem transcriptomes of *Eucalyptus* backcross population. From this co-expression analyses, 1136 *E. grandis* genes were co-expressed with the 24 xylan modification genes and these were clustered into five nonredundant expression modules (EMs). Amongst the EMs, few genes were associated with SCW biosynthesis, and the expression of these genes was correlated with few biotechnology traits, generating an expression pattern useful in determining the genes that could be manipulated to yield favorable bioprocessing traits (Mizrachi *et al.* 2017; Wierzbicki *et al.* 2019a; b).

Here, we aimed to investigate how manipulating cell wall biosynthetic genes can impact cell wall development, plant growth, and physiology. We used system genetics data described above to select genes for functional testing of their T-DNA insertion lines in *Arabidopsis thaliana*, namely, *UDP-XYL SYNTHASE 6* (*AtUXS6*), *XYLANASE 1* (*AtXYN1*), *ENDOMEMBRANE PROTEIN 70* (*AtEMP70*), *ENDO-BETA-MANNASE 2* (*AtMAN2*), and *ALPHA-L-ARABINOFURANOSIDASE 1* (*AtARAF1*). These were selected based on the observation that their gene expression was positively correlated with xylan acetylation and negatively correlated with pulp yield, and thus, knocking them out will potentially lead to reduced xylan acetylation, and subsequent improvement in saccharification efficiency (i.e. pulp yield). *UXS6* is predominately expressed in the xylem vessels together with *UXS5* and *UXS3,* and encodes the enzyme UDP-Xyl synthase which catalyses the irreversible conversion of UDP-GlcA to UDP-Xylose necessary for xylan backbone synthesis (Kuang *et al.* 2016; Zhong *et al.*

2017). Simultaneous mutation of the three *UXS* genes lead to drastic reduction in SCW thickening accompanied with reduced xylan content (Kuang *et al.* 2016; Zhong *et al.* 2017). Similar to *UXS6*, *XYN1* is predominantly expressed in xylem tissues and encodes xylanase enzyme which is responsible for hydrolysis of xylan to generate xylose (Suzuki *et al.* 2002). *Arabidopsis XYN1* overexpressing lines show enhanced transport from hypocotyl to leaves, however, no difference in the growth, xylem anatomy and vessel element growth are observed in these lines (Endo *et al.* 2019). Similar to *XYN1*, *AtARAF1* encodes glycoside hydrolase/glycosyl hydrolase enzymes and it is ubiquitously expressed throughout the plant (Fulton and Cobbett 2003; Chavez Montes *et al.* 2008). More importantly, *AtARAF1* is amongst the most highly expressed homologs in developing tension wood in poplar (Andersson-Gunnerås et al. 2006). It has been suggested that this overexpression leads to a decrease in xylan levels while favouring cellulose biosynthesis (Andersson-Gunnerås et al. 2006), which confirms the hemicellulose hydrolase activity of the encoded enzyme. Besides xylan hydrolase enzymes, *MAN2* encodes a mannanase, which digests manno-polysaccharides to facilitate plant growth. This enzymatic activity is necessary during seed germination, during which *MAN* is suggested to facilitate the hydrolysis of mannan-rich endosperm cell walls to facilitate radicle emergence (Iglesias-Fernández *et al.* 2011b). Finally, *EMP70*'s encoded protein or function in plants is currently unknown, although heterologous expression of a different *Arabidopsis EMP70* (AT3G13772) homolog in yeast leads to a substantial increase in cellular copper accumulation (Hegelund *et al.* 2010).

We hypothesized that given information from literature, and the correlation of these genes with bioprocessing traits, they directly have an impact on cell wall physicochemical properties and that as a result perturbing their expression would result in cell wall-related phenotype. *ARAF1* T-DNA insertion lines have been previously analysed, however, here, we show a temporal growth phenotype that has not been observed before, we further perform secondary cell wall chemistry analyses. Moreover, previous analyses of *AtMAN2* showed that although mutation in other *MANs* display an

early germination in *Arabidopsis*, *MAN2* mutants do not exhibit such phenotype (Iglesias-Fernández *et al.* 2011a; b). Here, in addition to demonstrating that *AtMAN2* mutants do not display altered germination time as previously shown, we perform in-depth cell wall morphology and chemistry analyses on these mutants. We further perform, for the first time, growth analyses of *XYN1* and *EMP70* T-DNA insertion mutants. We show that *emp70* knock-out mutants have a growth advantage compared to wild-type as shown by their early germination, larger and more rosette leaves, and significantly longer stems compared to wild-type. Conversely, *xyn1* mutants have significantly shorter roots which lead to their reduced rosette leaf size in early development. Furthermore, *xyn1* mutants had increased lignin content compared to wild-type, however, a mutation in any of these genes did not affect monosaccharide content nor xylem vessel morphology. Thus, while this study only established preliminary growth phenotypes and cell wall composition in *Arabidopsis*, it lays a strong foundation for subsequent studies specifically aimed at functionally characterizing these genes and uncovering their exact impact on bioprocessing efficiency. More importantly, the consistent and newly observed phenotypes observed in *araf1*, *emp70,* and *xyn1* mutants merit further investigation in a model tree.

2.3 Materials and Methods

2.3.1 Candidate gene selection

To select candidate genes to be targeted in *Arabidopsis,* data generated through systems genetics from *Eucalyptus urophylla x E. grandis x E. urophylla* backcross population systems was used. To generate this data, *E grandis* genome data from PLAZA 3.0 was used to identify 72 genes from four gene families, namely; 55 TBL, 2 RWA, 11 DUF579 and 4 GUX genes (Myburg *et al.* 2014). From these genes, phylogenetic reconstruction was performed to determine which of these genes were homologous to the SCW modification genes in *Eucalyptus, Arabidopsis,* and *Populus* (Wierzbicki *et al.* 2019a). In addition to phylogenetic analyses, the initial 72 genes were subjected to network and co-expression reconstruction to aid in identifying co-expressed genes in the transcriptomes of developing xylem of 156 *Eucalyptus* backcross individuals (Pinard *et al.* 2019). From the 72 putative xylem modification genes, 24 query genes passed the set correlation cut-off of 0.6 (*F*-score of 3.8). These were subsequently used as bait to pull out 1136 *E. grandis* genes that had significant co-expression with them, and clustered into five distinct expression modules (EMs; **Figure 2.1**).

In the current study, the log-transformed expression profiles of all 1136 genes within the EMs were subjected to a Pearson correlation analysis with seven industrially relevant traits, namely, lignin, xylose, cellulose, and glucose content as well as xylan acetylation, S:G ratio, and pulp yield. These traits were previously determined by either wet lab or near infrared analysis (NIRA) prediction. The correlation results indicated that the expression of 539 genes had a significant (either negative or positive) correlation with one or more of the traits and the correlation patterns were distributed between 6 distinct clusters that were termed biotechnology clusters (BCs, **Figure 2.2**).

After the identification of all the genes that were both co-expressed with known xylan modification genes and significantly correlated with the seven industrially relevant traits, potential candidate genes for functional testing in *Arabidopsis* were selected from the 88 genes in BC 3 (**Figure 2.2**). Genes in this cluster were chosen as their expression was positively correlated to xylan acetylation and negatively correlated to pulp yield (**Figure 2.2**). To further narrow down the number of genes to be functionally characterized in this study, *Arabidopsis* gene locus IDs for all 88 BC 3 genes were queried in the *Arabidopsis* Information Resource database (https://www.Arabidopsis.org/) to identify genes, based on gene ontology (GO) annotations, that are associated with the biosynthesis and/or modification of secondary cell wall biopolymers. Additionally, the literature on each of the candidate genes was gathered to build evidence for the possible roles of the genes in secondary cell wall development.

2.3.2 Plant material and genetic analysis

The Columbia (Col-0) ecotype background was used for both wild-type and all T-DNA insertion mutants. T-DNA insertion lines (SALK_058602C, *UXS6*; SALK_099151C, *XYN1*; SALK_082204C, *EMP70*; SALK_126628C, *MAN2*; SALK_039176, *ARAF1*) were obtained from the *Arabidopsis* Biological Resource Centre, ABRC (https://abrc.osu.edu/). To select T-DNA insertion positive transformants, seeds were stratified at 4°C for 7 days and grown on Murashige and Skoog (MS) medium (0.5x Murashige and Skoog salts, 10 g/l sucrose and 7 g/l agar) supplemented with 50 µg/ml kanamycin (Petersen *et al.* 2012). MS Plates were kept in the dark at 4°C for 3 days and then transferred to the growth chamber at 20 °C and 16-hour light periods until germination. Seeds that germinated and rooted on kanamycin were transferred to jiffy pellets and grown at 22°C in a growth room with a 14-hour photoperiod.

To confirm that the T-DNA insertion lines were homozygous for the insertion, genomic DNA was extracted from 2-week-old leaves using NucleoSpin™ Plant Kit (Machery-Nagel) and quantified by A260 spectrophotometry. PCRs were then performed with the allele-specific left and right primer pair (LP + RP) and with RP and T-DNA left border (LB) primer pair (RP + LB**)**. All primers were generated with the T-DNA primer design tool from SIGnAL SALK (http://signal.salk.edu/tdnaprimers.2.html; **Table 2.2**). Both PCRs were performed with Taq polymerase (ExcelTaq™) in 20 µl final volume at 95°C for 3 minutes; 34 cycles at 95°C for 30 seconds, followed by primer annealing at 58°C for 30 seconds and 72°C for 90 seconds, and finally, extension at 72°C for 3 minutes. PCR products were analyzed by agarose gel electrophoresis in a 1% (w/v) agarose gel. Seeds were then collected from one of the confirmed homozygous plants (from each mutant line) for a growth trial.

2.3.3 Trial conditions and physical measurements

For each confirmed homozygous T-DNA insertion line, 36 biological replicates and 24 wild-types were planted. All growth trials were conducted in a controlled growth room at 22°C and with a 14-hour photoperiod. At two weeks post-planting, genotypes for all plants were reconfirmed through PCR as previously described, rosette leaf diameters were then measured every week until there were no significant changes in the rosette diameters. Following this, bolting time was determined by recording the age of the plant when the inflorescence stem emerged. At the same time, the number and size of rosettes leaves longer than 1 mm were recorded. Once all plants had bolted, the primary inflorescence stem height was measured every week until there was no obvious change in height. Observations from these physical measurements suggested that *xyn1* and *emp70* mutant lines have some notable growth differences compared to wild-type, to further investigate this observation, an additional growth trial on soil was performed as described above and germination time was investigated. Further, seeds were sterilized and sown on 0.5 MS (Sigma-Aldrich) nutrient media (without antibiotic or sucrose) as

previously described in Bahmani *et al.* 2016, germination time was evaluated and root length of the seedlings was measured at 1 day, 3 days, 7 days, and 10 days after germination in ½ MS.

2.3.4 Analyses of cell wall morphology with microscopy

Internode segments were cut 3 cm from the rosettes of 8-week-old plants and kept in FAA (50% ethanol, 10% formalin, 5% acetic acid) for 48 hours. Stems were then prepared for sectioning and subsequent microscopic analyses through London resin (LR) white embedding for transmission electron microscopy (TEM) protocol (Skepper and Powell 2008). Briefly, stems were washed in 20, 40, 60, and 80% LR white resin in ethanol by inverting the tube for 5-10 minutes between the different concentrations of resin and ethanol solutions. Samples were then kept in 100% LR white resin mixture overnight in the fridge. The following day, fresh LR white resin was added to the samples and incubated in the oven at 60°C for 36 hours. The LR white was then mixed with LR white accelerator and allowed to polymerize. The solidified resin was then removed, and samples embedded in gelatin samples and kept in the oven for 36 hours before sectioning. Thin stem transverse sections were cut with a vibratome and fixed into microscope slides. Sections were stained with 0.02% toluidine blue O for 30 minutes and subsequently rinsed with distilled water and kept at room temperature until dry. Vessel elements were visualized with Axio Imager.M2 (Zeiss) light microscope and micrographs were captured with AxioCam ICc 5 (Zeiss) and analyzed on Zen Pro (Zeiss) software.

2.3.5 Quantification of cell wall monosaccharide, lignin, and acetyl content

To investigate the impact of gene mutation on vessel morphology and secondary cell wall chemistry, nine-week-old *Arabidopsis* stems were harvested into 50 ml falcon tubes and left to dry for four weeks at room temperature. Six plants were pooled together to represent one sample, three independent samples, and two technical replicates were then generated per mutant line, while ten wild-type

technical replicates were generated. Monosaccharide and lignin content were determined as detailed by Maloney *et al.* 2012. Dry stems were ground through a Wiley mill to pass a 0.4-mm screen, following by removal of extractives in hot acetone overnight. The monosaccharide and lignin content were then determined through the modified Klason (Coleman *et al.* 2009). This involved treatment of ground tissue (50 mg) with 72% H_2SO_4 followed by stirring every 10 minutes for 2 hours. Samples were then diluted with deionized water and autoclaved at 121 °C for an hour. Lignin content was then determined by spectrophotometry at an absorbance of 205 nm. Monosaccharide content was determined with anion exchange high-performance liquid chromatography. Finally, cell wall acetyl content was quantified following the protocol described in Johnson *et al.* 2017. Briefly, a saponification reaction was performed, acetone-extracted wood samples were reacted with 0.2M sodium hydroxide. This reaction was then incubated for 75 minutes at 120 °C with shaking at 500 rpm. Following this, each sample was acidified with 72% w/w sulphuric acid, followed by cooling for 5 minutes in an ice bath. Next, samples were separated into solid and liquid by centrifuging at 13 000*g* for 2 minutes and the supernatant was injected onto the HPX-87H column (Aminex, USA) on a liquid chromatography instrument (Dionex, USA). The resulting acetic acid peaks were manually integrated into areas determined against the known concentration standards.

2.4 Results

2.4.1 Selection of xylan-associated genes in *Arabidopsis thaliana*

By combining phylogenetic, co-expression and network reconstruction, 24 xylan modification genes were identified from *E. grandis*, and used as bait in co-expression analyses to identify 1136 SCW-related genes in the *Eucalyptus urophylla* x *E. grandis* x *E. urophylla* backcross individuals (Myburg *et al.* 2014; Mizrachi *et al.* 2017; Pinard *et al.* 2019; Wierzbicki *et al.* 2019a). Of these, 539 were either

positively or negatively correlated with one or more bioprocessing traits based on Pearson correlation analyses as visualized in the six distinct biotechnology clusters (**Figure 2.2**). Following database search and GO analyses for BC 3 genes, 27 candidate genes were selected, available scientific literature on each of these genes were reviewed to gather information supporting their potential implication in secondary cell wall biosynthesis and modification. Fourteen of 27 genes were overrepresented in secondary cell wall biosynthetic processes based exclusively on literature (**Table S2.1**). However, for the feasibility of this study, five genes, namely, *UDP-XYL SYNTHASE 6* (*AtUXS6,* AT2G28760), *XYLANASE 1* (*AtXYN1,* AT1G58370), *ENDOMEMBRANE PROTEIN 70* (AT2G01970, *AtEMP70*), *ENDO-BETA-MANNASE 2* (*AtMAN2*, AT2G20680), and *ALPHA-L-ARABINOFURANOSIDASE 1* (*AtARAF1,* AT3G10740) were selected for functional testing in *Arabidopsis* (**Table 2.1**). To assess phenotypes resulting from knockout mutations in these genes, growth trials were initially performed in *Arabidopsis*. Notably, the growth trial of *AtMAN2* mutant line was performed separately to those for other mutants, therefore, the results from its trial are presented separately from that of the other lines where applicable.

2.4.2 Growth phenotypes of mutants

After confirming that all T-DNA insertional lines were homozygous knock-outs (**Figure S2.1)**, some of the growth characteristics that could reveal the impact of the mutation on plant growth and development were investigated. Starting at 2 weeks post-planting, rosette diameters were measured weekly until there was no observable difference in the rosette leaf growth between consecutive weeks of measurements. After the growth trial, the rosette diameter data was analyzed with the two-sample equal variance t-test to establish the significance of rosette sizes between wild-type and mutants. The rosettes of *xyn1* mutants were 4.76% smaller than wild-type's, while rosettes of *uxs6* and *emp70* mutants were 6.18% and 5.45% larger compared to wild-type, respectively (**Figure 2.3**). There *araf1*

mutants were also significantly larger than wild-type, although this difference was minor (**Figure 2.3**), furthermore, no difference was observed in the rosette sizes of *man2* mutants compared to wild-type (**Figure 2.3**). Following this, plant bolting time and the number of leaves at bolting were measured immediately after the plants had bolted; all mutants bolted at the same time as wild-type (two-tailed paired t-test; **Figure S2.2**). Furthermore, no statistically significant difference in the number of rosette leaves was observed in any mutant line except *emp70*, which had an average of 1 more rosette leaf at bolting compared to wild-type based on two-sample equal variance t-test (**Figure 2.4**).

Shortly after bolting, the maximum inflorescence stem height was measured and *araf1* mutants had notably shorter stems; 100 mm shorter on average compared to wild-type (**Figure S2.3**), however, these mutants had caught up and grow as well as wild-type by the 7th week post-planting (two-sample equal variance t-test; **Figure 2.5**). Notably, 7-week-old *xyn1*, *emp70*, and *man2* mutants had longer inflorescence stems (4.44%, 6.54%, and 13.58% longer than wild-type, respectively) although this phenotype could not be replicated for the *man2* mutant line in a separate additional growth trial (**Figure 2.5**). In addition to these growth analyses, root length and germination time were analyzed as motivated by the growth phenotypes described above. *xyn1* roots were 38.57% shorter than those of wildtype, while all other mutant lines' roots were the same length as wild-type (two tailed paired t-test; **Figure 2.6, Figure S2.4**).

2.4.3 The *emp70* mutants germinate earlier than wild-type

Based on physical growth measurements described above, *emp70* plants generally grew better than wild-type as they had larger rosettes, more rosette leaves, and longer stems (**Table S2.4**). Interestingly, analyses of germination time show that 75% of *emp70* mutant seeds germinated 72 hours post-sowing, while the same proportion of wild-type seeds only germinated 120 hours after planting (**Figure 2.7**;

Table S2.3). Important to note that there were no differences in the treatment of mutant and wild-type seeds in terms of storage, time, humidity, temperature etc., so the observed difference in germination time was purely due to the genotype (two-sample equal variance t-test).

2.4.4 Secondary cell wall morphology is not altered in mutants

To assess whether the observed consistent phenotypes were caused by defects in the secondary cell wall structure, vascular bundle morphology of *xyn1*, *emp70*, and *araf1* mutants was investigated through microscopic analyses of transverse sections of bottom parts of stems. All mutant lines investigated showed the normal xylem vessel morphology, represented by the large open and relatively round xylem vessels (**Figure 2.8**). Similarly, there were no discrepancies in the vessel element thickness of mutants and wild-type (**Figure 2.8**; **Figure S2.5**). Thus, the xylem vessel architecture was not impacted by T-DNA insertion mutation in any of the three mutant lines.

2.4.5 The *xyn1* mutants have more stem lignin content

Following microscopic analyses of cell wall morphology, the composition of the cell wall was investigated focusing on structural non-cellulosic monosaccharide composition, lignin, and acetyl content. Stem dry weight was determined before grinding and quantification. Stems from each mutant line weighed the same as wild-type stems (**Figure 2.9**; **Figure 2.10**). This was an interesting observation since *xyn1* and *emp70* mutants' stems did not weight more despite being larger or taller than wild-type stems. It will be worthwhile to follow up this observation with stem density analyses. Similar to stem dry weight, there were no differences observed in the monosaccharide or acetyl content between wild-type and mutants (**Figure 2.9**; **Figure 2.10**; **Figure 2.11**), however, *xyn1* mutants had 18.5% more lignin compared to wild-type (two-sample equal variance t-test; **Figure 2.11**).

2.5 Discussion

In this chapter, we aimed to investigate how altering uncharacterized xylan-associated genes could affect plant growth and physiology of *Arabidopsis thaliana*. The *xyn1* knock-outs had abnormally short roots which may have contributed to the poor vegetative growth, marked by significantly shorter rosette leaves in the mutants. Conversely, the *Arabidopsis emp70* knockout lines germinated earlier than wild-type and showed a growth advantage throughout vegetative growth phase. While these phenotypes were consistent over the three independent growth trials conducted, they were rather mild. More importantly, no phenotypes were observed in *man2*, *araf1*, and *uxs6* mutants. This mild or complete lack of phenotype points to one of the possible limitations of the study, which is the use of the model plant *Arabidopsis thaliana*. While there are indeed many examples of genes affecting secondary cell wall properties in both *Arabidopsis* and woody plants, in the current study, the genes we chose might not have been playing an essential role in *Arabidopsis,* resulting in mild or absent phenotype. It is possible that knocking out the very same gene in *Eucalyptus*, which is a strong carbon sink, may yield an extreme phenotype.

2.5.1 The expression of candidate genes is positively correlated to xylan acetylation

Previous studies have shown that reducing the level of xylan acetylation could attenuate lignocellulose biomass recalcitrance and concomitantly increase saccharification efficiency (Pawar *et al.* 2016, 2017a; Donev *et al.* 2018). With this understanding, genes in BC 3 were selected for functional characterization since their expression profiles were positively correlated with xylan acetylation (**Figure 2.2**). The rationale was that homozygous T-DNA mutations or CRISPR/Cas9 knock-out of these genes would reduce xylan acetylation and thus improve lignocellulose saccharification. Indeed, correlation analyses in the current study indicate that the expression of BC 3 genes is negatively

correlated with pulp yield (**Figure 2.2**), which suggested that knocking these genes out may theoretically improve pulp yield. Interrogation of available literature on the five selected candidate genes showed that *XYN1* and *UXS6* are predominantly expressed in xylem tissues, while *ARAF1* expression is predominant throughout the whole vascular system (Suzuki *et al.* 2002; Fulton and Cobbett 2003; Chavez Montes *et al.* 2008; Zhong *et al.* 2017). We therefore hypothesized that alteration of these genes may affect xylem vessel development, which could ultimately impair water and nutrient transport and consequently alter plant development. Moreover, although *MAN2* and *EMP70* have not yet been showed to be highly expressed specifically in vascular tissues, *MAN2* expression is restricted to micropylar endosperm and radicle, and its expression decreases as the radicle emerges, however, the lack of phenotype in *MAN2* KO mutants suggest that they may not be playing any essential role in cell wall biosynthetic processes, instead, the related members such as *MAN5* and *MAN6* may be the major players (Iglesias-Fernández *et al.* 2011b; a). EMP70 protein family members are the most abundant group of Golgi proteins (Nikolovski *et al.* 2014), regardless of the lack of definite literature evidence for their involvement in plant development thus far, it can be proposed that the ubiquitous nature of encoded proteins specifically within the Golgi may translate to the significance of these proteins in cell wall biosynthetic processes that occur specifically within the Golgi (e.g., modification of lignin biosynthetic enzymes and xylan backbone biosynthesis).

2.5.2 The *xyn1* mutant plants show stunted growth during early vegetative growth stage

With all the above-mentioned evidence linking selected genes to secondary cell wall development, *Arabidopsis thaliana* T-DNA insertional mutants were used for analyses of growth and developmental phenotypes associated with a mutation in each gene. Amongst all T-DNA insertion mutant lines examined, only *xyn1* knock-outs had shorter rosettes compared to wild-type (**Figure 2.3**). To find the possible explanation for this phenotype, seed germination and root length were analyzed; no

differences in germination time was observed in either soil or MS (**Table S2.3**), however, mutant seedlings had notably shorter roots when grown on MS (**Figure 2.6**). The shorter roots may logically present a limitation for efficient water and nutrient transport to the rosette leaves. This is in agreement with the recent observation that overexpression of *XYN1* in *Arabidopsis* leads to enhanced transport of fluorescein from hypocotyl to leaves (Endo *et al.* 2019). In addition to shorter roots, the overall development of vascular bundles may be impaired, leading to limited transport and subsequently shorter rosette leaves. This hypothesis is consistent with previous reports that both *AtXYN1* and a related gene, *BETA-XYLOSIDASE 1* (*AtBXL1*), which also encode xylooligosaccharide-hydrolyzing enzyme, are predominantly expressed specifically in root vascular bundles (Suzuki *et al.* 2002; Goujon *et al.* 2003). However, the root vasculature was not explored in the current study to confirm if it was developing normally, this will indeed be an important future work. In addition to the possible role in root vasculature, *AtXYN1* is amongst eight of the most highly induced cell wall-related genes in the 35S*::LLA23* overexpressing *Arabidopsis* transgenic lines, which are more drought-resistant than wild-type (Wang *et al.* 2013). Similarly, Goujon *et al.* 2003 showed that mutation in *AtBXL1* led to significantly shorter stems compared to wild-type under drought conditions. Taken together, *AtXYN1* may be essential for normal plant development through regulation of the plant's response to water or lack thereof. In line with molecular function suggested by GO analyses (**Table S2.2**), such regulation may likely involve hemicellulose alteration mediated through endo-1,4-β-xylanase (EC 3.2.1.8) activity, allowing breakdown and rerouting of cell wall constituents into the development of secondary vasculature. This enzymatic activity describes the common step occurring during primary cell wall biosynthesis (reviewed in Verbančič *et al.* 2018), it is therefore likely that *XYN1* may be playing a role in primary cell wall biosynthesis instead of specifically secondary cell wall biosynthesis.

Normal *xyn1* plant growth was observed during the late vegetative stage, with inflorescence stems significantly longer than wild-type at 7-weeks post planting (**Figure 2.5**). The fact that potential growth phenotypes were only observed in the early stages of development may suggest that *At*XYN1 is required during the early developmental stages such as rooting, and in its absence, the activity of one of other putative xylanases (*At*XYN2, *At*XYN3, *At*XYN4 , and *At*XYN5) may compensate for the lost *At*XYN1 activity, allowing plant growth to continue normally (Suzuki *et al.* 2002). This possibility is partially supported by the observation that *AtXYN1* expression can be induced by the modest concentrations of a carbon source such as xylose (Xing *et al.* 2013), thus, *XYN1* may be amongst the very first, but not the only player during early cell wall biosynthesis, and in its absence, its role may be complemented by another *XYN1* gene, allowing plant growth to continue normally. Further work is required to test this hypothesis and to accurately verify the enzymatic activity of the other putative xylanases. More importantly, based on the observed phenotypes and additional evidence on the molecular function of XYN1 enzyme – XYN1 may be playing an even more important role in trees in wood forming tissue when the carbon sink is more significant towards cell wall synthesis, this is worth exploring further in model woody systems.

2.5.3 The *araf1* mutants have considerably shorter stems in early bolting but ultimately grow like wild-type

The 5-weeks-old *araf1* T-DNA insertion mutants had significantly shorter stems, this was a notable growth difference compared to wild-type plants (**Figure S2.3**). However, by the seventh week, the mutant stems were the same height as wild-type's (**Figure 2.5**). Comparably, *ARAF1* overexpressing lines have been shown to exhibit significantly delayed inflorescence emergence and altered stem architecture (Chávez Montes et al. 2008). These previous findings and ones generated in this study point to the putative biological role of *ARAF1* to inflorescence stem development. However, little is currently known about the exact biological function and regulation of *Arabidopsis ARAF1* to be able

to conclude this hypothesis. Nonetheless, the enzyme encoded by *AtARAF1* is specifically active during xylem vessel formation and capable of dissolving L-arabinose side-chain residues from arabinoxylan and several cell wall polysaccharides (Kubo et al. 2005; Ichinose et al. 2010). These observations suggest that the shorter inflorescence stems in *araf1* mutants observed in the current study may be due to impaired xylem vessels incapable of sufficient nutrient and water transport. In line with this prospect, the shorter stems in *araf1* in the early bolting stage is similar to the observation following downregulation of *AtBXL1* which were also shorter than wild-type (Goujon et al. 2003). This comparison is more relevant here not only because BXL1 has the α-L-arabinofuranosidase (EC:3.2.1.55) activity, but it is also co-expressed with ARAF1 based on the predicted functional protein-protein association network (Szklarczyk et al. 2019). More importantly, *ARAF1* homolog was found to be amongst the most highly expressed genes in developing tension wood in *Populus tremula* x *P. tremuloides*, suggesting that ARAF1 may target and hydrolyze polysaccharides required for hemicellulose biosynthesis, favoring cellulose biosynthesis and subsequent tension wood formation (Andersson-Gunnerås et al. 2006). Consistent with this suggestion, the upregulation of *ARAF1* during active xylem vessel formation in *Arabidopsis* has been associated with the possible involvement of *ARAF1* in alteration of arabinoxylan during xylem vessel development and possibly plant growth (Fulton and Cobbett 2003; Ichinose et al. 2010). Such alteration may involve side-chain cleavage and xylose release amongst other glycosyl hydrolase activities.

The fact that besides the temporal shorter stems, *araf1* mutant plants growth was otherwise similar to that of wild-type may be due to the compensation activity of *AtARAF2* (AT5G26120), which when compared to *AtARAF1*, its expression is limited to the vasculature of old root and stem tissue in wild-type *Arabidopsis* (Fulton and Cobbett 2003). *ARAF2*, like *ARAF1*, is believed to be involved in the coordinated dissolution of cell wall matrix during secondary xylem vessel development and abscission (Fulton and Cobbett 2003). It is thus possible that in *araf1* mutants, *AtARAF2* expression in roots and

stems allows plant growth to continue normally following bolting. This coincides with the observation that *AtARAF2* expression is observed in old roots, suggesting that its role in plant growth may be important late in plant development. On the contrary, while the compensation by *ARAF2* is a possibility, it is also unlikely since *araf2* mutants show no obvious phenotype, and *araf1araf2* double mutants do not exhibit an additive effect (Chávez Montes *et al.* 2008). It is notably surprising that while *araf1* mutants show a phenotype, the *araf1araf2* does not show any phenotype. More importantly, it would be interesting to investigate why *araf2* seem to rescue the *araf1* phenotype as speculated from the current study.

An alternative explanation for recovery of *araf1* mutants therefore exist; it may be due to the predominance of glucuronoxylan in *Arabidopsis* instead of arabinoxylan, the latter of which *ARAF1* is thought to predominantly target for modulation of plant growth in monocots (Fulton and Cobbett 2003). Therefore, *ARAF1* may not be playing a significant role in *Arabidopsis* cell wall modification and overall plant development as it does in monocots. Supporting this point, overexpression of *OsARAF1* in rice leads to 25% reduction in arabinose accompanied with 28% increase in the glucose yield (Sumiyoshi *et al.* 2013), showing the clear involvement of *ARAF1* in cell wall development of monocots.

2.5.4 Mutation in *EMP70* confers consistent growth advantage during vegetative growth

Amongst the genes chosen for phenotypic analyses in this study, *EMP70* had the least prior information available for reference. Regardless, the EMP70 (or TM9) family, represented by 12 isoforms of unknown function in *Arabidopsis thaliana*, is the most abundant protein family in the Golgi proteome (Gao *et al.* 2012; Nikolovski *et al.* 2012, 2014). Notably, in addition to being the most abundant in the Golgi, the *Arabidopsis* EMP70 members have also been shown to localize exclusively

to the Golgi (Dunkley *et al.* 2006; Gao *et al.* 2012). Furthermore, the EMP70 family is amongst the three identified type III protein families that likely function as channels or transporters (Nikolovski *et al.* 2012). These localization studies suggest that EMP70 may be involved in Golgi transport of various metabolites possibly necessary for polysaccharide biosynthesis. While such prospect has not yet been tested in *Arabidopsis* or any other plant species at the time of writing, studies of the TM9 physiological function in *Saccharomyces cerevisiae* have given an insight into the possible role of this protein family. A yeast *tmn1-3* (where *TMN3* is AT2G01970 homologue in yeast) triple mutant exhibited a 25% reduction in cellular copper (Cu) content (Hegelund *et al.* 2010). Furthermore, overexpression of *TMN1* in yeast leads to more than double the concentration of cellular Cu (Hegelund *et al.* 2010). Similarly, heterologous expression of the *Arabidopsis* TM9 homolog, *AtTM7* (AT3G13772) in yeast leads to an increase in cellular Cu content (Hegelund *et al.* 2010). These observations led to the conclusion that yeast and possibly *Arabidopsis* EMP70 homologs are involved in the regulation of intracellular Cu balance; a role that may involve Golgi membrane-trafficking (Gao *et al.* 2012; Nikolovski *et al.* 2014).

Cu plays an essential role as a cofactor of several enzymes including some involved in lignin biosynthesis (Yruela 2005), however, due to its redox reactions, Cu is extremely toxic to cells at high concentrations, fortunately, plants have evolved some sophisticated mechanisms to relieve copper toxicity. Thus, considering the significance of intracellular Cu balance, if *EMP70* plays an indispensable role in Cu homeostasis, a homozygous mutation in this gene could alter plant growth to some extent. Here, the very first attempt at functional testing of *Arabidopsis emp70* mutants is reported. Mutants had significantly bigger rosettes, longer inflorescence stems, and an average of one more rosette leaf compared to wild-type plants (**Figure 2.3**; **Figure 2.4**; **Figure 2.5**). The observed consistency in growth advantage questioned the possibility that these mutants may perhaps be germinating earlier than wild-type such that they continue to grow vegetatively bigger compared to

wild-type, or perhaps the root morphology (e.g. length and root hair density) may be rendering some growth advantage to the mutants (Bahmani *et al.* 2016). These hypotheses were both investigated, and expectedly, *emp70* mutant seeds germinated ~2 days earlier than wild-type when grown on soil (**Figure 2.7**). Germination is one of the most crucial steps of plant development and it is well documented that hormones abscisic acid (ABA) and Gibberellic acid (GA) play predominant roles in its regulation (Finch-Savage and Leubner-Metzger 2006). This primarily involves stringently controlling the concentration of ABA and GA, a process that is affected by several environmental cues such as light and water. In general, ABA promotes seed dormancy, while, at the right time, GA, an ABA antagonist, counteracts the activity of ABA to promote seed germination (Gao and Ayele 2014; Shu *et al.* 2015, 2016). Furthermore, the regulation of seed germination through the crosstalk between ABA and GA is highly sensitive to the concentration of heavy metals such as Cu (Xiong and Wang 2005). It has subsequently been demonstrated that Cu inhibits seed germination mostly through suppression of ABA catabolism in rice (Ye *et al.* 2014). Indeed, measurement of ABA under increased Cu concentration conditions showed that ABA content in germinating seeds was significantly higher than in control plants (Ye *et al.* 2014). In the current study, the *emp70* knock-outs would be expected to have significantly lower cellular Cu concentrations, as observed in yeast *tm1-3* knock-out lines (Hegelund *et al.* 2010). Lower Cu levels may lead to reduced ABA concentration in mutants through the suggested mechanism in rice (Ye *et al.* 2014), which could concomitantly lead to the observed enhancement of seed germination as promoted by the action of GA (**Figure 2.7**; **Figure 4.1**). Nonetheless, the crosstalk between ABA and Cu concentration on the regulation of seed germination are not well understood, and future work will be necessary to measure the levels of ABA and GA in the *Arabidopsis EMP70* mutants. Furthermore, the limited functional information about *EMP70* in plants makes it difficult to speculate the exact regulatory role of *EMP70* and how such role is affected by other regulatory cues and enzymes. Moreover, since *Arabidopsis emp70* phenotypic analyses performed here clearly showed a growth advantage compared to wild-type, further work is required to

link this phenotype to Cu transport or concentration, such future work may involve growing mutants in both Cu deficient and Cu supplemented medium followed by an assessment of phenotype.

2.5.5 No major growth phenotypes are observed in *uxs6* and *man2* knock-outs

The growth phenotype of the *Arabidopsis uxs6* T-DNA single mutants have been previously identified by Kuang *et al.* 2016, in the current study, this previously identified phenotype made *uxs6* a good control since the growth phenotype, secondary cell wall chemistry and vessel morphology was all known. Like in the previous study, *uxs6* mutants showed no obvious growth phenotype, nor alteration in secondary cell wall architecture (**Table S2.4)**, indicating that the growth conditions and subsequent analyses were properly performed. Besides the lack of phenotype in *uxs6* mutants, *man2* T-DNA insertion mutants also showed no major growth phenotype. ENDO-BETA-MANNANASE (MAN; EC Number: 3.2. 1.78) plays a fundamental role in plant growth and development by digesting mannan-polysaccharides. In *Arabidopsis*, the *MAN* family is represented by eight members, amongst these, *AtMAN2*, *AtMAN7*, *AtMAN5*, and *AtMAN6* are expressed in germinating seeds (Iglesias-Fernández *et al.* 2011b). *In situ* hybridisation showed that transcript accumulation of these *MAN* genes is restricted to micropylar endosperm and radicle, and their expression disappears immediately after radicle emergence (Iglesias-Fernández *et al.* 2011a). These observations suggested that *AtMANs* play an important role in the germination of *Arabidopsis* through hydrolysis of mannan-rich endosperm cell walls, thereby facilitating radicle emergence (Iglesias-Fernández *et al.* 2011b; a). The involvement of endo-β-mannanase in endosperm cap weakening to facilitate germination has also been confirmed in several species such *Capsicum* spp. and *Coffea arabica* L. (Da Silva *et al.* 2005; Caixet *et al.* 2014). Earlier analyses of the *Arabidopsis* T-DNA insertion in these four *AtMAN* genes reported that mutants (except mutation in *AtMAN2*) germinate later than wild-type (Iglesias-Fernández *et al.* 2011a). Amongst the mutants with negatively affected germination time course, *AtMAN6* exhibited the most

delay in germination (Iglesias-Fernández *et al.* 2011a; b). The observation that only *AtMAN6* knock-out show extreme germination delay may suggest that while *AtMAN2* may be required for endosperm weakening during germination, it may not be playing a significant overall role in this process, hence no phenotype was observed in either the previous or current study (**Table S2.4**). Without ruling out the possibility that *AtMAN2* may be necessary for seed development, the lack of germination phenotype in *AtMAN2* knock-out mutants indicates that this gene may play a different role from that of the other three *MANs*, which may, or may not involve the mannan endo-1,4-β-mannosidase activity. Further work is needed to uncover such a potential role. In the interim, information from STRING (https://string-db.org/.0t67) show that *AtMAN2* have protein-protein interactions with four hexokinases (Szklarczyk *et al.* 2019). With diverse roles of hexokinase family, it is not obvious what functional role of MAN2 is associated with these hexokinases (Moore *et al.* 2003; Cho *et al.* 2006). However, as HEXOKINASE 1 (HXK1) and HXK2 mediate the effects of sugar on plant growth and development (Granot *et al.* 2013; Kunz *et al.* 2015; Laurian *et al.* 2019), it would be worthwhile to find out how the role of *AtMAN2* possibly relates to this or any other HXK role in plant growth and development. While these questions require further studies, the current study has confirmed that, although *AtMAN2* belong to the same group of genes that are expressed in germinating seeds and whose T-DNA mutants show delayed germination (Iglesias-Fernández *et al.* 2011b), *man2* mutants do not show this delayed germination phenotype (**Table S2.3**), neither do the plants show any other alteration of growth or physiology (**Table S2.4**). More importantly, the current study has for the first time, demonstrated that in addition to the lack of effect on plant growth, the secondary cell wall morphology and chemistry is not affected in *man2* mutants (**Figure 2.8**; **Figure 2.10; Figure 2.11**).

2.5.6 Xylem vessel morphology is not altered in mutants

To investigate whether any of the consistent growth phenotypes were due to impaired secondary cell walls, stem transverse sections of *xyn1*, *emp70* and *araf1* were examined through light microscopy. All three mutant lines had no morphological defects in the secondary cell wall structure (**Figure 2.8**). Considering that both *xyn1* and *araf1* mutants were able to recover and grow as well as wild-type, it was expected that they would most likely be no overall alteration of the cell wall structure. However, if the shorter rosettes observed in *xyn1* culminated in an overall reduction in plant growth similar to *irx* mutants (Petersen *et al.* 2012), it would then be more likely that this phenotype is caused by collapsed xylem vessels and the resulting impairment of water and nutrient transport, this would be reflected on the micrographs accordingly. In contrast to *xyn1* and *araf1*, *emp70* showed a consistent growth advantage, however, as previously discussed, the limited functional information about this gene family makes it difficult to speculate on its exact role and similarly, what could be expected in terms of xylem vessel morphology cannot be easily deduced. Finally, with the lack of obvious growth phenotype in *uxs6* and *man2* mutant lines, no obvious changes in cell wall morphology would be expected for these lines.

2.5.7 Cell wall chemistry of all mutants resembles that of wild-type

In all mutants, the non-cellulosic monosaccharide content was not affected (**Figure 2.9**; **Figure 2.10**). Since ARAF1 and XYN1 both possess the xylan hydrolase activity (**Table S2.2**), it was expected that in knock-out mutants corresponding to these genes, monosaccharide levels will not be impacted. Furthermore, the observation that vessel element thickness was not altered in mutants hinted that the cell wall composition, specifically D-Xyl level, is less likely to be altered. In agreement, no differences in D-Xyl content were observed in *araf1-1* (T-DNA insertional mutant in Ws background) when monosaccharide composition of dry seeds, roots and stems were analyzed (Chavez Montes et al. 2008).

Conversely, in 35S::*ARAF1* overexpressing lines, both D-Xyl and L-Ara content were significantly higher compared to wild-type, although the difference was minor (Chavez Montes et al. 2008). This is expected as xylose and arabinose would accumulate when arabinoxylan is hydrolyzed through ARAF1's xylan hydrolase and xylan 1,4-β-xylosidase activity. With this understanding, a *man2* overexpressing line may also present increased mannan content. Together, the lack of major alteration of monosaccharide content in all mutants indicates that even if these genes are important in cell wall development, their contribution to the overall monosaccharide biosynthesis can be compensated for if mutated. *UXS* functional redundancy has already been provided as an explanation for the overall lack of phenotype in *uxs6* single mutants. Furthermore, in instances where a mutation in these genes may alter monosaccharide content, there may be significant toleration of plasticity in the monosaccharide amount, allowing the overall growth to continue normally as observed in all mutants (Kuang *et al.* 2016). Importantly, the current study is limited in that it only assessed monosaccharide content in stems, it would be of interest to further expand this analysis into other organs and different growth stages, this is particularly necessary in the case of *MAN2*, which is predominantly expressed in germinating seeds (Iglesias-Fernández et al. 2011b). Similarly, the observation that *emp70* mutants germinate earlier than wild-type may necessitate analyses of cell wall chemistry of the seeds during the germination stage.

2.5.8 The *xyn1* mutants have higher lignin content

Lignin and xylan acetylation are the major contributors to biomass recalcitrance and their reduction has proven to be effective in improving saccharification yields (Acker *et al.* 2014; Pawar *et al.* 2017a; b). To find out if a mutation in selected genes can likely lead to improvement of biomass saccharification, lignin content was quantified. *xyn1* mutants had higher lignin content compared to wild-type, while lignin content of all other mutants resembled that of wild-type (**Figure 2.11**). It is

possible that the taller stems observed late during *xyn1* vegetative development (**Figure 2.5**) may have been due to the increased secondary cell wall development, although such increase may not have been significant enough to be reflected either in xylem vessel morphology or monosaccharide quantification. Nonetheless, the increased lignin content observed in *xyn1* stems may be reflecting such increase in secondary cell wall development. In addition to lignin quantification, acetate levels were also quantified and no difference in the levels were observed between mutants and wild-type. The lack of impact on the acetate content in all the mutants was expected considering that none of these genes are implicated in acetylation processes at least based on available literature. However, the observation in *Eucalyptus* population that the expression of these genes is positively correlated with xylan acetylation would suggest that knocking them out may lead to reduction of acetyl content at least in *Eucalyptus* or trees in general (**Figure 2.2**). The fact that this is not the case may indicate that, although associated to xylan acetylation, these genes may not play a major role in processes associated with acetylation. Finally, while saccharification analyses are pending, it is possible that the sugar release efficiency will be reduced in *xyn1* mutants as they have high lignin content, a trait known to generally impede saccharification as it reduces access to cellulose during bioprocessing (Yang and Wyman 2004). Conversely, besides cellulose accessibility, the relative sugar content in the biomass may affect the amount of sugar released, since none of the mutants had altered sugar content (**Table S2.4**), the saccharification yield will most likely not be affected.

2.5.9 The possible involvement of candidate genes in primary cell wall development

The lack of consistent and/or significant phenotype observed from all mutants here may suggest that the targeted genes may be playing a more secondary role in primary cell wall development, instead of specifically involved in secondary cell wall biosynthesis. Such roles may involve modification, reorganization, and metabolic salvage of wall sugars during plant cell growth, all of which are well

documented activities necessary throughout primary cell wall biosynthesis (Extensively reviewed in Barnes and Anderson 2018b). This primary wall remodeling is known to stimulate cell expansion, an important step during secondary cell wall deposition (Sun *et al.* 2010; Paque *et al.* 2014). However, how such primary cell wall recycling impacts the overall biosynthesis of the secondary cell wall is largely ununderstood. Broad research that integrates important aspects of both primary and secondary cell wall development is therefore highly recommended. Nevertheless, amongst the selected candidate genes, *ARAF1* is most likely to be involved in primary cell wall biosynthesis in which xylan contains the Ara*p* substitution (Mortimer *et al.* 2015). It is therefore highly recommended to explore this possibility perhaps through both gene knockout and overexpression of *ARAF1*.

In addition to the possible role in primary cell wall development, it will be important to find out whether double mutants of candidate genes and their close homologs could yield any phenotypes. From the selected candidate genes, *AtMAN2*, *AtUXS6,* and *AtEMP70* have one *Arabidopsis* close paralog each; ATG28320 (*MAN5*), AT2G28755 (*UDP-D-glucuronate carboxy-lyase-like*) and AT1G14670 (*EMP70*) respectively (**Figure 3.1, Figure S 3.1**). The *AtMAN2* paralog is the *AtMAN5*, which is expressed during seed germination as previously described. More importantly, different from *AtMAN2, AtMAN5* mutation leads to delayed seed germination (Iglesias-Fernández *et al.* 2011b; Wang *et al.* 2015). It would be interesting to see if overexpression of *MAN2* can rescue the delayed seed germination phenotype of the *MAN5* since *MAN2* do not display any alteration to germination. Interestingly, the *UXS6* paralog is amongst one of the most highly induced genes (together with *XYN1*) in *Arabidopsis* overexpressing *LLA23* genes, these plants are highly resistant to drought (Yang *et al.* 2008). This poses a question about the possible role of this homolog in water transport, and it will be worthwhile to assess the phenotype of double mutant of this gene and *UXS6*. Finally, the *EMP70* homolog is uncharacterised just like *EMP70*. Since the phenotype of *emp70* is now established in the current study, a double knockout in this gene and homolog may lead to an addictive effect, with plants

germinating significantly earlier and plants growing better than wild-type. However, this is considering that the either of these genes act redundantly.

2.6 References

Acker R. Van, J. C. Leple, D. Aerts, V. Storme, G. Goeminne, *et al.*, 2014 Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. Proc. Natl. Acad. Sci. 111: 845–850. https://doi.org/10.1073/pnas.1321673111

Andersson-Gunnerås S., E. J. Mellerowicz, J. Love, B. Segerman, Y. Ohmiya, *et al.*, 2006 Biosynthesis of cellulose-enriched tension wood in *Populus*: Global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. Plant J. 45: 144–165. https://doi.org/10.1111/j.1365-313X.2005.02584.x

Bahmani R., D. G. Kim, J. A. Kim, and S. Hwang, 2016 The density and length of root hairs are enhanced in response to cadmium and arsenic by modulating gene expressions involved in fate determination and morphogenesis of root hairs in *Arabidopsis*. Front. Plant Sci. 7: 1–16. https://doi.org/10.3389/fpls.2016.01763

Barnes W. J., and C. T. Anderson, 2018 Release, recycle, rebuild: Cell-wall remodeling, autodegradation, and sugar salvage for new wall biosynthesis during plant development. Mol. Plant 11: 31–46. https://doi.org/10.1016/j.molp.2017.08.011

Basen M., A. M. Rhaesa, I. Kataeva, C. J. Prybol, I. M. Scott, *et al.*, 2014 Degradation of high loads of crystalline cellulose and of unpretreated plant biomass by the thermophilic bacterium *Caldicellulosiruptor besci*i. Bioresour. Technol. 152: 384–392. https://doi.org/10.1016/j.biortech.2013.11.024

Bonawitz N. D., and C. Chapple, 2010 The Genetics of lignin biosynthesis: Connecting genotype to phenotype. Annu. Rev. Genet. 44: 337–363. https://doi.org/10.1146/annurev-genet-102209-163508

Boudet A. M., 2000 Lignins and lignification: Selected issues. Plant Physiol. Biochem. 38: 81–96. https://doi.org/10.1016/S0981-9428(00)00166-2

Bromley J. R., M. Busse-Wicher, T. Tryfona, J. C. Mortimer, Z. Zhang, *et al.*, 2013 GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. Plant J. 74: 423–434. https://doi.org/10.1111/tpj.12135

Brown D. M., Z. Zhang, E. Stephens, P. Dupree, and S. R. Turner, 2009 Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. Plant J. 57: 732–746. https://doi.org/10.1111/j.1365-313X.2008.03729.x

Caixet F., E. V. D. R. Von Pinh, R. M. Guimarães, P. Henrique, A. R. Pereira, *et al.*, 2014 Physiological and biochemical alterations during germination and storage of habanero pepper seeds. African J. Agric. Res. 9: 627–635. http://dx.doi.org/10.5897/AJAR2013.7133

Chavez Montes R. A., P. Ranocha, Y. Martinez, Z. Minic, L. Jouanin, *et al.*, 2008 Cell wall modifications in *Arabidopsis* plants with altered α-L-Arabinofuranosidase activity. Plant Cell Physiol. 147: 63–77. https://doi.org/10.1104/pp.107.110023

Chen F., and R. A. Dixon, 2007 Lignin modification improves fermentable sugar yields for biofuel production. Nat. Biotechnol. 25: 759–761. https://doi.org/10.1038/nbt1316

Chen X., M. E. Vega-Sánchez, Y. Verhertbruggen, D. Chiniquy, P. E. Canlas, *et al.*, 2013 Inactivation of *OsIRX10* leads to decreased xylan content in rice culm cell walls and improved biomass saccharification. Mol. Plant 6: 570–573. https://doi.org/10.1093/mp/sss135

Chiniquy D., P. Varanasi, T. Oh, J. Harholt, J. Katnelson, *et al.*, 2013 Three novel rice genes closely related to the *Arabidopsis IRX9*, *IRX9L*, and *IRX14* genes and their roles in xylan biosynthesis. Front. Plant Sci. 4: 1–13. https://doi.org/10.3389/fpls.2013.00083

Cho Y. H., S. D. Yoo, and J. Sheen, 2006 Regulatory functions of nuclear HEXOKINASE1 complex in glucose signaling. Cell 127: 579–589. https://doi.org/10.1016/j.cell.2006.09.028

Chong S. L., M. Derba-Maceluch, S. Koutaniemi, L. D. Gómez, S. J. McQueen-Mason, *et al.*, 2015 Active fungal GH115 aα-glucuronidase produced in *Arabidopsis thaliana* affects only the UX1-reactive glucuronate decorations on native glucuronoxylans. BMC Biotechnol. 15: 1–13. https://doi.org/10.1186/s12896-015-0154-8

Donev E., M. L. Gandla, L. J. Jönsson, and E. J. Mellerowicz, 2018 Engineering non-cellulosic polysaccharides of wood for the biorefinery. Front. Plant Sci. 871. https://doi.org/10.3389/fpls.2018.01537

Dunkley T. P. J., S. Hester, I. P. Shadforth, J. Runions, T. Weimar, *et al.*, 2006 Mapping the *Arabidopsis* organelle proteome. Proc. Natl. Acad. Sci. 103: 6518–6523. https://doi.org/10.1073/pnas.0506958103

Ebringerová A., and T. Heinze, 2000 Xylan and xylan derivatives - Biopolymers with valuable properties, 1: Naturally occurring xylans structures, isolation procedures and properties. Macromol. Rapid Commun. 21: 542–556. https://doi.org/10.1002/1521-3927(20000601)21:9<542::AID-MARC542>3.0.CO;2-7

Endo S., Y. Iwai, and H. Fukuda, 2019 Cargo-dependent and cell wall-associated xylem transport in *Arabidopsis*. New Phytol. 222: 159–170. https://doi.org/10.1111/nph.15540

Finch-Savage W. E., and G. Leubner-Metzger, 2006 Seed dormancy and the control of germination. New Phytol. 171: 501–523. https://doi.org/10.1111/j.1469-8137.2006.01787.x

Fulton L. M., and C. S. Cobbett, 2003 Two *α-L-ARABINOFURANOSIDASE* genes in *Arabidopsis thaliana* are differentially expressed during vegetative growth and flower development. J. Exp. Bot. 54: 2467–2477. https://doi.org/10.1093/jxb/erg269

Gandla M. L., C. Martín, and L. J. Jönsson, 2018 Analytical enzymatic saccharification of lignocellulosic biomass for conversion to biofuels and bio-based chemicals. Energies 11. https://doi.org/10.3390/en11112936

Gao C., K. Y. Li, S. Qu, M. W. Y. San, C. K. Y. Yu, *et al.*, 2012 The Golgi-Localized Arabidopsis Endomembrane Protein12 Contains Both Endoplasmic Reticulum Export and Golgi Retention Signals at Its C Terminus. Plant Cell 24: 2086–2104. https://doi.org/10.1105/tpc.112.096057

Gao F., and B. T. Ayele, 2014 Functional genomics of seed dormancy in wheat: Advances and prospects. Front. Plant Sci. 5: 1–11. https://doi.org/10.3389/fpls.2014.00458

González-Vila F. J., G. Almendros, J. C. Del Rio, F. Martín, A. Gutiérrez, *et al.*, 1999 Ease of delignification assessment of wood from different *Eucalyptus* species by pyrolysis (TMAH)-GC/MS and CP/MAS 13C-NMR spectrometry. J. Anal. Appl. Pyrolysis 49: 295–305. https://doi.org/10.1016/S0165-2370(98)00097-7

Goujon T., Z. Minic, A. El Amrani, O. Lerouxel, E. Aletti, *et al.*, 2003 *AtBXL1*, a novel higher plant (*Arabidopsis thaliana*) putative beta-xylosidase gene, is involved in secondary cell wall metabolism and plant development. Plant J. 33: 677–690. https://doi.org/10.1046/j.1365-313X.2003.01654.x

Granot D., R. David-Schwartz, and G. Kelly, 2013 Hexose kinases and their role in sugar-sensing and plant development. Front. Plant Sci. 4: 1–17. https://doi.org/10.3389/fpls.2013.00044

Gu Y., N. Kaplinsky, M. Bringmann, A. Cobb, A. Carroll, *et al.*, 2010 Identification of a cellulose synthase-associated protein required for cellulose biosynthesis. Proc. Natl. Acad. Sci. 107: 12866–12871. https://doi.org/10.1073/pnas.1007092107

Hao Z., and D. Mohnen, 2014 A review of xylan and lignin biosynthesis: Foundation for studying *Arabidopsis* irregular xylem mutants with pleiotropic phenotypes. Crit. Rev. Biochem. Mol. Biol. 49: 212–241. https://doi.org/10.3109/10409238.2014.889651

Hegelund J. N., T. P. Jahn, L. Baekgaard, M. G. Palmgren, and J. K. Schjoerring, 2010 Transmembrane nine proteins in yeast and *Arabidopsis* affect cellular metal contents without changing vacuolar morphology. Physiol. Plant. 140: 355–367. https://doi.org/10.1111/j.1399-3054.2010.01404.x

Helle S., D. Cameron, J. Lam, B. White, and S. Duff, 2003 Effect of inhibitory compounds found in biomass hydrolysates on growth and xylose fermentation by a genetically engineered strain of *S. cerevisiae*. Enzyme Microb. Technol. 33: 786–792. https://doi.org/10.1016/S0141-0229(03)00214-X

Himmel M. E., S. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, *et al.*, 2007 Biomass recalcitrance: Engineering plants and enzymes for biofuels production. Nature 454: 804–807. https://doi.org/10.1126/science.1137016

Ichinose H., N. Nishikubo, T. Demura, and S. Kaneko, 2010 Characterization of α-L-arabinofuranosidase related to the secondary cell walls formation in *Arabidopsis thaliana*. Plant Biotechnol. 27: 259–266. https://doi.org/10.5511/plantbiotechnology.27.259

Iglesias-Fernández R., M. C. Rodríguez-Gacio, C. Barrero-Sicilia, P. Carbonero, and A. Matilla, 2011a Three *ENDO-β-MANNANASE* genes expressed in the micropylar endosperm and in the radicle influence germination of *Arabidopsis thaliana* seeds. Planta 233: 25–36. https://doi.org/10.1007/s00425-010-1257-z

Iglesias-Fernández R., M. del C. Rodríguez-Gacio, C. Barrero-Sicilia, P. Carbonero, and A. J. Matilla, 2011b Molecular analysis of *ENDO-β-MANNANASE* genes upon seed imbibition suggest a cross-talk between radicle and micropylar endosperm during germination of *Arabidopsis thaliana*. Plant Signal. Behav. 6: 80–82. https://doi.org/10.4161/psb/14126

Johnson A. M., H. Kim, J. Ralph, and S. D. Mansfield, 2017 Natural acetylation impacts carbohydrate recovery during deconstruction of *Populus trichocarpa* wood. Biotechnol. Biofuels 10: 1–12. https://doi.org/10.1186/s13068-017-0734-z

Kang B. H., E. Nielsen, M. L. Preuss, D. Mastronarde, and L. A. Staehelin, 2011 Electron tomography of RabA4b- and PI-4Kβ1-labeled *trans* Golgi Network Compartments in *Arabidopsis*. Traffic 12: 313–329. https://doi.org/10.1111/j.1600-0854.2010.01146.x

Kang X., A. Kirui, M. C. Dickwella Widanage, F. Mentink-Vigier, D. J. Cosgrove, *et al.*, 2019 Lignin-polysaccharide interactions in plant secondary cell walls revealed by solid-state NMR. Nat. Commun. 10: 1–9. https://doi.org/10.1038/s41467-018-08252-0

Kopycki J. G., D. Rauh, A. A. Chumanevich, P. Neumann, T. Vogt, *et al.*, 2008 Biochemical and structural analysis of substrate promiscuity in plant Mg^{2+}-dependent O-methyltransferases. J. Mol. Biol. 378: 154–164. https://doi.org/10.1016/j.jmb.2008.02.019

Kuang B., X. Zhao, C. Zhou, W. Zeng, J. Ren, *et al.*, 2016 Role of UDP-Glucuronic acid decarboxylase in xylan biosynthesis in *Arabidopsis*. Mol. Plant 9: 1119–1131. https://doi.org/10.1016/j.molp.2016.04.013

Kubo M., M. Udagawa, N. Nishikubo, G. Horiguchi, M. Yamaguchi, *et al.*, 2005 Transcription switches for protoxylem and metaxylem vessel formation Minoru. Genes Dev. 1855–1860. https://doi.org/10.1101/gad.1331305.

Kunz S., P. Gardeström, E. Pesquet, and L. A. Kleczkowski, 2015 Hexokinase 1 is required for glucose-induced repression of *bZIP63*, *At5g22920*, and *BT2* in *Arabidopsis*. Front. Plant Sci. 6: 1–12. https://doi.org/10.3389/fpls.2015.00525

Laurian R., K. Dementhon, B. Doumèche, A. Soulard, T. Noel, *et al.*, 2019 Hexokinase and glucokinases are essential for fitness and virulence in the pathogenic yeast *Candida albicans*. Front. Microbiol. 10: 1–18. https://doi.org/10.3389/fmicb.2019.00327

Lee C., Q. Teng, W. Huang, R. Zhong, and Z.-H. Ye, 2010 The *Arabidopsis* family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. Plant Physiol. 153: 526–541. https://doi.org/10.1104/pp.110.155309

Lee C., Q. Teng, R. Zhong, and Z. H. Ye, 2011 The four *Arabidopsis REDUCED WALL ACETYLATION* genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. Plant Cell Physiol. 52: 1289–1301. https://doi.org/10.1093/pcp/pcr075

Lee C., Q. Teng, R. Zhong, and Z. H. Ye, 2012a *Arabidopsis* GUX proteins are glucuronyltransferases responsible for the addition of glucuronic acid side chains onto xylan. Plant Cell Physiol. 53: 1204–1216. https://doi.org/10.1093/pcp/pcs064

Lee C., Q. Teng, R. Zhong, Y. Yuan, M. Haghighat, *et al.*, 2012b Three *Arabidopsis* DUF579 domain-containing GXM proteins are methyltransferases catalyzing 4-O-methylation of glucuronic acid on xylan. Plant Cell Physiol. 53: 1934–1949. https://doi.org/10.1093/pcp/pcs138

Lee C., R. Zhong, and Z. H. Ye, 2012c Biochemical characterization of xylan xylosyltransferases involved in wood formation in poplar. Plant Signal. Behav. 7. https://doi.org/10.4161/psb.19269

Lee C., Q. Teng, R. Zhong, Y. Yuan, and Z. Ye, 2014 Functional roles of rice glycosyltransferase family GT43 in xylan biosynthesis. Plant Signal. Behav. e27809: 1–13. https://doi.org/10.4161/psb.27809

Lunin V., H.-T. Wang, V. Bharadwaj, P. M. Alahuhta, M. J. Pena, *et al.*, 2020 Molecular mechanism of polysaccharide acetylation by the *Arabidopsis* xylan O-acetyltransferase XOAT1. Plant Cell. https://doi.org/10.1105/tpc.20.00028

Maloney V. J., A. L. Samuels, and S. D. Mansfield, 2012 The endo-1,4-β-glucanase KORRIGAN exhibits functional conservation between gymnosperms and angiosperms and is required for proper cell wall formation in gymnosperms. New Phytol. 193: 1076–1087. https://doi.org/10.1111/j.1469-8137.2011.03998.x

McDougall G. J., I. M. Morrison, D. Stewart, J. D. B. Weyers, and J. R. Hillman, 1993 Plant fibres: Botany, chemistry and processing for industrial use. J. Sci. Food Agric. 62: 1–20. https://doi.org/10.1002/jsfa.2740620102

Mizrachi E., L. Verbeke, N. Christie, A. C. Fierro, S. D. Mansfield, *et al.*, 2017 Network-based integration of systems genetics data reveals pathways associated with lignocellulosic biomass accumulation and processing. Proc. Natl. Acad. Sci. 114: 1195–1200. https://doi.org/10.1073/pnas.1620119114

Moore B., L. Zhou, F. Rolland, Q. Hall, W. H. Cheng, *et al.*, 2003 Role of the *Arabidopsis* glucose sensor HXK1 in nutrient, light, and hormonal signaling. Science. 300: 332–336. https://doi.org/10.1126/science.1080585

Mortimer J. C., G. P. Miles, D. M. Brown, Z. Zhang, M. P. Segura, *et al.*, 2010 Absence of branches from xylan in *Arabidopsis gux* mutants reveals potential for simplification of lignocellulosic biomass. Proc. Natl. Acad. Sci. 107: 17409–17414. https://doi.org/10.1073/pnas.1005456107

Mortimer J. C., N. Faria-Blanc, X. Yu, T. Tryfona, M. Sorieul, *et al.*, 2015 An unusual xylan in Arabidopsis primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14. Plant J. 83: 413–426. https://doi.org/10.1111/tpj.12898

Myburg A. A., D. Grattapaglia, G. A. Tuskan, U. Hellsten, R. D. Hayes, *et al.*, 2014 The genome of *Eucalyptus grandis*. Nature 510: 356–362. https://doi.org/10.1038/nature13308

Nikolovski N., D. Rubtsov, M. P. Segura, G. P. Miles, T. J. Stevens, *et al.*, 2012 Putative glycosyltransferases and other plant Golgi apparatus proteins are revealed by LOPIT proteomics. Plant Physiol. 160: 1037–1051. https://doi.org/10.1104/pp.112.204263

Nikolovski N., P. V. Shliaha, L. Gatto, P. Dupree, and K. S. Lilley, 2014 Label-free protein quantification for plant golgi protein localization and abundance. Plant Physiol. 166: 1033–1043. https://doi.org/10.1104/pp.114.245589

Nunes C. A., C. F. Lima, L. C. A. Barbosa, J. L. Colodette, A. F. G. Gouveia, *et al.*, 2010 Determination of Eucalyptus spp lignin S/G ratio: A comparison between methods. Bioresour. Technol. 101: 4056–4061. https://doi.org/10.1016/j.biortech.2010.01.012

Paque S., G. Mouille, L. Grandont, D. Alabadí, C. Gaertner, *et al.*, 2014 AUXIN BINDING PROTEIN1 links cellwall remodeling, auxin signaling, and cell expansion in *Arabidopsis*. Plant Cell 26: 280–295. https://doi.org/10.1105/tpc.113.120048

Pauly M., and K. Keegstra, 2008 Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J. 54: 559–568. https://doi.org/10.1111/j.1365-313X.2008.03463.x

Pawar P. M. A., M. Derba-Maceluch, S. L. Chong, L. D. Gómez, E. Miedes, *et al.*, 2016 Expression of fungal acetyl xylan esterase in *Arabidopsis thaliana* improves saccharification of stem lignocellulose. Plant Biotechnol. J. 14: 387–397. https://doi.org/10.1111/pbi.12393

Pawar P. M. A., M. Derba-Maceluch, S. L. Chong, M. L. Gandla, S. S. Bashar, *et al.*, 2017a *In muro* deacetylation of xylan affects lignin properties and improves saccharification of aspen wood. Biotechnol. Biofuels 10: 1–11. https://doi.org/10.1186/s13068-017-0782-4

Pawar P. M. A., C. Ratke, V. K. Balasubramanian, S. L. Chong, M. L. Gandla, *et al.*, 2017b Downregulation of *RWA* genes in hybrid aspen affects xylan acetylation and wood saccharification. New Phytol. 214: 1491–1505. https://doi.org/10.1111/nph.14489

Pellny T. K., A. Patil, A. J. Wood, J. Freeman, K. Halsey, *et al.*, 2020 Loss of *TaIRX9b* gene function in wheat decreases chain length and amount of arabinoxylan in grain but increases cross-linking. Plant Biotechnol. J. https://doi.org/10.1111/pbi.13393

Peña M. J., R. Zhong, G. K. Zhou, E. A. Richardson, M. A. O'Neill, *et al.*, 2007 *Arabidopsis irregular xylem8* and *irregular xylem9*: Implications for the complexity of glucuronoxylan biosynthesis. Plant Cell 19: 549–563. https://doi.org/10.1105/tpc.106.049320

Petersen P. D., J. Lau, B. Ebert, F. Yang, Y. Verhertbruggen, *et al.*, 2012 Engineering of plants with improved properties as biofuels feedstocks by vessel-specific complementation of xylan biosynthesis mutants. Biotechnol. Biofuels 5: 1–19. https://doi.org/10.1186/1754-6834-5-84

Petrik D. L., T. Tryfona, P. Dupree, and C. T. Anderson, 2020 *BdGT43B2* functions in xylan biosynthesis and is essential for seedling survival in *Brachypodium distachyon*. Plant Direct 4: 1–16. https://doi.org/10.1002/pld3.216

Pinard D., A. C. Fierro, K. Marchal, A. A. Myburg, and E. Mizrachi, 2019 Organellar carbon metabolism is coordinated with distinct developmental phases of secondary xylem. New Phytol. 222: 1832–1845. https://doi.org/10.1111/nph.15739

Qing Q., and C. E. Wyman, 2011 Supplementation with xylanase and β-xylosidase to reduce xylo-oligomer and xylan inhibition of enzymatic hydrolysis of cellulose and pretreated corn stover. Biotechnol. Biofuels 4: 1–12. https://doi.org/10.1186/1754-6834-4-18

Ralph J., G. Brunow, P. J. Harris, R. A. Dixon, P. F. Schatz, *et al.*, 2008 Lignification: are lignins biosynthesized via simple combinatorial chemistry or via proteinaceous control and template replication?, pp. 36–66 in *Recent Advances in Polyphenol Research*, John Wiley & Sons, Ltd.

Rennie E. A., and H. V. Scheller, 2014 Xylan biosynthesis. Curr. Opin. Biotechnol. 26: 100–107. https://doi.org/10.1016/j.copbio.2013.11.013

Río J. C. Del, A. Gutiérrez, M. Hernando, P. Landín, J. Romero, *et al.*, 2005 Determining the influence of eucalypt lignin composition in paper pulp yield using Py-GC/MS. J. Anal. Appl. Pyrolysis 74: 110–115. https://doi.org/10.1016/j.jaap.2004.10.010

Shu K., Y. J. Meng, H. W. Shuai, W. G. Liu, J. B. Du, *et al.*, 2015 Dormancy and germination: How does the crop seed decide? Plant Biol. 17: 1104–1112. https://doi.org/10.1111/plb.12356

Shu K., X. D. Liu, Q. Xie, and Z. H. He, 2016 Two Faces of one seed: Hormonal regulation of dormancy and germination. Mol. Plant 9: 34–45. https://doi.org/10.1016/j.molp.2015.08.010

Silva E. A. A. Da, P. E. Toorop, J. Nijsse, J. D. Bewley, and H. W. M. Hilhorst, 2005 Exogenous gibberellins inhibit coffee (*Coffea arabica* cv. Rubi) seed germination and cause cell death in the embryo. J. Exp. Bot. 56: 1029–1038. https://doi.org/10.1093/jxb/eri096

Skepper J. N., and J. M. Powell, 2008 Ultrastructural immunochemistry. Cold Spring Harb. Protoc. 3: 1–7. https://doi.org/10.1101/pdb.top47

Smith P. J., H. T. Wang, W. S. York, M. J. Peña, and B. R. Urbanowicz, 2017 Designer biomass for next - generation biorefineries : leveraging recent insights into xylan structure and biosynthesis. Biotechnol. Biofuels 10: 1–14. https://doi.org/10.1186/s13068-017-0973-z

Smith P. J., M. A. O'Neill, J. Backe, W. S. York, M. J. Peña, *et al.*, 2020 Analytical techniques for determining the role of Domain of Unknown Function 579 Proteins in the Synthesis of O-methylated plant polysaccharides. SLAS Technol. https://doi.org/10.1177/2472630320912692\

Studer M. H., J. D. DeMartini, M. F. Davis, R. W. Sykes, B. Davison, *et al.*, 2011 Lignin content in natural *Populus* variants affects sugar release. Proc. Natl. Acad. Sci. 108: 6300–6305. https://doi.org/10.1073/pnas.1009252108

Sumiyoshi M., A. Nakamura, H. Nakamura, M. Hakata, H. Ichikawa, *et al.*, 2013 Increase in cellulose accumulation and improvement of saccharification by overexpression of arabinofuranosidase in rice. PLoS One 8. https://doi.org/10.1371/journal.pone.0078269

Sun Y., X. Y. Fan, D. M. Cao, W. Tang, K. He, *et al.*, 2010 Integration of brassinosteroid signal transduction with the transcription network for plant growth regulation in *Arabidopsis*. Dev. Cell 19: 765–777. https://doi.org/10.1016/j.devcel.2010.10.010

Suzuki M., A. Kato, N. Nagata, and Y. Komeda, 2002 A xylanase, *AtXyn1*, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the *Arabidopsis thaliana* genome. Plant Cell Physiol. 43: 759–767. https://doi.org/10.1093/pcp/pcf088

Szklarczyk D., A. L. Gable, D. Lyon, A. Junge, S. Wyder, *et al.*, 2019 STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47: D607–D613. https://doi.org/10.1093/nar/gky1131

Urbanowicz B. R., M. J. Peña, S. Ratnaparkhe, U. Avci, J. Backe, *et al.*, 2012 4-O-methylation of glucuronic acid in *Arabidopsis* glucuronoxylan is catalyzed by a domain of unknown function family 579 protein. Proc. Natl. Acad. Sci. 109: 14253–14258. https://doi.org/10.1073/pnas.1208097109

Urbanowicz B. R., M. J. Peña, H. A. Moniz, K. W. Moremen, and W. S. York, 2014 Two *Arabidopsis* proteins synthesize acetylated xylan in vitro. Plant J. 80: 197–206. https://doi.org/10.1111/tpj.12643

Vanholme R., V. Storme, B. Vanholme, L. Sundin, J. H. Christensen, *et al.*, 2012 A systems biology view of responses to lignin biosynthesis perturbations in *Arabidopsis*. Plant Cell 24: 3506–3529. https://doi.org/10.1105/tpc.112.102574

Wang C. S., S. W. Hsu, and Y. F. Hsu, 2013 New insights into desiccation-associated gene regulation by lilium longiflorum ASR during pollen maturation and in transgenic *Arabidopsis*. Elsevier. https://doi.org/10.1016/b978-0-12-407704-1.00002-6

Wierzbicki M. P., N. Christie, D. Pinard, S. D. Mansfield, E. Mizrachi, *et al.*, 2019a A systems genetics analysis in *Eucalyptus* reveals coordination of metabolic pathways associated with xylan modification in wood-forming tissues. New Phytol. 223: 1952–1972. https://doi.org/10.1111/nph.15972

Wierzbicki M. P., V. Maloney, E. Mizrachi, and A. A. Myburg, 2019b Xylan in the middle: Understanding xylan biosynthesis and its metabolic dependencies toward improving wood fiber for industrial processing. Front. Plant Sci. 10: 1–29. https://doi.org/10.3389/fpls.2019.00176

Xing S., G. Li, X. Sun, S. Ma, G. Chen, *et al.*, 2013 Dynamic changes in xylanases and β-1,4-endoglucanases secreted by *Aspergillus niger* An-76 in response to hydrolysates of lignocellulose polysaccharide. Appl. Biochem. Biotechnol. 171: 832–846. https://doi.org/10.1007/s12010-013-0402-0

Xiong Z. T., and H. Wang, 2005 Copper toxicity and bioaccumulation in Chinese cabbage (Brassica pekinensis Rupr.). Environ. Toxicol. 20: 188–194. https://doi.org/10.1002/tox.20094

Yang C. Y., C. H. Wu, G. Y. Jauh, J. C. Huang, C. C. Lin, *et al.*, 2008 The LLA23 protein translocates into nuclei shortly before desiccation in developing pollen grains and regulates gene expression in Arabidopsis. Protoplasma 233: 241–254. https://doi.org/10.1007/s00709-008-0016-5

Yang B., and C. E. Wyman, 2004 Effect of Xylan and Lignin removal by batch and flowthrough pretreatment on the enzymatic digestibility of corn stover cellulose. Biotechnol. Bioeng. 86: 88–98. https://doi.org/10.1002/bit.20043

Ye N., H. Li, G. Zhu, Y. Liu, R. Liu, *et al.*, 2014 Copper suppresses abscisic acid catabolism and catalase activity, and inhibits seed germination of rice. Plant Cell Physiol. 55: 2008–2016. https://doi.org/10.1093/pcp/pcu136

Yruela I., 2005 Copper in plants. Brazilian J. Plant Physiol. 17: 145–156. https://doi.org/10.1590/s1677-04202005000100012

Zeng Y., M. E. Himmel, and S. Y. Ding, 2017 Visualizing chemical functionality in plant cell walls. Biotechnol. Biofuels 10: 1–16. https://doi.org/10.1186/s13068-017-0953-3

Zhong R., Q. Teng, M. Haghighat, Y. Yuan, S. T. Furey, *et al.*, 2017 Cytosol-localized UDP-xylose synthases provide the major source of UDP-xylose for the biosynthesis of Xylan and xyloglucan. Plant Cell Physiol. 58: 156–174. https://doi.org/10.1093/pcp/pcw179

2.7 Tables and figures

Table 2.1 Candidate gene homologs selected for functional testing in *Arabidopsis thaliana* in the current study.

***Eucalyptus* Gene ID[a]**	***Arabidopsis* Gene ID[b]**	**Name and Description[c]**	**T-DNA insertion line[d]**
Eucgr.G02921	AT2G28760	*UXS6*, UDP-XYL synthase 6	SALK_058602C
Eucgr.F00109	AT1G58370	*XYN1*, Xylanase 1	SALK_099151C
Eucgr.K02293	AT2G01970	*EMP70*, Endomembrane protein 70 protein family	SALK_082204C
Eucgr.A01823	AT2G20680	*MAN2*, Endo-β-Mannase 2	SALK_126628C
Eucgr.A02178	AT3G10740	*ARAF1*, alpha-L-arabinofuranosidase 1	SALK_039176*

[a]*E. grandis* gene ID for genes possibly associated with secondary cell wall xylan modifications, identified by systems genetics (Wierzbicki *et al.* 2019a)

[b]*Arabidopsis thaliana* ortholog gene ID retrieved from EucGenIE (https://eucgenie.org).

[c]Assigned gene name and description from TAIR.

[d]Confirmed homozygous T-DNA insertion lines with insertions in coding regions. These are the selected insertion lines amongst others. All lines and information thereof were obtained from ARBC

*SALK_039176 is not listed as confirmed homozygous line on ARBC, however, it was confirmed homozygous though PCR in this study.

Table 2.2 Sequences of allele-specific primers used for genotyping T-DNA insertional lines.

T-DNA line	Left Primer (LP)	Right Primer (RP)
SALK_082204C	AAGGATCGGGTGATTGAAATC	TGCTGTTTTTACCGGCAATAC
SALK_039176	TGGAAACGAGACAAGAATTGG	CAATACCGTGATGATGTCGTG
SALK_058602C	GGCTTAGTTAACAAATTTTGTGG	TGCTCCAACTCGCTTAGCTAG
SALK_099151C	TCAAGCAACTGTTCAGCAATG	TGGTTCTACGGGGTAATGATG
SALK_126628C	TATGCCCATATGAGAGGCAAG	ATTTTCATATGCCAGGTGCTG

All sequences are given in 5'-3', LP+RP pair generates a PCR product of ~1000bp representing a wild-type allele. Pairing RP with T-DNA border primer Lb generates a ~600bp amplicon which represent the T-DNA inserted.

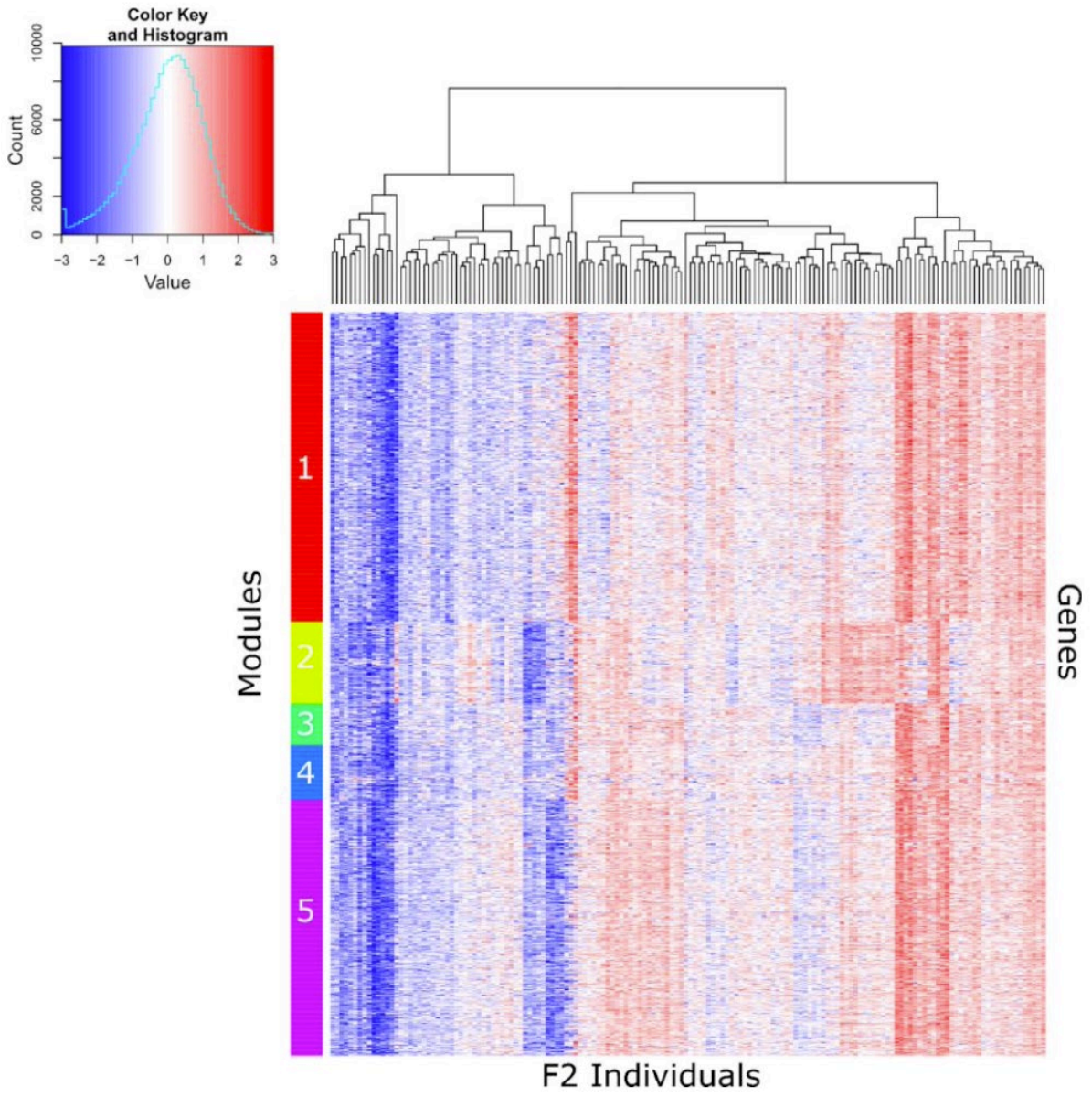


Figure 2.1 Co-expression and clustering of 1136 genes grouped into five nonredundant expression modules.

Correlation cut-off of 0.6 and F-score of 3.8 were set as criteria for this co-expression and clustering analysis. The columns show expression profile from of 156 *Eucalyptus urophylla* x *E. grandis* x *E. urophylla* tree population. These trees were clustered according to the genes' expression profiles in each tree, shown as rows. Furthermore, genes were grouped into expression modules (EMs) based on the similarity of their expression profiles. Adapted from Wierzbicki *et al.* 2019a.

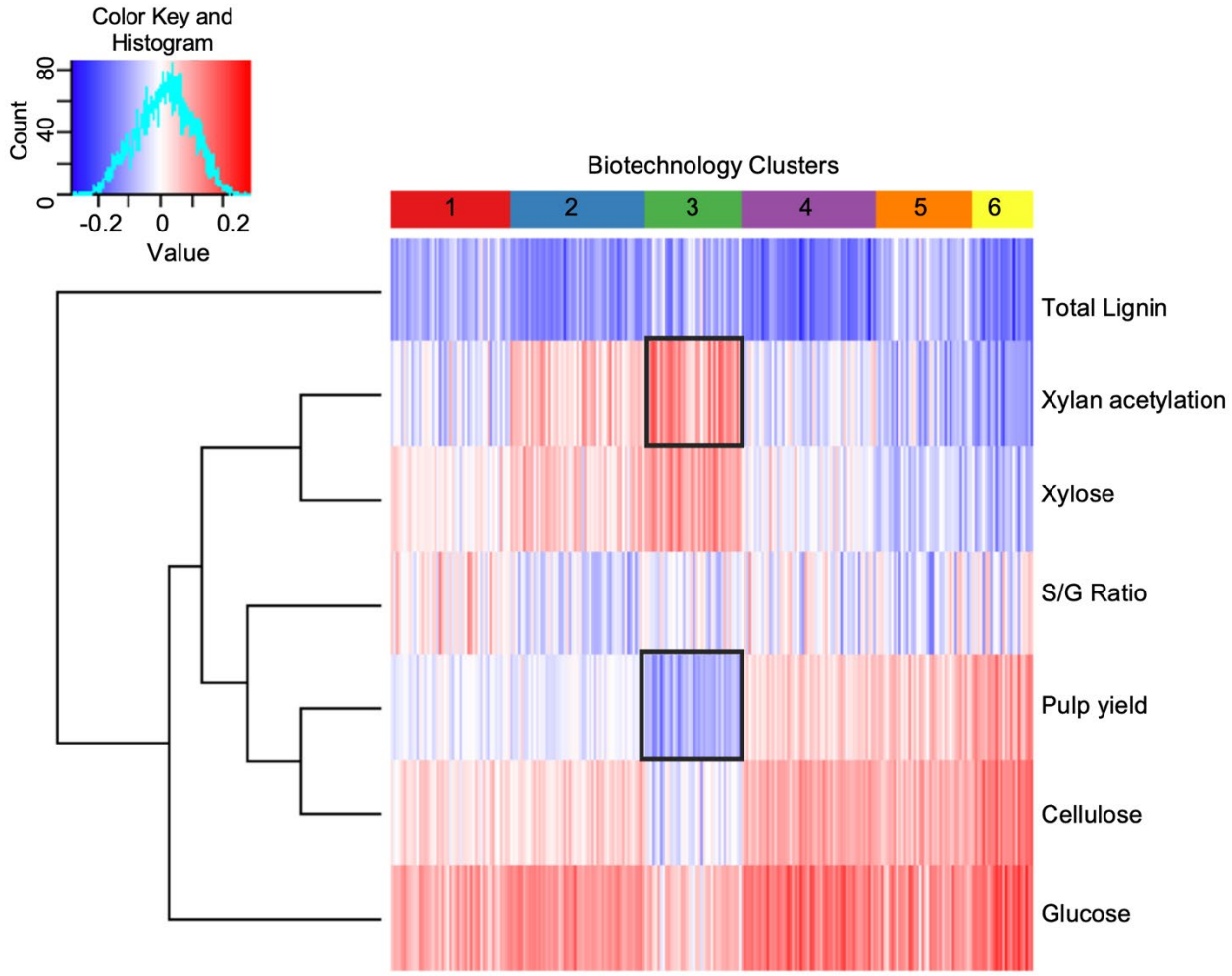


Figure 2.2 Correlation heat map showing 539 genes whose expression is either positively or negatively correlated with bioprocessing traits.

Pearson correlation analyses was performed on the 1136 genes shown in the EMs in the previous figure. Of the 1136 genes, 539 were clustered into six biotechnology clusters based on the similarities of their correlation patterns; this heat map shows these 539 genes. Red, white, and blue colors depict the R-value of correlation analyses, with red being positive values, white being 0 and blue being negative R values. Biotechnology cluster 3 was chosen for the selection of candidate genes to be functionally tested in *Arabidopsis*, selection of this cluster was explicitly based on the correlation pattern to xylan acetylation and pulp yield (Black borders). Figure modified from Wierzbicki *et al.* 2019a.

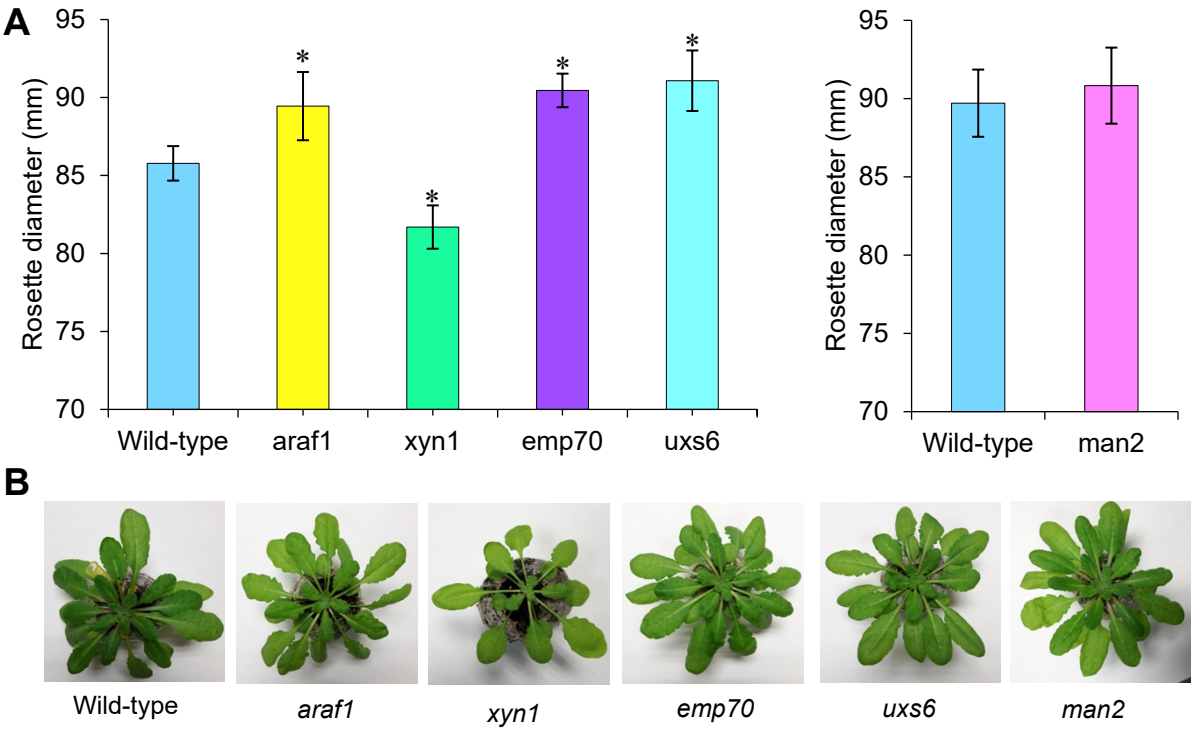


Figure 2.3 Average rosette leaf diameters of mutants measured 4 weeks post-planting.

(**A**) Rosette diameters were measured from the widest rosettes in each plant, *xyn1* mutant line had notably smaller rosettes, while *emp70* and *uxs6* had larger rosettes. Average rosette diameter ± SE, n_{WT} = 24, n_{mutant} = 35 (**t*-test, P <0.05). (**B**) Representative rosette leaves of mutants and wild-type at 4 weeks post-planting.

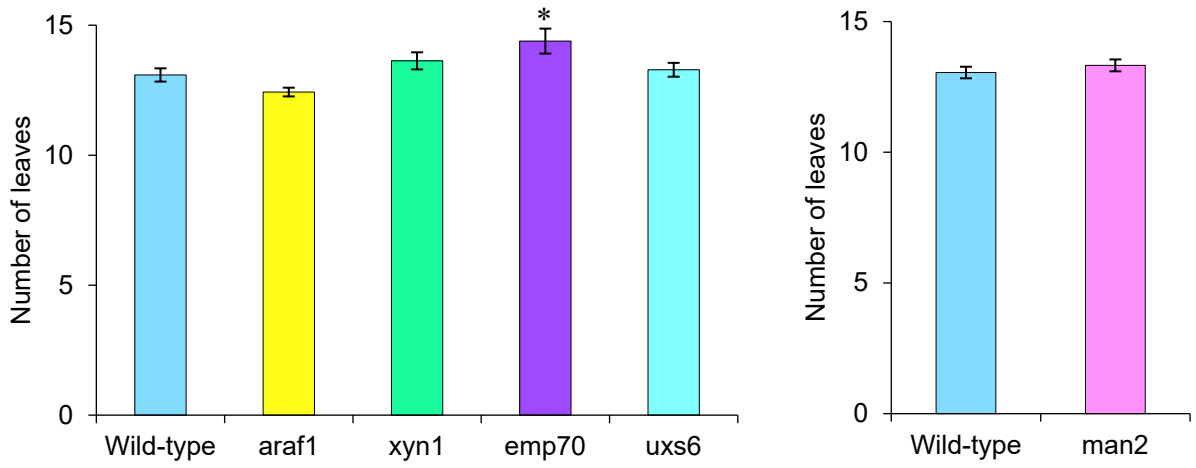


Figure 2.4 Average number of rosette leaves at the time of bolting.

The number of rosette leaves greater than 1 mm in length were counted immediately following inflorescence emergence. Figures show the average number of rosette leaves ± SE, n_{WT} = 24, n_{mutant} = 35. *Indicates significant difference in leaf numbers of *emp70* mutant compared with wild-type (**t*-test, P <0.05).

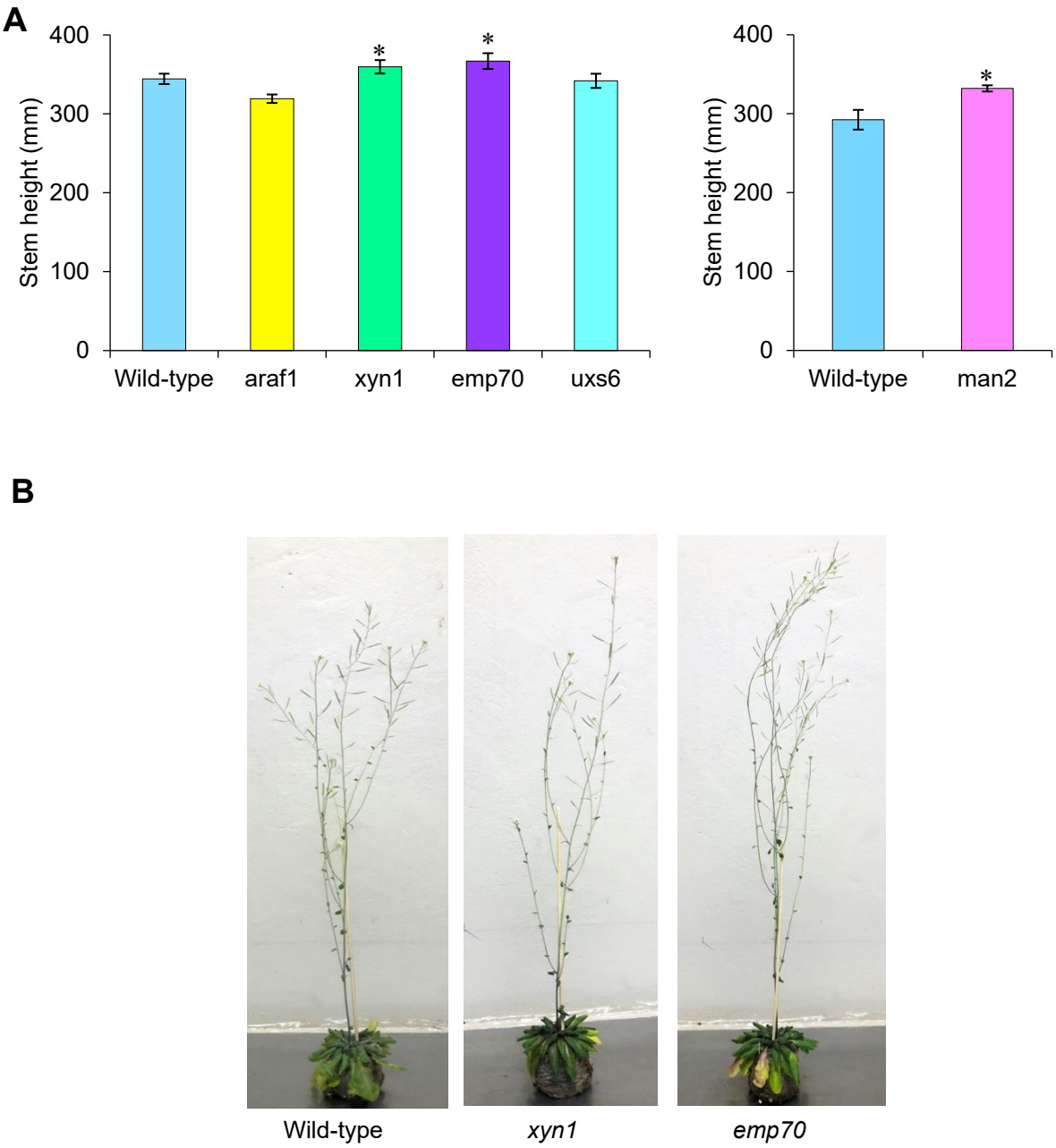


Figure 2.5 Maximum inflorescence stem height measured 7 weeks post planting.

(A) Inflorescence height was measured weekly until there was no apparent significant difference in height. Average stem height ± SE, $n_{WT} = 20$, $n_{mutant} = 30$ (**t*-test, $P < 0.05$). **(B)** Representative inflorescence stems for *xyn1* and *emp70* mutant lines which consistently had significantly longer stems. The taller stems phenotype of *man2* is not presented as it was not replicated in a separate trial.

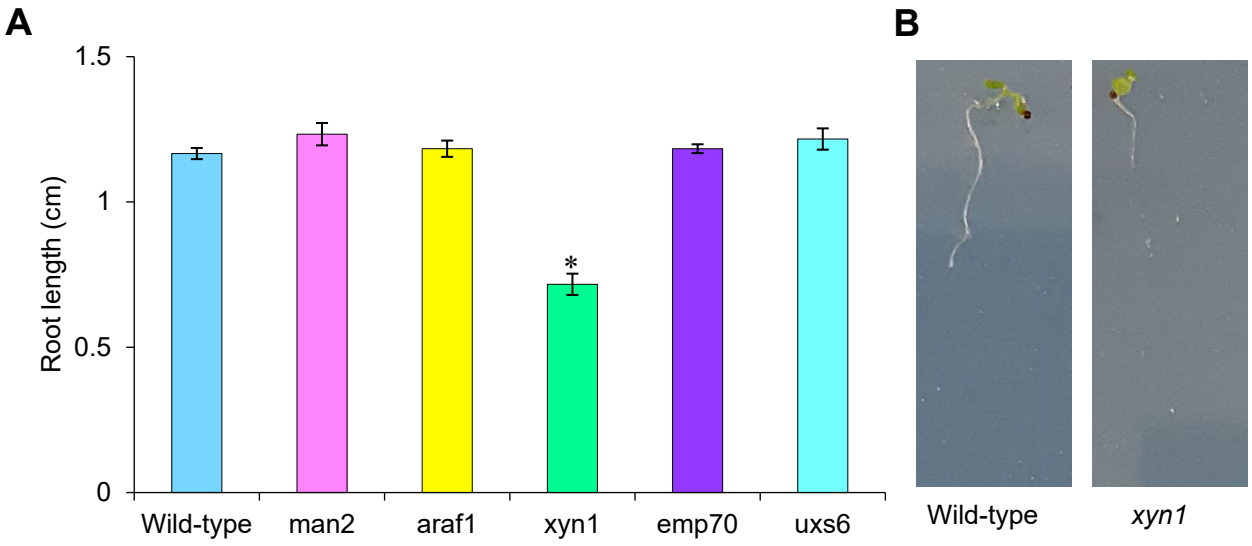


Figure 2.6 Root lengths measured 10 days post germination on ½ MS media.

(A) Mutants germinated in ½ MS and root lengths measured 10 days post-germination. Average root length ± SE, from 2 independent trials with n=6 each. *Indicates mutant line whose roots are significantly shorter than wild-type's (**t*-test, P <0.05). **(B)** Representative 10 days-old seedlings showing the obvious difference in root length between *xyn1* mutant and wild-type. Refer to **Figure S2.4** for the presentation of root lengths of all mutant lines.

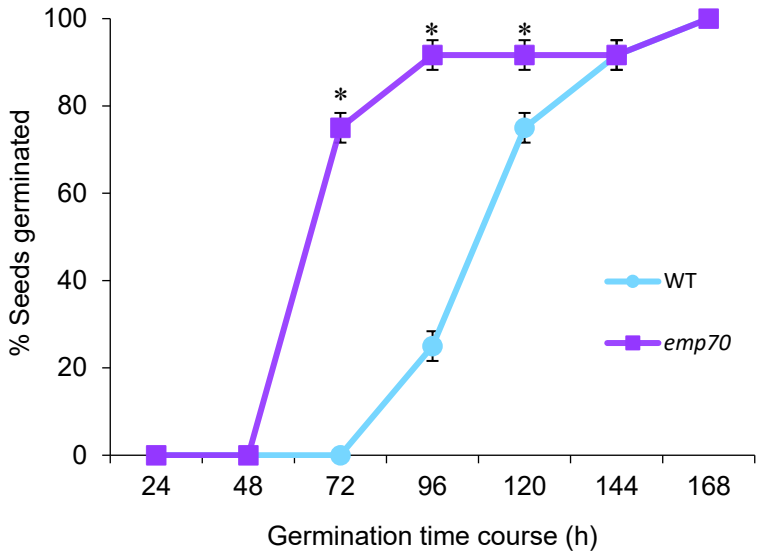


Figure 2.7 Analyses of germination time course of *emp70* knock-out mutant seeds.

Seeds were sown on soil and the number of seeds germinated counted every day until all seeds have germinated. Data shows an average number of seeds germinated ± SE, n = 6 (t-test, P <0.05) from two independent experiments. *emp70* mutant line germinate earlier than wild-type.

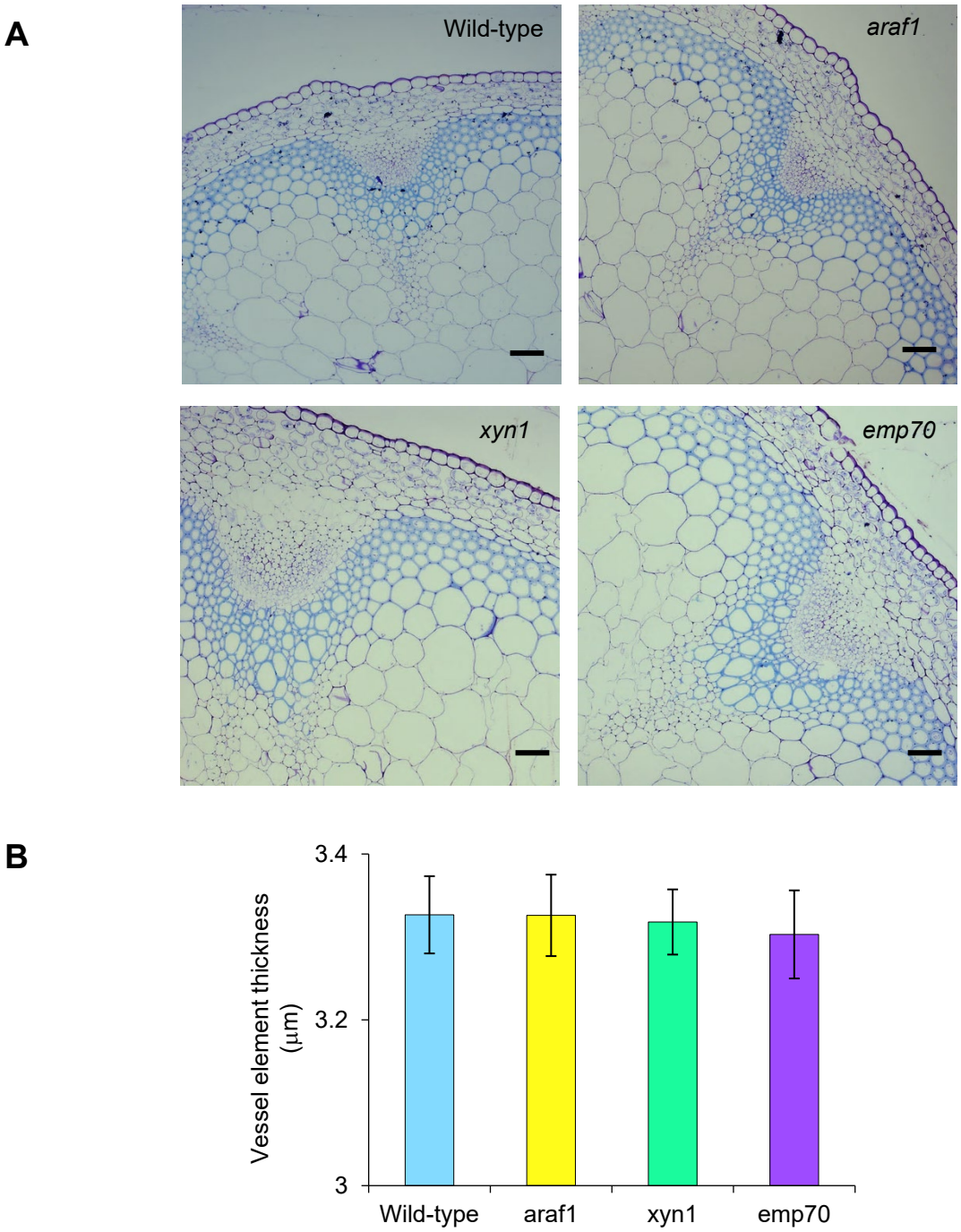
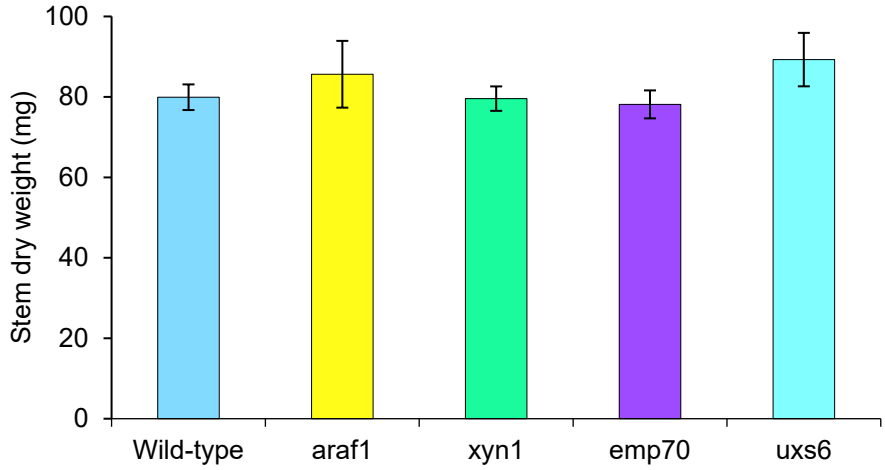


Figure 2.8 Analysis of xylem vessel morphology of mutants through microscopy.

(A) Transverse sections of 8-week-old bottom internodes of wild-type, *araf1*, *xyn1* and *emp70* mutants, respectively. Two biological replicates and 15 technical replicates per line were analyzed. Images visualized through light microscope at 20X. No apparent defects in secondary cell wall morphology were observed. Scale bars represent 50 µm. **(B)** Vessel element thickness measured at 40X showing no difference in thickness between cell walls from wild-type and mutant. Average vessel element thickness mm ± SE**,** n=10**,** (*t*-test, P <0.05).

A



B

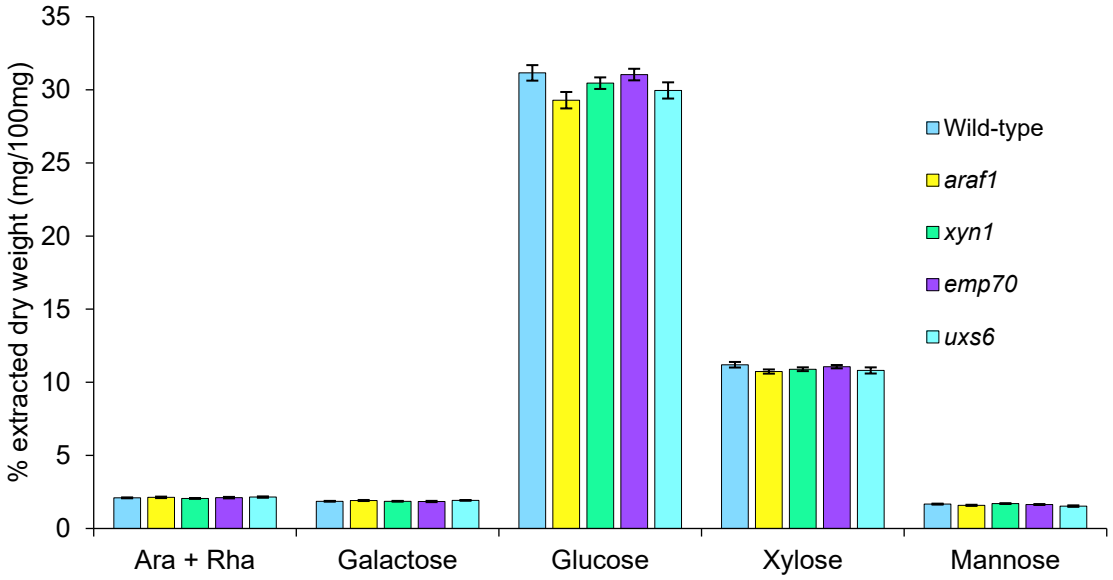
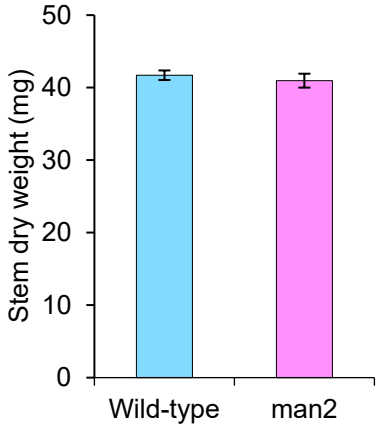


Figure 2.9 Cell wall monosaccharide composition of *xyn1*, *emp70*, *uxs6*, *araf1* and wild-type.

(A) Stems from nine-week-old plants were cut, dried, and weighed on a balance scale. Average stem weight ± SE, n_{WT} = 16, n_{mutant} = 6 (t-test, P <0.05). **(B)** Cell wall material was prepared from the dried stem and wood composition determined through the Klason method. Sugar content is given as the percentage of extracted dry weight (mg/100mg) ± SE, n_{WT} = 16, n_{mutant} = 6 (t-test, P <0.05).

A



B

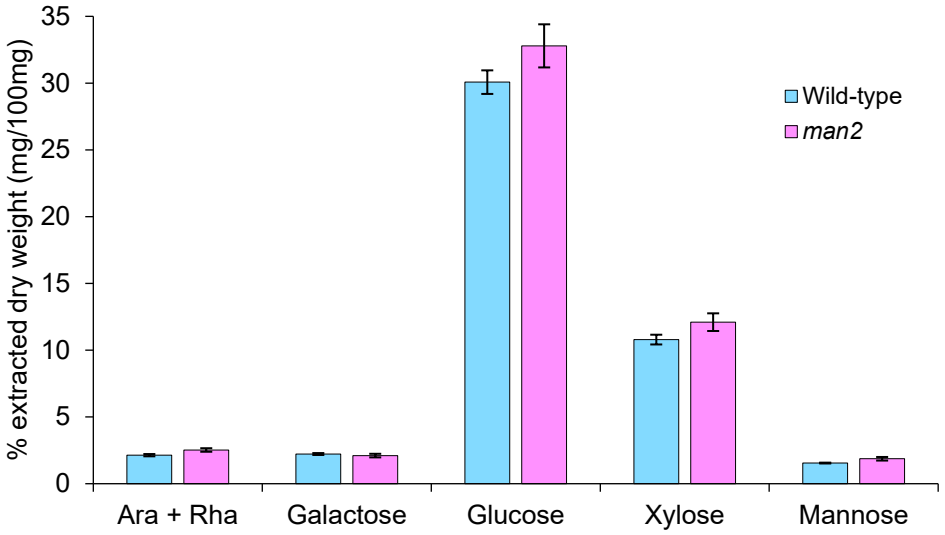


Figure 2.10 Cell wall monosaccharide composition of *man2* and wild-type.

(A) Stems from nine-week-old plants were cut, dried, and weighed on a balance scale. Average stem weight ± SE, n = 6 (t-test, P <0.05). **(B)** Cell wall material was prepared from the dried stem and wood composition determined through the Klason method. Sugar content is given as the percentage of dry extracted weight (mg/100mg) ± SE, n_{WT} = 16, n_{mutant} = 6 (t-test, P <0.05).

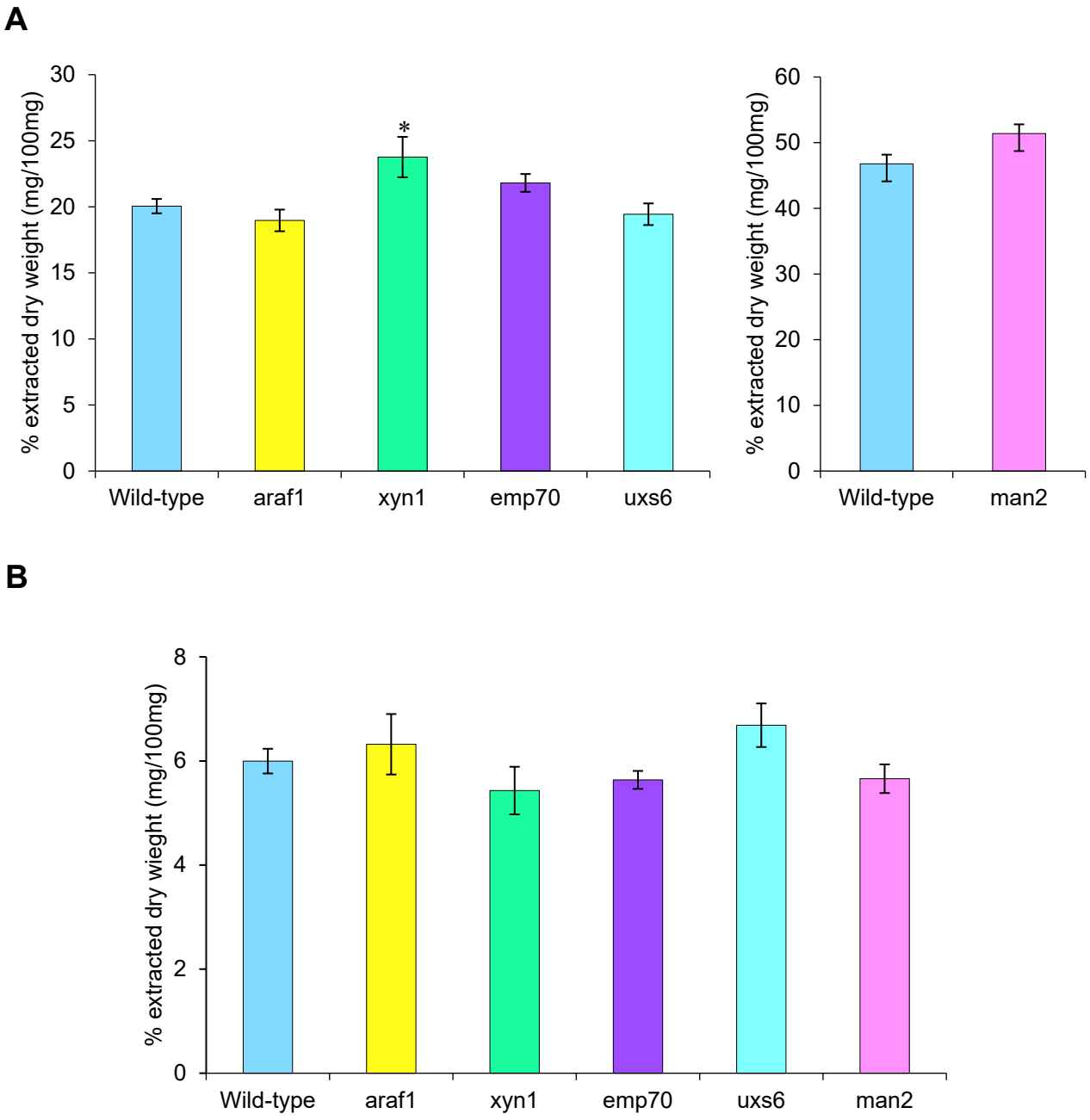


Figure 2.11 Quantification of cell wall lignin and acetyl content.

(A) Dried 9-week-old stems were ground, and lignin recovered through the Klason method; lignin content presented as % extracted dry weight ± SE, $n_{WT} = 16$, $n_{mutant} = 6$ (**t*-test, $P < 0.05$). (**B**) Acetyl content presented as % extracted dry weight ± SE, $n_{WT} = 10$, $n_{mutant} = 4$ (*t*-test, $P < 0.05$). No apparent differences were observed between mutant lines and wild-type's acetic acid content.

2.8 Supplementary Tables and Figures

Table S2.1 Candidate *Arabidopsis* genes implicated in secondary cell wall biosynthesis and/or modification based on literature.

Eucalyptus Gene ID[a]	*Arabidopsis* Gene ID[b]	Name and Description[c]	Literature[d]
Eucgr.G02921	AT2G28760	*UXS6*, *UDP-XYL SYNTHASE 6*	Zhong *et al.* 2017
Eucgr.F00109	AT1G58370	*XYN1*, *XYLANASE 1*	Endo *et al.* 2019
Eucgr.K02293	AT2G01970	*EMP70*, *ENDOMEMBRANE PROTEIN 70 PROTEIN FAMILY*	(Nikolovski *et al.* 2014)
Eucgr.A01823	AT2G20680	*MAN2*, *ENDO-BETA-MANNASE 2*	Ichinose *et al.* 2010
Eucgr.A02178	AT3G10740	*ARAF1*, *ALPHA-L-ARABINOFURANOSIDASE 1*	Kubo *et al.* 2005
Eucgr.K00317	AT5G53340	*HPGT1*, *HYDROXYPROLINE O-GALACTOSYLTRANSFERASE 1*	Ogawa-Ohnishi and Matsubayashi 2015
Eucgr.J00367	AT2G28100	*FUC1*, *ALPHA-L-FUCOSIDASE 1*	Kato *et al.* 2018
Eucgr.H02683	AT3G23820	*GAE6*, *UDP-D-GLUCURONATE 4-EPIMERASE 6*	Bethke *et al.* 2015
Eucgr.A01558	AT5G03760	*CSLA09*, *CELLULOSE SYNTHASE-LIKE A9*	Kim *et al.* 2014
Eucgr.D00298	AT1G30000	*MNS3*, *ALPHA MANNOSIDASE 3*	Liebminger *et al.* 2009
Eucgr.E00471	AT4G00110	*GAE3*, *UDP-D-GLUCURONATE 4-EPIMERASE 3*	Mølhøj *et al.* 2004
Eucgr.H03918	AT3G16630	*KINESIN-13A*, *KINESIN-13A*	Fujikura *et al.* 2014
Eucgr.G00653	AT5G64570	*BXL4*, *BETA-D-XYLOSIDASE 4*	Breitenbach *et al.* 2014

[a]*E. grandis* gene IDs for putative secondary cell wall modification/biosynthetic genes, identified by systems genetics (Wierzbicki *et al.* 2019a)

[b]*Arabidopsis thaliana* ortholog gene ID retrieved from EucGenIE (https://eucgenie.org).

[c]Assigned gene name and description from TAIR.

[d]Selected research studies with information indicating a possible direct role of candidate genes in secondary cell wall development and modification.

Candidate genes ultimately selected for functional testing in current study are shaded in blue.

Table S2.2 Gene ontology (GO) terms for selected genes. Cellular component, biological process and molecular function indicated. Information gathered from TAIR, Phytozome and Dicots PLAZA.

Gene	Cellular component	Biological process	Molecular function
UXS6	Plasma membrane	Nucleotide-sugar metabolism	Cellular metabolism, coenzyme binding
XYN1	Extracellular region, cell wall	Xylan catabolic process	Endo-1,4-β-xylanase activity
EMP70	Golgi membrane, endosome membrane, integral component of membrane, plant-type cell wall, vacuolar membrane, plasmodesma, endosome, trans-Golgi network	Transport	Unknown
MAN2	Extracellular region	Mannan catabolic process, mannan metabolic process	Hydrolase activity, hydrolysing o-glycosyl compounds, mannan endo-1,4-β-mannosidase activity
ARAF1	Apoplast, extracellular region, plant-type cell wall, vacuole	L-arabinose metabolic process, xylan catabolic process, maybe involved in cell wall modification	α-L-arabinofuranosidase activity, hydrolase activity, acting on glycosyl bonds, xylan 1,4-β-xylosidase activity

Table S2.3 Germination time and the percent seed germinated of wild-type *Arabidopsis* and T-DNA insertion mutant seeds.

Time (hours)	*man2*	*araf1*	*xyn1*	*uxs6*	*emp70*	WT
24	0	0	0	0	**0**	0
48	0	0	0	0	**0**	0
72	0	0	0	0	**75.0***	0
96	16.7	8.3	8.3	25	**91.7***	25
120	75	58.3	75	58.3	**91.7**	75
144	91.7	83.3	91.7	91.7	**100**	91.7
168	100	100	100	100	**100**	100

*Shows the significant difference (**t*-test, $P < 0.05$) in the average % seed germinated between *emp70* mutant line and wild-type seeds.

Table S2.4 Summary of traits observed in *Arabidopsis* mutant growth trial.

Trait	Significant increase ($P \leq 0.05$)	Significant decrease ($P \leq 0.05$)	Figure
Growth			
Germination % over time	***emp70***	None	**2.7**
Root lengths	None	***xyn1***	**2.6**
Rosette diameter	***emp70***	***xyn1***	**2.3**
Bolting time	None	None	**S2.2**
Rosette leaf number at bolting	***emp70***	None	**2.4**
Inflorescence stem height	***emp70; xyn1***	***araf1****	**2.5; S2.3**
Cell wall morphology and composition			
Vessel element thickness	None	None	**2.8**
Stem dry weight	None	None	**2.9; 2.10**
Monosaccharide content	None	None	**2.9; 2.10**
Lignin content	***xyn1***	None	**2.11**
Acetyl content	None	None	**2.11**

**araf1* stems caught up with wild-type at the end of measurement period.
All statistical analyses were performed with t-test, significant difference in the trait deduced from P-value (Paired T-test), where $P \leq 0.05$ represents a significant difference.

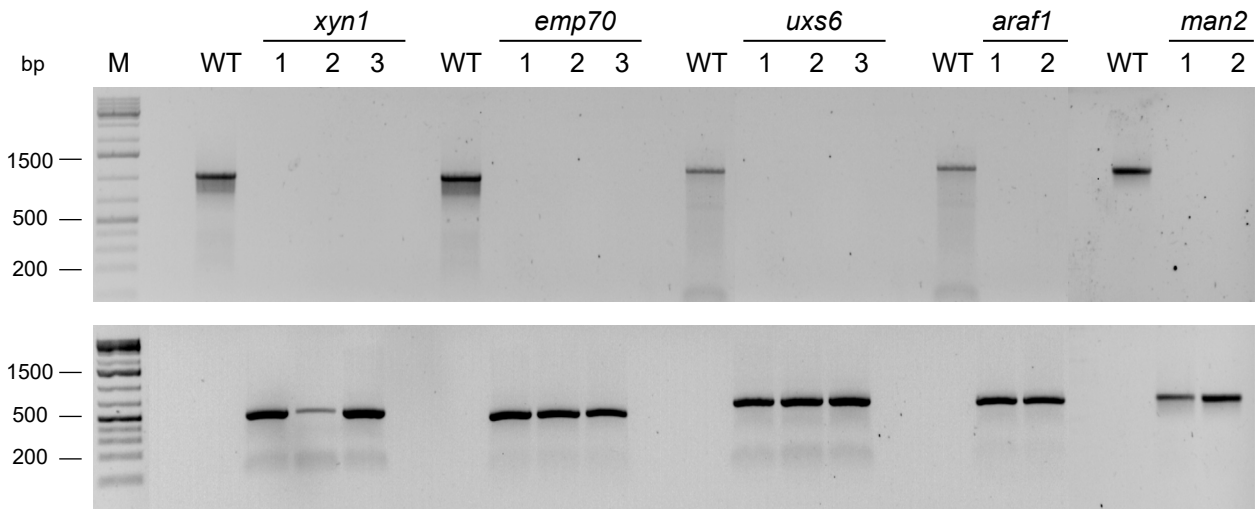


Figure S2.1 Confirmation of the homozygous genotype of T-DNA insertion mutants through PCR.

Top gel image shows the 1kb band generated with allele-specific PCR primers, LP and RP. Bottom gel image shows a 600 bp band generated with right allele-specific (RP) and T-DNA insertion border primer (LB). M is the Gene Ruler 1kb plus DNA ladder (Thermo Scientific). For all mutant lines but *araf1* and *man2*, DNA was extracted from three randomly chosen plants (1, 2, and 3). The absence of 1 kb band in mutant DNA samples confirms that all lines are homozygous for the T-DNA insertion.

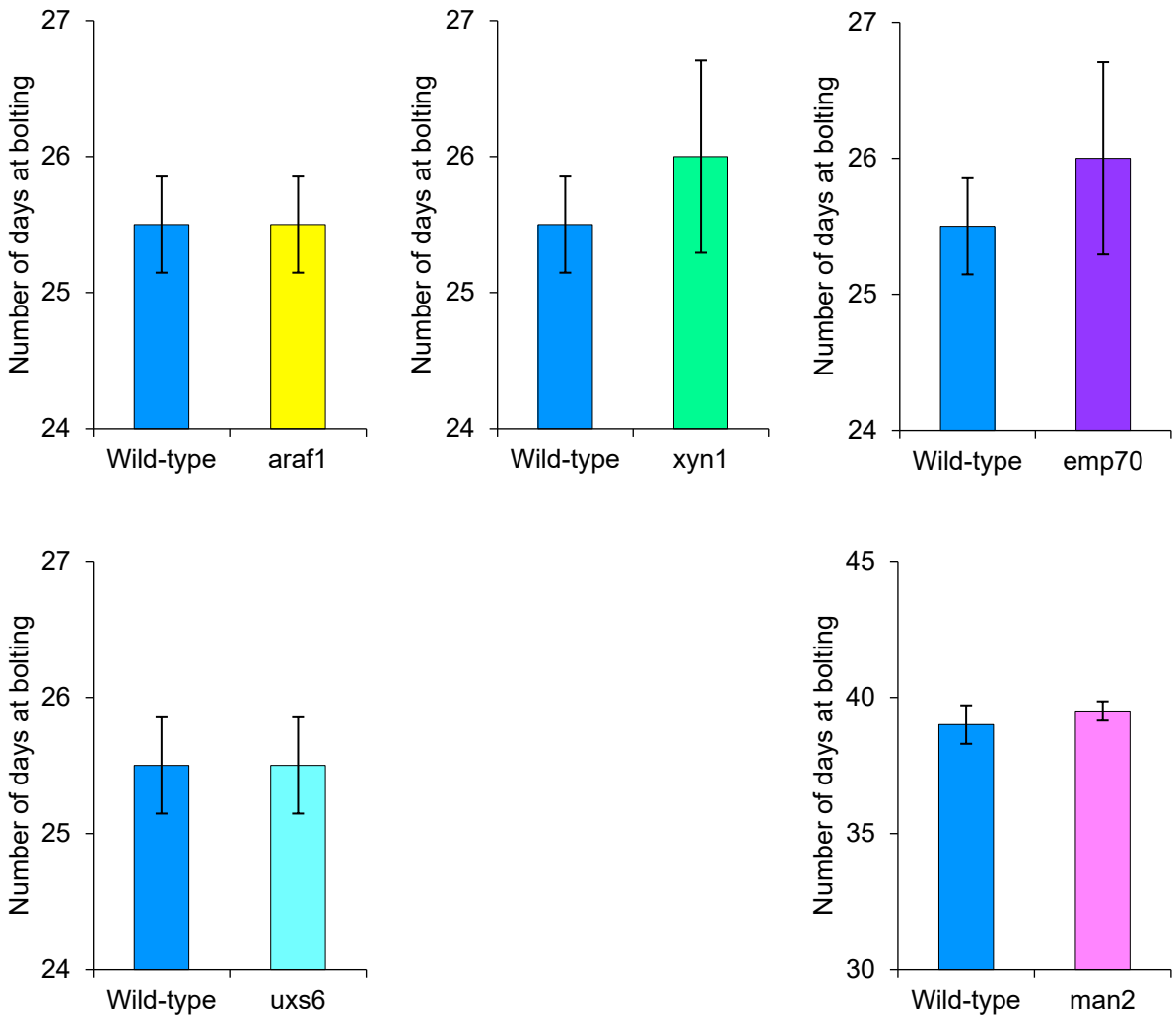


Figure S2.2 Bolting time of wild-type and mutant lines recorded at the day of bolting.

The day the inflorescence stem emerged was recorded, bolting time is presented as the number of days it took the plants to bolt post-planting. Note the difference in the chat scale for *man2* mutant as this was a separate trial. Data presented as an average number of days at bolting ± SE, observations made in two independent trials with 18 biological replicates per line in each trial (*t*-test, $P < 0.05$).

Wild-type ***araf1***

Figure S2.3 Maximum stem height of 5-week-old *araf1* mutants.

Maximum inflorescence height measured after all plants had bolted. *araf1* mutants had significantly shorter stems relative to wild-type, $n_{WT} = 20$, $n_{mutant} = 30$ (*t*-test, $P < 0.05$), however, plants caught up and stems were as tall as wild-type at seventh week (**Figure 2.5**).

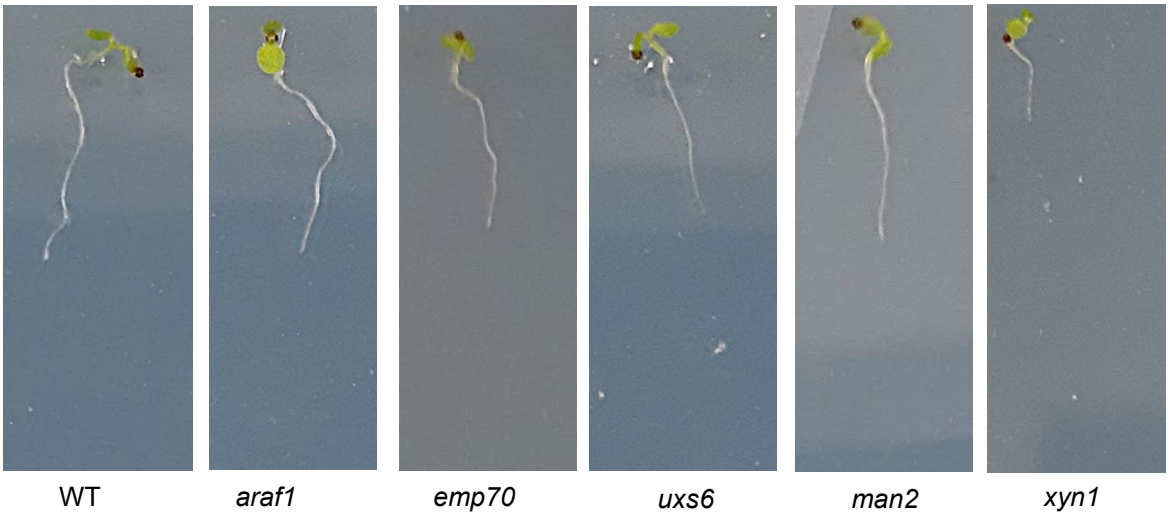


Figure S2.4 Root lengths measured 10 days post-germination.

Mutants germinated in ½ Murashige and Skoog media and root lengths measured 10 days post-germination, roots from 10-week-old wild-type and mutant lines seedlings are represented. Only *xyn1* mutant line roots were significantly shorter than that of wild-type seedlings, while all other mutant lines roots have identical root lengths to wild-type.

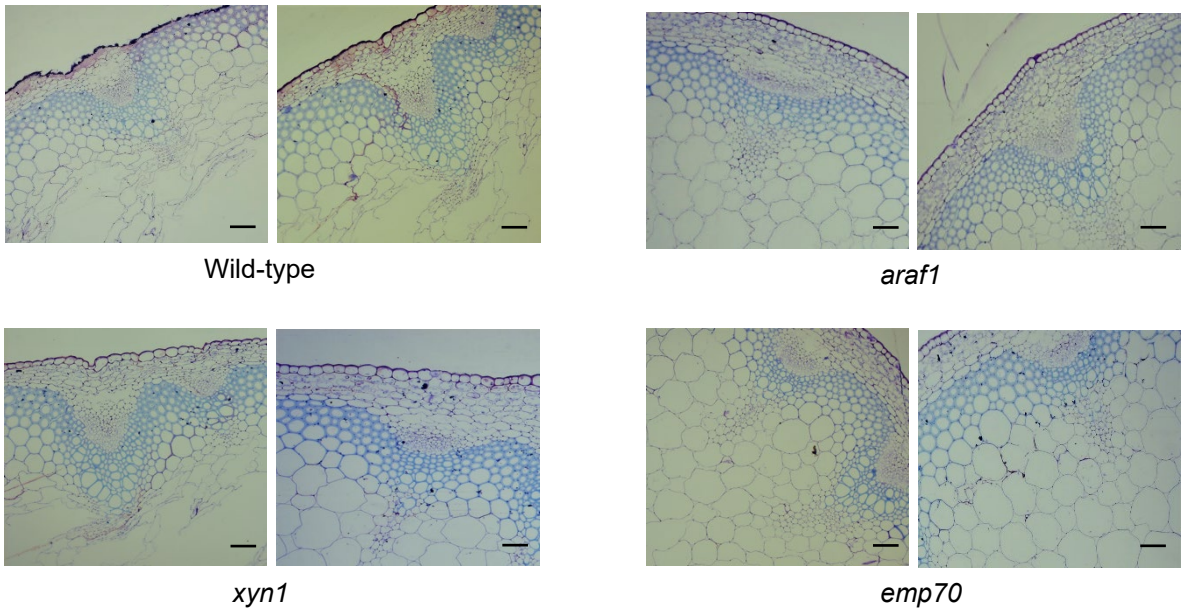


Figure S2.5 Micrographs showing xylem vessel morphology of mutants and wild-type.

Transverse sections of 8-week-old bottom internodes used for analyses. Two biological replicates and 15 technical replicates per line. No apparent defects in secondary cell wall morphology were observed. Images visualized through light microscope at 20X, Scale bars represent 50 μm.

Chapter 3

Application of CRISPR/Cas9 gene editing in *Populus*

Nganeavhutshilo Nangammbi[1], Alexander A. Myburg[1], Eshchar Mizrachi[1] and Victoria J. Maloney[1]

[1]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), Forest Molecular Genetics (FMG), University of Pretoria, Private bag X20, Pretoria, 0028

3.1 Summary

While the work in *Arabidopsis* in the previous chapter was useful for analyses of mutant growth phenotypes and elucidating the possible significance of selected genes, it is the fast-growing woody plants with high biomass yields such as poplar and *Eucalyptus* that are vital to produce bio-based products. Modifying these secondary cell wall-related genes in woody plant species is extremely necessary in understanding their roles in cell wall development and potentially yield desirable bioprocessing phenotypes. This modification requires an understanding of the biological roles of cell wall biosynthetic genes, and how modifying these genes could impact traits that affect bioprocessing. While the genes of interest in the current study were selected from a study originally performed in *Eucalyptus* as previously described, the poorly established transformation system for *Eucalyptus* in our lab necessitated the use of poplar for the functional characterization work in this chapter. Optimization of *Eucalyptus* transformation protocol are underway and future work will focus on functionally testing candidate genes in *Eucalyptus* as a more important source of lignocellulosic biomass feedstock.

In the previous chapter, gene expression profiles of genes associated with xylan modification were correlated with several biotechnology traits, and analyses of growth phenotypes, and cell wall chemistry were performed. Here, we used CRISPR/Cas9 gene editing tool to target *XYN1* and *MAN2* homologs in hybrid poplar (*Populus alba* x *P. tremula*; P717) to assess the effect of a mutation in these genes in woody plant growth. We first perform phylogenetic analyses to establish the evolutionary relationships amongst *Eucalyptus*, *Arabidopsis* and *Populus* homologs. Understanding the relationship between *Eucalyptus* and *Populus* homologs was essential for the future gene editing work to be performed in *Eucalyptus*. Phylogenetic analysis was followed by gRNA design, CRISPR/Cas9 construct assembly and several P717 transformation rounds. Low numbers of putative transformant shoots were recovered from leaf explants, from these, CRISPR/Cas9-*MAN2* and CRISPR/Cas9-*XYN1*

lines expressing Cas9 were obtained, however, the target locus was not modified in either of these lines. The data generated here presents some of the initial steps towards understanding the roles of the selected xylan-associated genes specifically in woody plants, which are more significant carbon sinks.

3.2 Introduction

In the previous chapter, we analyzed growth phenotypes of *Arabidopsis thaliana* T-DNA insertional mutants as the initial step towards elucidating the possible biological roles of the selected genes. Furthermore, the impact of the mutation on secondary cell wall structure was investigated through cell wall morphology and chemistry analyses. Therefore, the work in *Arabidopsis* was useful for analyses of possible biological roles of selected genes, and follow-up studies will be useful in establishing the exact function and direct impact of modifying these genes on bioconversion efficiency. However, the overall secondary cell wall development in *Arabidopsis*, which is herbaceous, differs from that in woody plants such as poplar (reviewed in Barros *et al.* 2015). Nevertheless, the molecular mechanisms and signaling hubs involved in secondary cell wall development are conserved between these species (Barra-Jiménez and Ragni 2017). In both species, secondary cell wall development originates from the vascular cambium, which produces xylem and phloem (Taylor 2002). In *Arabidopsis*, a substantial secondary cell wall development continues in the hypocotyl, forming secondary xylem closely resembling the wood anatomy of woody plants (Taylor 2008). Secondary cell wall development in *Arabidopsis* roots and stems is, however, very limited (Joshi *et al.* 2004). The overall extent of secondary cell wall development in *Arabidopsis* is therefore minimal relative to that of woody plants which form extensive secondary cell walls necessary for supporting the tree throughout its extended lifespan, and thus, presents a more significant carbon sink. These differences in the secondary cell wall between *Arabidopsis* and woody plants, therefore, warrants separate functional analyses studies specifically in woody plants. Moreover, woody plants represent the most abundant potential feedstock to produce a wide range of important bio-based products as mentioned previously, consequently, subsequent analyses of the impact on bioprocessing of gene modification specifically in woody plants are highly valuable.

Several genetic engineering studies have been focusing extensively on understanding the biological roles of secondary cell wall biosynthetic genes in woody plants and how modification of such genes impact lignocellulose biomass production. Amongst others, hybrid poplar has been one of the most studied woody plant model in most of these studies, this is attributed to its fast growth rate and increased vegetative propagation capacity (Park *et al.* 2004). Simultaneous downregulation of *RWA* genes, which are responsible for acetylation of xylan, leads 25% reduction in xylan acetylation in hybrid aspen (Pawar *et al.* 2017b), this can indeed possibly lead to a reduction of biomass recalcitrance. Furthermore, heterologous expression of yeast acetyl esterase, AnAXE, in hybrid aspen (*Populus tremula L.* x *tremuloides Michx*.) lead to transgenic plants with reduced xylan and a 25% increase in glucose yield (Pawar *et al.* 2017a). A similar study in *Populus tremula* x *P. tremuloides* has shown that by targeting *IRX9* and *IRX14* homologs, transgenic plants with reduced Xyl and increased lignocellulose bioprocessing efficiency can be generated (Ratke *et al.* 2018). Besides targeting xylan-biosynthetic genes, one of the most significant improvements in bioprocessing efficiency was achieved through downregulation of cinnamoyl-CoA reductase (CCR) necessary for lignin biosynthesis in *Populus tremula* x *P. alba*, in this study, a 161% increase in ethanol yield was achieved following saccharification (Acker *et al.* 2014). Thus, genetic modification of cell wall biosynthetic genes of woody plants represents the potential means of improving bioprocessing efficiency, to achieve such significant and sustainable improvements, precise and efficient genetic engineering tools are necessary.

Until the last decade, most of the gene editing-based functional characterization studies have mostly relied on transcription-like effector nucleases (TALENs) and the zinc finger nucleases (ZFNs). These approaches have indeed proven to be useful, however, they both suffer from limited gene specificity and gene silencing stability, which necessitates extensive screening and subsequent characterization of many transgenic lines for the desired trait (see detailed review by Gaj *et al.* 2013). Alternatively, in

the past decade, CRISPR/Cas9 genome editing technology has been exploited extensively for its advantages over TALENs and ZFNs. CRISPR offers simple targeted gene modification, high editing efficiency , and stable gene mutations (Jiang *et al.* 2013). Moreover, compared to other traditional methods such as RNA interference (RNAi), CRISPR/Cas9 induces targeted DNA double-stranded breaks (DSBs), thereby generating biallelic mutations and ensuring no residual gene expression (Jiang *et al.* 2013; Fan *et al.* 2015). Under the guidance of a 20-bp target-specific single guide RNA (sgRNA), a Cas9 endonuclease cleaves both DNA strands, generating double stranded-breaks (DSBs). The resulting DSBs can be repaired by the error-prone nonhomologous end-joining (NHEJ) or homology-directed repair (HDR) (Jiang *et al.* 2013). A repair through NHEJ, which is more predominant, often lead to random deletions and insertions (indels) and stop codons within the exon, thereby generating mutations through gene disruption (Jiang *et al.* 2013). Thus, the power of CRISPR/Cas9 can be harnessed to generate precise biallelic mutations of cell wall biosynthetic genes in woody plants, an important step in functional characterization and subsequent cell wall modification aimed at improvement of bioprocessing.

In this chapter, the goal was to investigate the effects of knocking out *Populus XYN1* and *MAN2* gene homologs on plant growth and physiology. This involved phylogenetic analyses which showed that both genes have no paralogs, enabling the design and use of only one gRNA to map the target site. Next, with reference to RNA-seq data from *P. tremula*, we show that both genes have increased expression specifically during secondary cell wall deposition. gRNAs were then designed to target these genes in P717, and plants containing Cas9 transgene were obtained although no mutations were detected in these transgenic plants.

3.3 Materials and Methods

3.3.1 Identifying candidate gene homologs in *Populus trichocarpa*

Populus genes orthologous to the *Eucalyptus* and *Arabidopsis* candidate genes were identified through phylogenetics. Briefly, *Eucalyptus* candidate gene IDs were used as a query in The *Populus* Genome Integrative Explorer, PopGenIE (http://popgenie.org/) to retrieve the amino acid sequences of the candidate genes in *Populus trichocarpa.* Additionally, the amino acid sequences of *Eucalyptus grandis* orthologs were obtained from EucGenIE (https://eucgenie.org). A phylogenetic tree representing all *A. thaliana, P. trichocarpa* and *E. grandis* candidate gene homologs was generated from the protein sequences with Geneious tree builder on Geneious Prime® version 2020.2.2 with the following parameters: Neighbour-joining tree build model and Jukes-Cantor genetic distance model. Based on the phylogenetic analyses, two *Populus* candidate gene orthologs; Potri.013G130400 (*MAN2*) and Potri.002G113100 (*XYN1*) were selected for genetic editing with CRISPR/Cas9 since they had no paralogs. This was essential since the lack of paralogs meant only one gRNA is required to map the target site, thereby avoiding the technical difficulties associated with CRISPR multiplexing, in which numerous gRNAs are designed and expressed for multi-locus editing
(Mojica *et al.* 2009).

3.3.2 *In silico* analyses of candidate gene expression during secondary cell wall development

Gene expression profiles of both candidate genes during secondary cell development was investigated using raw high-spatial-resolution RNA sequencing data published by Sundell *et al.* 2017. In the study, Sundell *et al.* 2017 assayed gene expression across the wood-forming tissues of a wild-growing aspen (*Populus tremula*) using RNA-seq. This involved cryosection to obtain a continuous sequence of

samples across differentiated phloem and mature xylem, followed by analyses of transcriptional modules composed of functionally characterized genes (Sundell *et al.* 2017). Expression profiles were generated for annotated genes and several potentially novel cell wall biosynthetic genes (Sundell *et al.* 2017). Taking advantage of these data, gene expression profiles were obtained for *MAN2* and *XYN1* and compared with the expression profile of known secondary cell wall biosynthetic gene Potri.002G257900 (*PtCESA4*).

3.3.3 Designing single guide RNAs targeting *Populus* homologs

Single guide RNAs (sgRNAs) were designed to target the coding regions of *XYN1* and *MAN2* in *Populus tremula* and *Populus alba* genome. This involved first identifying the coding regions of both genes, to do this, *P. tremula* and *P. alba* gene models and transcripts were obtained from the Aspen database (http://aspendb.uga.edu/) and aligned through Muscle alignment in Geneious Prime® 2020.2.2. Several 19 bp sgRNAs target regions were selected downstream of the SpCas9 PAM sequence (5'-NGG-3') within *P. tremula* and *P. alba* coding sequences. sgRNA was also designed to target the phenotypically characterized *Populus tomentosa* phytoene desaturase gene 8 (*PtoPDS*) P717 homolog (Potri.014G148700) as a control. All designed 19 bp sgRNAs, were appended with extra guanine (G) at the 5' end such that they were in the form 5'-GN_{19}-3' (Mali *et al.* 2013b; Ran *et al.* 2013b). The predicted CRISPR efficiency of each potential sgRNA was then evaluated through the online CRISPR/Cas9 efficiency prediction tool (https://fgr.hms.harvard.edu/crispr-efficiency). Furthermore, potential off-target activity associated with each sgRNA was predicted as follows; the probe search function on the Aspen database was used to investigate the position and identity of possible off-target sites in the P717 genome. Additionally, genes containing the off-target sequences were analyzed in Geneious Prime to determine if the off-targets were preceded by the SpCas9 PAM sequence.

3.3.4 Screening P717 sgRNA target sequence

To confirm that the sgRNA sequences match the target P717 sequence, as this was the plant material to be used in the study, primers flanking the sgRNA target regions were designed for PCR and subsequent sequencing of the target regions (**Table 3.1**). PCR was then performed with wild-type P717 DNA as a template, amplicons were subsequently recovered with Gene JET Gel Extraction Kit (Thermo Scientific). Purified amplicons were sequenced through Sanger sequencing at Macrogen Europe using designed PCR primers (**Table 3.1**).

3.3.5 Amplicon prep and CRISPR/Cas9 construct assembly

pUC gRNA shuttle (plasmid #47024) vector and p201N Cas9 (plasmid #59175) binary vector were isolated and purified from bacterial stabs as described on Addgene. *Mt* U6.6 (referred to simply as U6), a gRNA promoter, was PCR-amplified with SwaI_MtU6 and MtU6R primers using pUC gRNA shuttle vector as a PCR template (Jacobs *et al.* 2015). Similarly, the gRNA scaffold was amplified from the same vector with Scaffold F and SpeI_Scaffold R primers (**Table 3.1**). Amplicon sizes were verified through gel electrophoresis, followed by amplicon purification with Gene JET Gel Extraction Kit (Thermo Scientific).

To generate CRISPR/Cas9 constructs targeting candidate genes in *P. alba* x *P. tremula*, p201N Cas9 was linearized through digestion with SwaI (New England Biolabs ®) followed by clean-up, and subsequent digestion with SpeI-HF (New England Biolabs ®). The digested plasmid was analyzed on 1% agarose gel to verify that it was successfully linearized. U6, scaffold, and sgRNA were assembled into the linearized p201N Cas9 using the Gibson® Assembly Cloning Kit (NEB #E5510S). Completed Gibson assembly reactions were then used to transform competent NEB® 5-α competent *E. coli* cells

as per the Gibson assembly transformation protocol, followed by plating on LB agar plates supplemented with 50 mg/ml kanamycin. Putative positive transformants were screened through PCR to confirm whether they contained the inserts (i.e., U6, sgRNA, and scaffold). Briefly, eight overnight colonies were randomly selected and screened through colony PCR with plasmid specific primers: IsceI R and Ubi3p218R (**Table 3.1**). Agarose gel electrophoresis was used to visualize and confirm the size of amplicons, insert amplicons were then purified with Gene JET Gel Extraction Kit (Thermo Scientific) and Sanger sequenced with Ubi3p218R primer.

3.3.6 *Agrobacterium* transformation and generation of transgenic plants

Agrobacterium tumefaciens EHA105 was transformed with either CRISPR/Cas9-*XYN1*, CRISPR/Cas9-*MAN2* or CRISPR/Cas9-*PDS* construct through the modified freeze-thaw method (Jyothishwaran *et al.* 2007). This was followed by PCR to confirm the successful transform of the *Agrobacterium.* After confirmation, P717 leaves were then transformed through *Agrobacterium*-mediated transformation and leaf disc inoculation approach (Coleman *et al.* 2008). Briefly, leaf discs were infected with transformed EHA105 and transferred to woody plant media (WPM) (0.1 mM NAA: 1-Naphthaleneacetic acid, 0.1 mM BA: 6-Benzylaminopurine, and 0.1 mM TDZ: Thidiazuron) and kept in the dark for 3 days. Infected discs were then co-cultivated in WPM supplemented with 250mg/l Cefotaxime and 500 mg/l Carbenicillin (WPM CC), thereafter, leaf discs were transferred to WPM CC with 25 mg/l Kanamycin (WPM CCK) for shoot induction. Shoots developed after three weeks and transferred to root induction media (WPM CCK but only containing 0.01 mM BA hormone) and kept in this media for 3 weeks before final transfer to magenta boxes containing just WPM.

3.3.7 Genomic DNA extraction and mutation screening

Following plant growth in WPM magenta jars for 3 weeks, DNA was extracted from transgenic and wild-type P717 leaves using NucleoSpin^{TM} Plant Kit as in the previous chapter. The presence of Cas9 transgene was confirmed through PCR with Cas9-specific primers under these PCR conditions: 95°C for 3 minutes; 34 cycles at 95°C for 30 seconds, followed by primer annealing at 56°C for 30 seconds and 72°C for 90 seconds, followed by primer extension at 72°C for 3 minutes. Subsequent PCR was performed with gene-specific primers to detect possible mutations on the target site (**Table 3.1**). PCR amplicons were separated on 1% agarose gel and bands corresponding to the target region were recovered and sequenced as described previously. DNA sequence reads from transgenic and wild-type plants were imported and analyzed with Geneious Prime® 2020.2.2 to identify possible mutations at the target exon.

3.4 Results

3.4.1 *Populus trichocarpa* XYN1 and MAN2 do not have close paralogs

To understand the evolutionary history of the selected genes and identify *Populus* orthologs to be targeted through CRISPR/Cas9, a phylogenetic tree representing *Eucalyptus grandis*, *Arabidopsis thaliana*, and *Populus trichocarpa* protein homologs were constructed. It was established that *Populus* XYN1, MAN2, and UXS6 have no paralogs **(Figure 3.1**; **Figure S3.1**), which is consistent with Du *et al.* 2013. Notably, while these genes had no *Populus* paralogs, they each had two *Eucalyptus* homologs (**Figure 3.1)**. From these homologs, Eucgr.F00109 (*XYN1*), Eucgr.A02178 (*MAN2*) and Eucgr.G02921 (*UXS6*) correspond to the original genes identified with systems genetics (Wierzbicki *et al.* 2019a).

As stated earlier, the *Populus* genes (*XYN1*, *MAN2* and *UXS6*) were prioritized mainly to avoid CRISPR multiplexing since they had no paralogs. Furthermore, *AtXYN1* mutant line was the only line that showed impaired growth; with smaller rosettes and significantly shorter roots (**Figure 2.3**; **Figure 2.6**), these observations further motivated the decision to target this gene in *Populus*. Conversely, *Arabidopsis man2* mutants did not show any obvious phenotype, however, *MAN2* is interesting in that it is the only *MAN* expressed in germinating seeds whose mutation do not affect germination time course in *Arabidopsis* (Iglesias-Fernández *et al.* 2011a; b). Therefore, it was deemed necessary to investigate the phenotype of mutation of this gene homolog in poplar, a more significant carbon sink.

3.4.2 *XYN1* and *MAN2* show characteristic SCW deposition gene expression profile

Using RNA sequencing gene expression data from wild-growing *P. tremula* (Sundell *et al.* 2017), gene expression profiles across the developing phloem and secondary cell wall forming tissues were generated for *MAN2* (Potri.013G130400) and *XYN1* (Potri.002G113100) orthologs (**Figure 3.2**). Both genes showed a noticeable increase in expression at the onset of secondary cell wall deposition (**Figure 3.2**). High gene expression levels were maintained throughout secondary cell wall deposition, like the expression pattern of the functionally characterized secondary cell wall deposition gene, *CESA4* (**Figure 3.2**).

3.4.3 Guide RNA design

sgRNAs to target *XYN1*, *MAN2*, and *PDS* in P717 were designed in the form GN_{19} followed by the PAM sequence (**Figure 3.3**). Selected *XYN1*, *MAN2*, and *PDS* sgRNAs had predicted target modification efficiency scores of 7.710, 9.286, and 7.313 respectively, which suggested that all sgRNAs would be highly efficient in directing Cas9 and introducing mutations in the target site. To

assess the potential for off-target cleavage, off-target sequences with similar sequences to the sgRNA guiding sequence were searched in the P717 genome through the aspen dB online tool. Several potential off-target sites were identified, two potential off targets for *PDS*, one for *XYN1* and two for *MAN2* had at most three mismatches to the guiding sequence (**Table S3.1**; **Figure 3.4**). All potential off-target sites for *PDS* and *XYN1* had mismatch at the seed sequence, proximal to the PAM region, while off-target sites for *MAN2* sgRNA had mismatch at the 5' distal to PAM (**Figure 3.4**). To investigate the potential for Cas9 to bind and cleave at these putative off-target sites, PAM sequence following potential off-target sites was investigated, none of the potential off-target sites were followed by the canonical 5'-NGG-3' PAM sequence, thus, theoretically, Cas9 would not recognize and cleave in these regions (**Figure 3.4**).

3.4.4 CRISPR/Cas9 constructs

Before CRISPR/Cas9 constructs were to be generated, it was confirmed through target site amplification and sequencing that designed sgRNAs could target candidate genes in P717 (**Figure 3.5**). U6 promoter and scaffold amplicons were generated through PCR using pUC shuttle vector as a template (**Figure 3.6**). Then, both amplicons and sgRNA were assembled into linearized p201N Cas9 vector to generate separate constructs targeting *XYN1*, *MAN2*, and *PDS* in P717 (**Figure 3.7**), henceforth referred to as CRISPR/Cas9-*XYN1*, CRISPR/Cas9-*MAN2*, and CRISPR/Cas9-*PDS* respectively. Successful assembly of inserts into vector was verified through PCR, all eight screened *XYN1* and *MAN2* constructs contained inserts, while two *PDS* constructs potentially carried inserts (**Figure 3.8**). One confirmed positive transformant from each line was sequenced to further confirm the presence of inserts, and more importantly, check for single nucleotide polymorphisms (SNPs) in the inserts. All screened constructs contained inserts with no SNPs (**Figure S3.2**). Next, transformation

of *Agrobacterium* with the screened constructs was performed and, transformed *Agrobacterium* colonies were further screened to confirm they carried the constructs with the inserts (**Figure 3.9**).

3.4.5 Genetic modification detection and analyses of preliminary phenotype

By employing the previously described CRISPR/Cas9 gene editing protocol (Jacobs *et al.* 2015), CRISPR constructs targeting P717 *MAN2, XYN1,* and *PDS* were successfully generated (**Figure 3.8**). In three trials (each starting with ~80 leaf discs per line), a total of nine CRISPR/Cas9-*MAN2* and two CRISPR/Cas9-*XYN1* putative transformants were recovered. While CRISPR/Cas9-*MAN2* shoots were obtained from each trial, the CRISPR/Cas9-*XYN1* shoots were only obtained in one of the three trials. (**Figure 3.10**). Following five weeks in WPM+NAA/BA/TDZ/C containing 25 mg/l Kanamycin, CRISPR/Cas9-*XYN1* and CRISPR/Cas9-*MAN2* putative mutant shoots formed, however, only 3 and 2 shoots formed in each line, respectively (**Figure 3.10**). Notably, no CRISPR/Cas9-*PDS* shoots were recovered during the transformation trial. When the recovered shoots were grown in root-inducing media supplemented with kanamycin, an obvious root was formed in CRISPR/Cas9-*MAN2* line while CRISPR/Cas9-*XYN1* and wild-type plants did not develop such a root (**Figure 3.10**), suggesting that the CRISPR/Cas9-*MAN2* was more likely to be transgenic. Nevertheless, genomic DNA was extracted from both T0 putative transgenic plants (one CRISPR/Cas9-*MAN2* and one CRISPR/Cas9-*XYN1*) followed by Cas9 and subsequent mutation screening. Both screened plants were transgenic as they expressed the Cas9 transgene, however, none of them were edited at the target locus (**Figure 3.11**).

3.5 Discussion

To address the limitation associated with using *Arabidopsis* as a model plant for functional characterization, we aimed to edit *MAN2* and *XYN1* gene homologs in *Populus* using CRISPR/Cas9-gRNA approach. Transgenic *Populus* plants were generated, however, none of the transgenics contained a mutation at the target locus. Failure to recover *Populus* mutants may imply that these genes are essential in *Populus*, and as a result, mutation thereof may be lethal, which would not be surprising considering that *Populus* is a more significant carbon sink.

3.5.1 *Populus XYN1* and *MAN2* gene expression increases during secondary cell wall deposition

In this section, the goal was to generate CRISPR/Cas9 knock-outs of *XYN1* and *MAN2* in P717 with a long-term goal of establishing the impact of mutation of these genes on cell wall development and plant growth and physiology. *In silico* gene expression analyses of these genes showed a rapid increase in *XYN1* and *MAN2* expression during cell wall deposition (**Figure 3.2**), this indicates that both these genes may be important in cell wall biosynthesis in *Populus*. Supporting the role of *XYN1* in cell wall deposition, *XYN1* has been shown to be co-expressed with the Vascular-related NAC-domain 7 (*VND7*), a master regulator of secondary cell wall deposition (Turco *et al.* 2017). The low gene expression maintained throughout all transcriptome reprogramming events before cell wall deposition may suggest that the functions of these genes are restricted to secondary cell wall deposition (**Figure 3.2**). Notably, the expression of *MAN2* also spikes during early cell death which presents a possibility that in addition to cell wall deposition, *MAN2* may play a role in programmed cell death (**Figure 3.2**).

3.5.2 Modification of candidate gene homologs

gRNAs were designed to target the sense strand of *XYN1* and *MAN2* genes. To be able to distinguish between *P. alba* and *P. tremula* alleles, the target regions for designed gRNA were selected with the consideration that they should be surrounded by some SNPs between *alba* and *tremula* alleles. Several sgRNAs were designed for each gene and subjected to an online CRISPR efficiency prediction tool which gives the cumulative p-value efficiency score for the gRNA sequence (Housden *et al.* 2015). This tool assesses the target modification efficiency of sgRNA based on the presence or absence of U6 terminator sequence (TTTT) and whether translational stop codons exists in any of the reading frames (Housden *et al.* 2015). The consideration for the presence of U6 terminator sequence was particularly important since the expression of the sgRNA was to be driven by U6 RNA polymerase III, which, in the presence of terminator sequence, would lead to nonfunctional sgRNA (Housden *et al.* 2015). sgRNAs with the highest predicted efficiency were chosen for targeting selected genes (**Figure 3.3**). It is worth noting that an extra Guanine nucleotide was appended to the 5' of the gRNA sequences since U6 RNA polymerase III prefers G as a first base of the transcript (Guschin *et al.* 2010). In addition to sgRNAs targeting *XYN1* and *MAN2*, sgRNA was also designed to CRISPR-target *PDS* as a positive control for all subsequent transformation events. *PDS* was chosen for this purpose because its mutants have been shown to display an obvious albino phenotype in several plant species including *Populus* (Qin *et al.* 2007; Shan *et al.* 2013; Fan *et al.* 2015). Thus, the obvious albino phenotype would be useful to confirm not only cloning success but also transformation efficiency. With all sgRNAs designed, the potential for off-target cleavage activities was evaluated through the aspen online database. Several potential cleavage regions were identified, however, none of them were immediately followed by a SpCas9-recognizable PAM (**Figure 3.4**), therefore, these regions would most likely not be recognized by Cas9 (Reviewed in Ran *et al.* 2013b). While this *in silico* approach is useful for rapid detection of potential off-targets as seen here, it is noteworthy that it does not precisely detect potential

off-target mutations that occur *in vivo* (Fu *et al.* 2013; Hsu *et al.* 2013; Mali *et al.* 2013a; Cho *et al.* 2014). Therefore, more efficient methods are required for the detection of off-target mutations at that level (Reviewed in Zhang *et al.* 2015).

3.5.3 Putative transformants recovery and mutation screening of T0 plants

The poor transformants recovery observed here may generally be attributed to the stringent plant resistance regulatory mechanisms that prevent the stable integration of transgenes. Alternatively, the concentration of kanamycin was suspected to be perhaps preventing shoot development (as control CRISPR/Cas9-*PDS* shoots were not recovered), to investigate this possibility, explants were transferred to media containing half the starting concentration of kanamycin, however, there was no improvement in shoot regeneration success. An important alternative in understanding the underlying factor behind such poor shoot regeneration will require including appropriate transformation controls (e.g., an empty vector with no sgRNAs), which will be useful in determining if the transfection process itself or the transfection reagents have a toxic effect on plant cells, possibly hindering shoot development. While there may indeed be a technical issue preventing shoot regeneration, it cannot be ruled out that the CRISPR/Cas9-induced mutation in *XYN1* may potentially be lethal, and thus, mutants may not be recovered.

Indeed, no mutants were recovered in the current study. The lack of genetic modification in the CRISPR/Cas9-*XYN1* transgenic plant was somewhat intriguing considering that the transgenic plant was significantly stunted (**Figure 3.11**), giving strong hints that it could potentially be mutant. Such lack of mutation may have to do with the location of integration of Cas9-sgRNA in the genome, this will require Southern blotting to confirm, however, such analyses are beyond the scope of this study. Notably, although no mutation was observed in *Populus*, the poor growth observed in *AtXYN1* may be

explained by other factors, which are unknown at this moment. Furthermore, since in poplar secondary cell walls are even larger carbon sink compared to *Arabidopsis*, *XYN1* may play a more secondary role in poplar, thus, a further attempt at generating mutants in *Populus XYN1* is highly recommended. In contrast to the significantly poor CRISPR/Cas9-*XYN1* putative transformants recovery, numerous CRISPR/Cas9-*MAN2* T0 shoots were recovered, however, the majority of these died as seedlings. The one plant that survived and contained Cas9 transgene but was not edited, just like with CRISPR/Cas9-*XYN1* line described above (**Figure 3.11**). Considering that only one plant was screened for mutation from each line, the likelihood of such one plant being genetically modified was indeed very low.

3.5.4 Possible future solutions to the lack of target site modification

Amongst others, the inefficiency and poor quality of the specific sgRNA in directingCas9 to the target exon and mutating the gene may be one of the major limitations leading to no target site modifications. This is possible since different sgRNAs targeting a similar gene often possess different efficiencies as previously observed when four different sgRNAs were used to target *PDS* in *Populus tomentosa* (Fan *et al.* 2015). To address this, the sgRNA can either be pre-screened through the transient tobacco approach or the target site can be targeted with 3 different sgRNAs through established multiplexing methods (Lowder *et al.* 2015; Liang *et al.* 2019). This way, an efficient sgRNA can be identified, and the efficiency of CRISPR/Cas9-mediated genome editing can be improved. Furthermore, a screening marker can be included to visualize the *in vivo* Cas9-mediated activity, this is essential in identifying positive transformants (Hahn *et al.* 2017; Yu and Zhao 2019). In addition to the sgRNA efficiency and quality, an alternative explanation for the lack of target site modification is the possibility that the target gene may be missing some essential parts, therefore, it will be important to verify the integrity of the target gene. Additionally, a possible future work will be to confirm the gene expression of Cas9 and sgRNA in the transgenic poplars. It may also be essential to perform gene expression modulation

rather than knock-out, this may yield more biotechnologically relevant and potentially desirable phenotypes.

3.6 References

Acker R. Van, J. C. Leple, D. Aerts, V. Storme, G. Goeminne, *et al.*, 2014 Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. Proc. Natl. Acad. Sci. 111: 845–850. https://doi.org/10.1073/pnas.1321673111

Barra-Jiménez A., and L. Ragni, 2017 Secondary development in the stem: when *Arabidopsis* and trees are closer than it seems. Curr. Opin. Plant Biol. 35: 145–151. https://doi.org/10.1016/j.pbi.2016.12.002

Barros J., H. Serk, I. Granlund, and E. Pesquet, 2015 The cell biology of lignification in higher plants. Ann. Bot. 115: 1053–1074. https://doi.org/10.1093/aob/mcv046

Cho S. W., S. Kim, Y. Kim, J. Kweon, H. S. Kim, *et al.*, 2014 Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. 24: 132–141. https://doi.org/10.1101/gr.162339.113

Coleman H. D., J.-Y. Park, R. Nair, C. Chapple, and S. D. Mansfield, 2008 RNAi-mediated suppression of p-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. Proc. Natl. Acad. Sci. 105: 4501–4506. https://doi.org/10.1073/pnas.0706537105

Du Q., W. Pan, J. Tian, B. Li, and D. Zhang, 2013 The UDP-Glucuronate decarboxylase gene family in *Populus*: Structure, expression, and association genetics. PLoS One 8: e60880. https://doi.org/10.1371/journal.pone.0060880

Fan D., T. Liu, C. Li, B. Jiao, S. Li, *et al.*, 2015 Efficient CRISPR/Cas9-mediated targeted mutagenesis in *Populus* in the first generation. Sci. Rep. 5: 1–7. https://doi.org/10.1038/srep12217

Fu Y., J. A. Foden, C. Khayter, M. L. Maeder, D. Reyon, *et al.*, 2013 High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. 31: 822–826. https://doi.org/10.1038/nbt.2623

Gaj T., C. A. Gersbach, and C. F. Barbas, 2013 ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 31: 397–405. https://doi.org/10.1016/j.tibtech.2013.04.004

Guschin D. Y., A. J. Waite, G. E. Katibah, J. C. Miller, M. C. Holmes, *et al.*, 2010 A Rapid and general assay for monitoring endogenous gene modification, pp. 247–256 in *Engineered Zinc Finger Proteins: Methods and Protocols*, edited by Mackay J. P., Segal D. J. Humana Press, Totowa, NJ.

Hahn F., O. Mantegazza, A. Greiner, P. Hegemann, M. Eisenhut, *et al.*, 2017 An efficient visual screen for CRISPR/CAS9 activity in *Arabidopsis thaliana*. Front. Plant Sci. 8: 1–13. https://doi.org/10.3389/fpls.2017.00039

Housden B. E., A. J. Valvezan, C. Kelley, R. Sopko, Y. Hu, *et al.*, 2015 Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi. Sci. Signal. 8: 1–10. https://doi.org/10.1126/scisignal.aab3729

Hsu P. D., D. A. Scott, J. A. Weinstein, F. A. Ran, S. Konermann, *et al.*, 2013 DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. 31: 827–832. https://doi.org/10.1038/nbt.2647

Iglesias-Fernández R., M. C. Rodríguez-Gacio, C. Barrero-Sicilia, P. Carbonero, and A. Matilla, 2011a Three *ENDO-β-MANNANASE* genes expressed in the micropylar endosperm and in the radicle influence germination of Arabidopsis thaliana seeds. Planta 233: 25–36. https://doi.org/10.1007/s00425-010-1257-z

Iglesias-Fernández R., M. del C. Rodríguez-Gacio, C. Barrero-Sicilia, P. Carbonero, and A. J. Matilla, 2011b Molecular analysis of *ENDO-β-MANNANASE* genes upon seed imbibition suggest a cross-talk between radicle and micropylar endosperm during germination of *Arabidopsis thaliana*. Plant Signal. Behav. 6: 80–82. https://doi.org/10.4161/psb/14126

Jacobs T. B., P. R. LaFayette, R. J. Schmitz, and W. A. Parrott, 2015 Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnol. 15: 1–10. https://doi.org/10.1186/s12896-015-0131-2

Jiang W., H. Zhou, H. Bi, M. Fromm, B. Yang, *et al.*, 2013 Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. Nucleic Acids Res. 41: e188–e188. https://doi.org/10.1093/nar/gkt780

Joshi C. P., S. Bhandari, P. Ranjan, U. C. Kalluri, X. Liang, *et al.*, 2004 Genomics of cellulose biosynthesis in poplars. New Phytol. 164: 53–61. https://doi.org/10.1111/j.1469-8137.2004.01155.x

Jyothishwaran G., D. Kotresha, T. Selvaraj, S. M. Srideshikan, P. K. Rajvanshi, *et al.*, 2007 A modified freeze–thaw method for efficient transformation of *Agrobacterium tumefaciens*. Curr. Sci. 93: 770–772.

Mali P., J. Aach, P. B. Stranges, K. M. Esvelt, M. Moosburner, *et al.*, 2013a CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. 31: 833–838. https://doi.org/10.1038/nbt.2675

Mali P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, *et al.*, 2013b RNA-guided human genome engineering via Cas9. Science. 339: 823–826. https://doi.org/10.1126/science.1232033

Mojica F. J. M., C. Díez-Villaseñor, J. García-Martínez, and C. Almendros, 2009 Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology 155: 733–740. https://doi.org/10.1099/mic.0.023960-0

Park S., S. Oh, and K.-H. Han, 2004 Large-scale computational analysis of poplar ESTs reveals the repertoire and unique features of expressed genes in the poplar genome. Mol. Breed. 14: 429–440. https://doi.org/10.1007/s11032-004-0603-x

Pawar P. M. A., M. Derba-Maceluch, S. L. Chong, M. L. Gandla, S. S. Bashar, *et al.*, 2017a *In muro* deacetylation of xylan affects lignin properties and improves saccharification of aspen wood. Biotechnol. Biofuels 10: 1–11. https://doi.org/10.1186/s13068-017-0782-4

Pawar P. M. A., C. Ratke, V. K. Balasubramanian, S. L. Chong, M. L. Gandla, *et al.*, 2017b Downregulation of *RWA* genes in hybrid aspen affects xylan acetylation and wood saccharification. New Phytol. 214: 1491–1505. https://doi.org/10.1111/nph.14489

Qin G., H. Gu, L. Ma, Y. Peng, X. W. Deng, *et al.*, 2007 Disruption of phytoene desaturase gene results in albino and dwarf phenotypes in *Arabidopsis* by impairing chlorophyll, carotenoid, and gibberellin biosynthesis. Cell Res. 17: 471–482. https://doi.org/10.1038/cr.2007.40

Ran F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, *et al.*, 2013 Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8: 2281–2308. https://doi.org/10.1038/nprot.2013.143

Ratke C., B. K. Terebieniec, S. Winestrand, M. Derba-Maceluch, T. Grahn, *et al.*, 2018 Downregulating aspen xylan biosynthetic GT43 genes in developing wood stimulates growth via reprograming of the transcriptome. New Phytol. 219: 230–245. https://doi.org/10.1111/nph.15160

Shan Q., Y. Wang, J. Li, Y. Zhang, K. Chen, *et al.*, 2013 Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 31: 686–688. https://doi.org/10.1038/nbt.2650

Sundell D., N. R. Street, M. Kumar, E. J. Mellerowicz, M. Kucukoglu, *et al.*, 2017 AspWood: High-Spatial-Resolution Transcriptome Profiles Reveal Uncharacterized Modularity of Wood Formation in *Populus tremula*. Plant Cell 29: 1585–1604. https://doi.org/10.1105/tpc.17.00153

Taylor G., 2002 *Populus*: *Arabidopsis* for forestry. Do we need a model tree? Ann. Bot. 90: 681–689. https://doi.org/10.1093/aob/mcf255

Taylor N. G., 2008 Cellulose biosynthesis and deposition in higher plants. New Phytol. 178: 239–252. https://doi.org/10.1111/j.1469-8137.2008.02385.x

Turco G. M., K. Kajala, G. Kunde-Ramamoorthy, C. Y. Ngan, A. Olson, *et al.*, 2017 DNA methylation and gene expression regulation associated with vascularization in Sorghum bicolor. New Phytol. 214: 1213–1229. https://doi.org/10.1111/nph.14448

Wierzbicki M. P., N. Christie, D. Pinard, S. D. Mansfield, E. Mizrachi, *et al.*, 2019 A systems genetics analysis in *Eucalyptus* reveals coordination of metabolic pathways associated with xylan modification in wood-forming tissues. New Phytol. 223: 1952–1972. https://doi.org/10.1111/nph.15972

Yu H., and Y. Zhao, 2019 Fluorescence Marker-assisted isolation of Cas9-free and CRISPR-edited *Arabidopsis* plants, pp. 147–154 in *Plant Genome Editing with CRISPR Systems: Methods and Protocols*, edited by Qi Y. Springer New York, New York, NY.

Zhang X. H., L. Y. Tee, X. G. Wang, Q. S. Huang, and S. H. Yang, 2015 Off-target effects in CRISPR/Cas9-mediated genome engineering. Mol. Ther. - Nucleic Acids 4: e264. https://doi.org/10.1038/mtna.2015.37

3.7 Tables and figures

Table 3.1 Primer sequences used for all amplification reactions in this research chapter.

Primer designation	Sequence (5′-3′)	Application[a]
XYN_F	TCACTTGTTCCTGTCCAAGC	sgRNA target screening in P717 (**3.3.4**) and screening of putative transgenic plants (**3.3.7**)
XYN_R	CCCAGTCCTTATCTGTTGCC	
MAN_F	TTGATGCCAACATTCCACCA	
MAN_R	GCCCGAGTGGTAAGGTAAAG	
PDS_F	AGAGCAAAAGTCACTGCCAA	
PDS_R	GCGGAGAAGAACGAAAGGAT	
Scaffold_F	GTTTTAGAGCTAGAAATAGCAAGTT	Amplicon prep for Gibson assembly (**3.3.5**)
SpeI_ScaffoldR*	AAAAAAAGCACCGACTCGGTG	
SwaI_MtU6F*	ATGCCTATCTTATATGATCAATGAGG	
MtU6R	AAGCCTACTGGTTCGCTTGAAG	
ISceI_R	GTGATCGATTACCCTGTTATCCCTAG	p201N Cas9 construct screening for inserts (**3.3.5**)
Ubi3p218_R	ACATGCACCTAATTTCACTAGATGT	
Cas9R1	GCTGTGGCGATCGGTATTGC	Cas9 transgene screening **(3.3.7)**
Cas9F1	CATATGATCAAATTTCGGGGACACTTC	

[a]The purpose of primer in the current study and the relevant sections in which the primer was used.
*SpeI_Scaffold reverse primer consist of the 5' GTCATGAATTGTAATACGACTCA overhang, while SwaI_MtU6F consist of 5' GATATTAATCTCTTCGATGAAATTT overhang. Overhangs are necessary for annealing scaffold and U6 to linearized p201N Cas9 vector in Gibson assembly, see Jacobs et al. 2015.

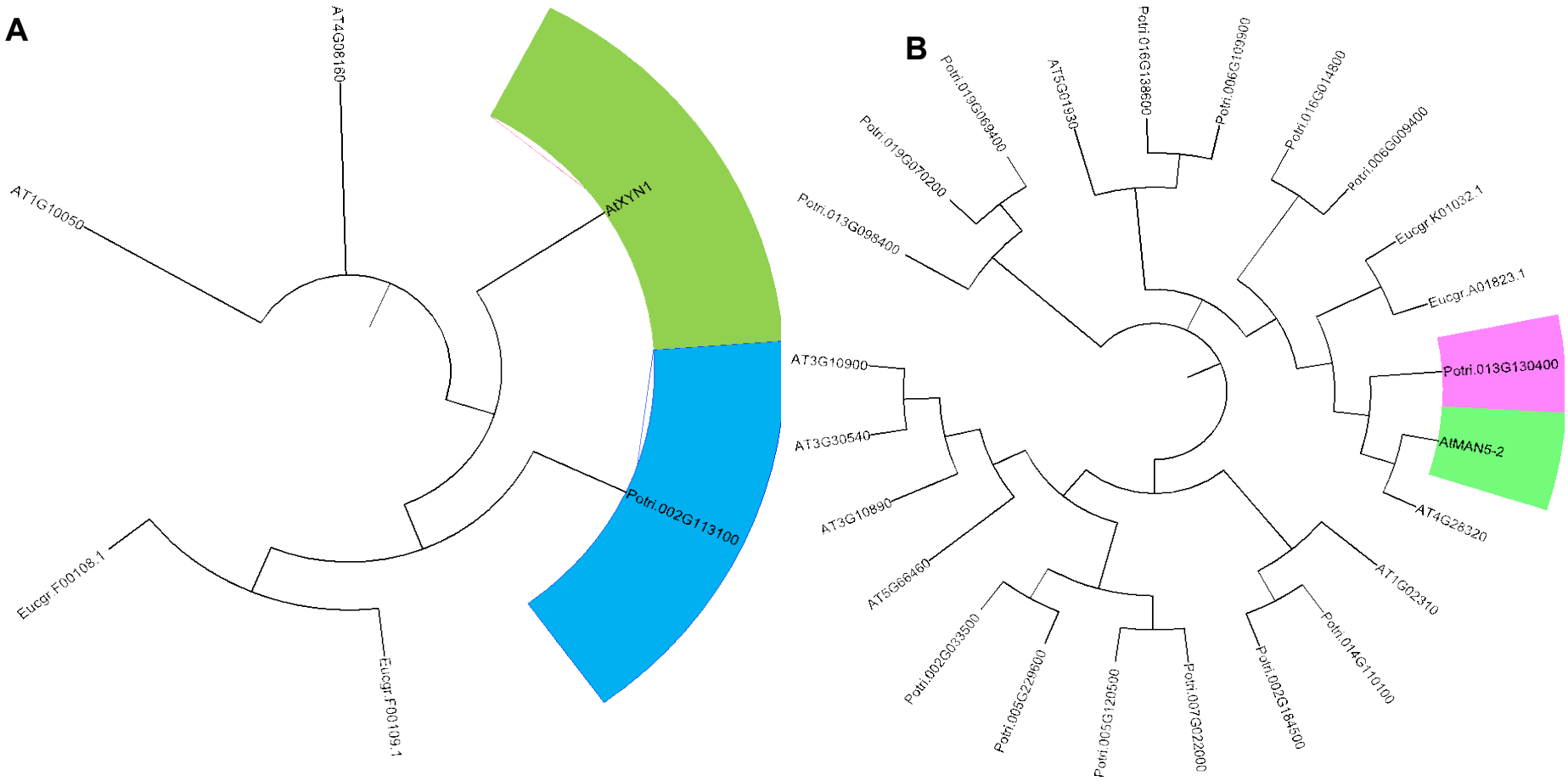


Figure 3.1 Phylogenetic trees representing homologs of MAN2 and XYN1 in *Arabidopsis*, *Eucalyptus* and *Populus*.

Phylogenetic trees showing *Arabidopsis thaliana*, *Eucalyptus grandis*, and *Populus trichocarpa* homologs. *Arabidopsis* and *Populus* homologs are coloured. *Populus* XYN1 (Potri.002G113100, tree **A**) and MAN2 (Potri.013G130400, tree **B**) have no paralogs and were chosen for CRISPR editing in the current study, Phylogenetic tree generated with Geneious Prime® 2020.0 and modified in iTOL v5.

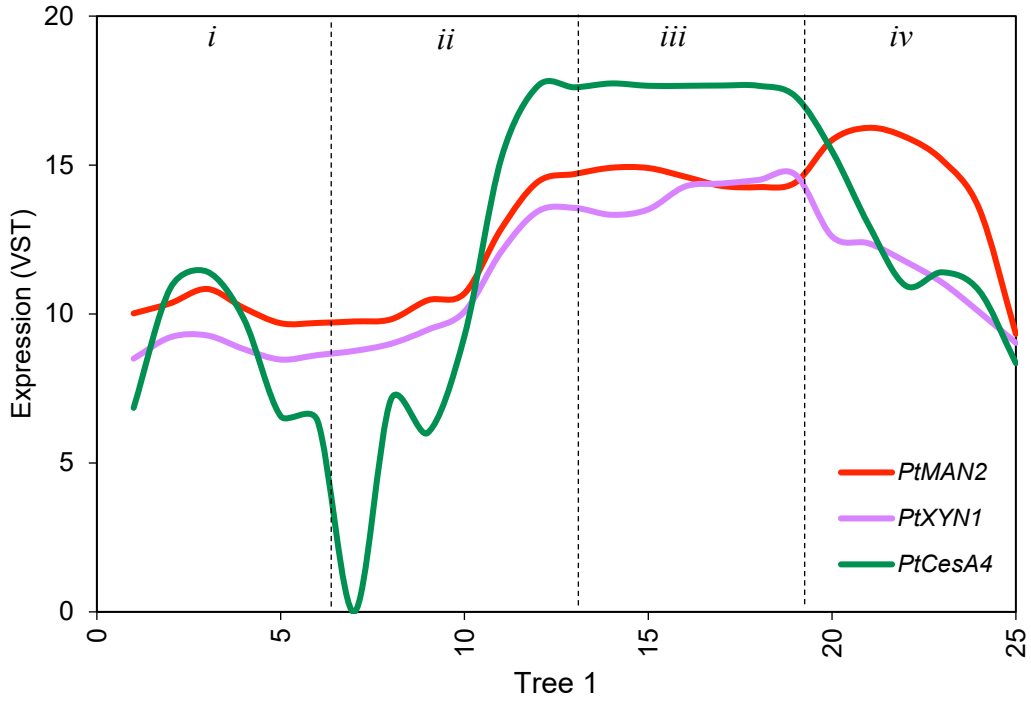


Figure 3.2 Gene expression profiles of *PtMAN2* and *PtXYN1* generated from *P. tremula (Pt)* RNAseq data.

The expression profile for *MAN2* and *XYN1 P. trichocarpa* orthologs. Gene expression profile of *CESA4* included as a reference. **i-iv** are tree developmental stages; **i:** phloem and xylem differentiation, **ii:** cell expansion, **iii:** secondary cell wall deposition, **iv**: cell death. Dotted lines mark the transition between developmental stages. VST: variance stabilizing transformation

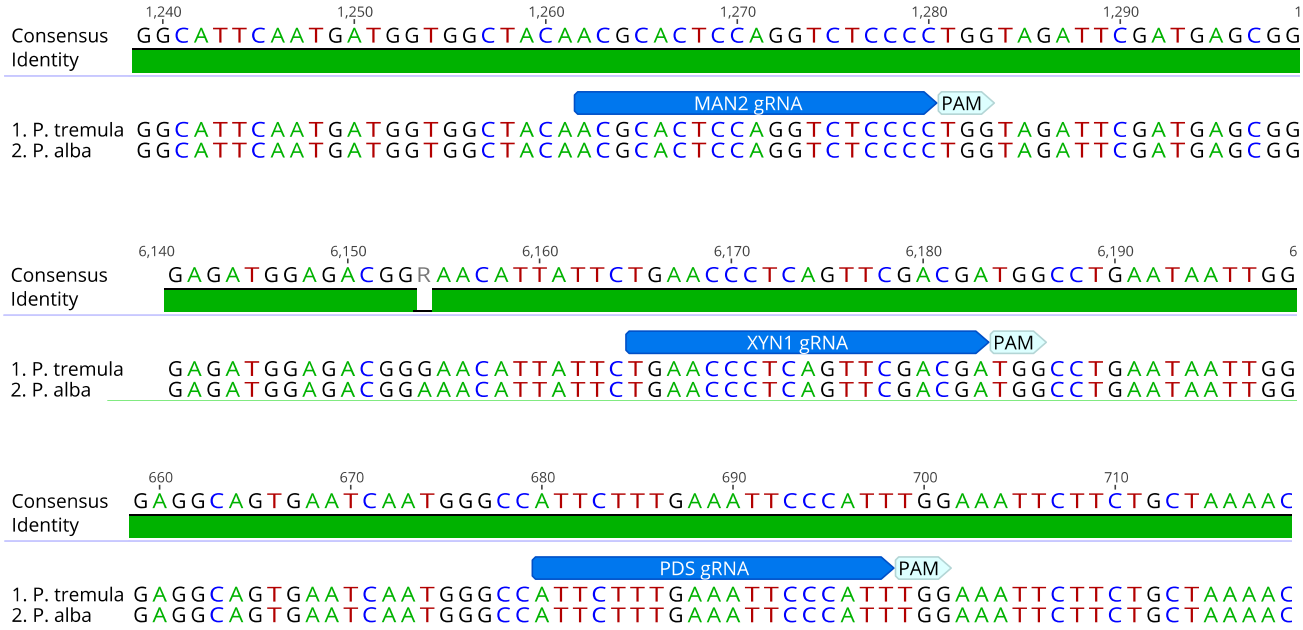


Figure 3.3 Sequences of *MAN2*, *XYN1* and *PDS* sgRNA obtained through Sanger sequencing.

sgRNAs were designed in 5'-GN_{19} -3' form to target genes in P717. N_{19} and PAM annotated. *MAN2* and *PDS* sgRNAs target exon 1 of these respective genes, while *XYN1* sgRNA targets exon 8. Images generated and modified in Geneious Prime® 2020.0.

Target **ATTCTTTGAAATTCCCATT|TGG|Potri.014G148700_*PDS***

Tremula ATTCTTTGAAATACCCTTT|CTT Off-target 1: Potri.006G132700
Alba ATTCTTTGAAATACCCTTT|CTT

Tremula ATTCTTTGAGTTTCCCATT|CAG Off-target 2: Potri.008G160200
Alba ATTCTTTGAGTTTCCCAAT|CAG

Target **TGAACCCTCAGTTCGACGA|TGG|Potri.002G113100_*XYN1***

Tremula TGAACTCTCAGTTCTACAA|AAA Off-target 1: Potri.006G231200
Alba TGAACTCTCAGTTCTACAA|AAA

Target **ACGCACTCCAGGTCTCCCC|TGG|Potri.013G130400_*MAN2***

Tremula ATGCAATCCAGGACTCCCC|TGC Off-target 1: Potri.002G257800
Alba ATGCAATCCAGGACTCCCC|TGC

Tremula ACCAACACCAGGTCTCCCC|GGC Off-target 2: Potri.012G017400
Alba ACCAACACCAGGTCTCCCC|GGC

Figure 3.4 *In silico* prediction of off-target cleavage activity.

P. tremula and *P. alba* genomes were searched for potential sgRNA target sites. Putative target DNA sequences were analyzed to verify whether they are immediately preceded by any of the canonical forms of the SpCas9 PAM sequence. Only potential targets with at most 3 mismatches to the target are shown. Mismatched nucleotide indicated in red, potential PAM sequence is shown in blue.

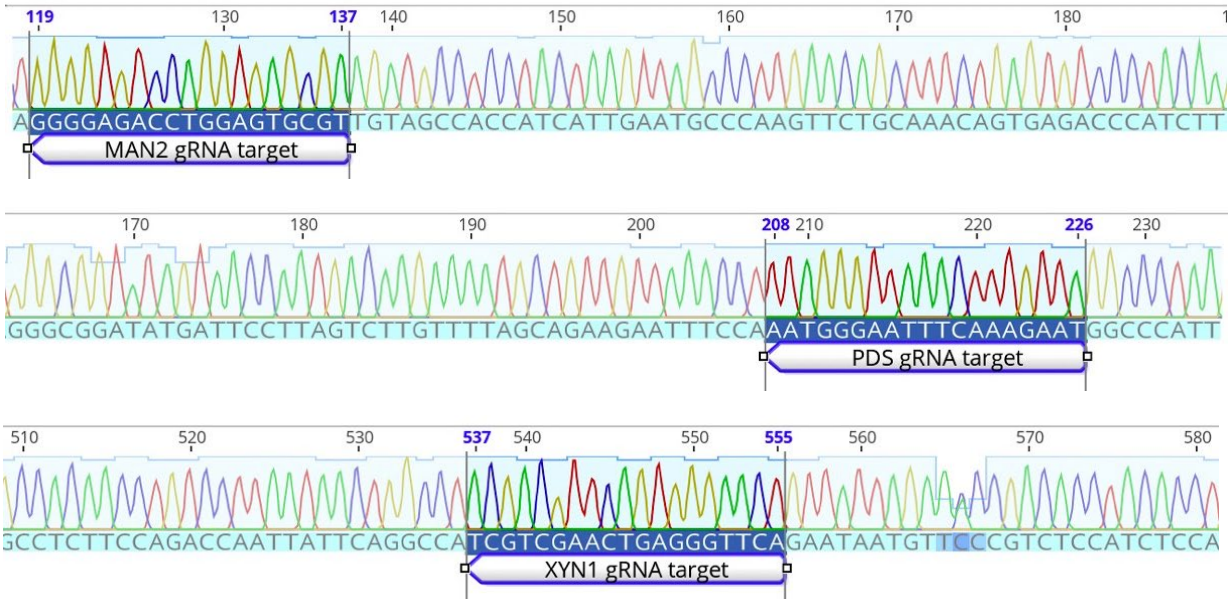


Figure 3.5 Screening sgRNA target regions in P717 through sequencing

P717 sgRNA target region was amplified and sequenced with gene-specific primers. sgRNA targets in the anti-sense DNA strand are highlighted. All target sequences match the designed sgRNA sequences (**Figure 3.3** and **Figure 3.4**).

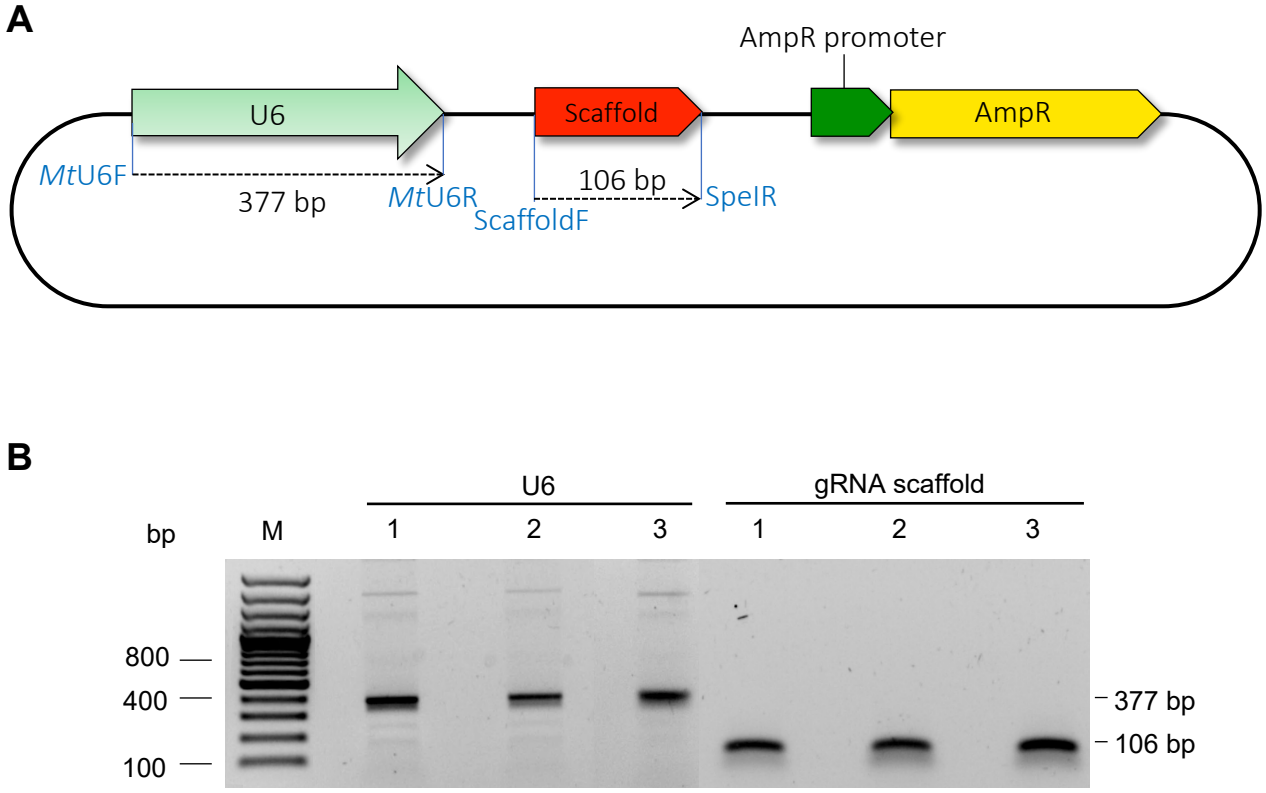


Figure 3.6 Preparation of U6 and gRNA scaffold for Gibson assembly.

(A) pUC gRNA shuttle vector schematic: Ampicillin was used to select the plasmid; U6 and gRNA scaffold were amplified from the plasmid with the amplicon-specific primers as shown in blue (**B**) Amplicon sizes of U6 and scaffold confirmed with PCR, U6 amplicon is 377 bp while the scaffold is 106 bp. M is the Gene Ruler 100 bp DNA Ladder (Thermo Scientific).

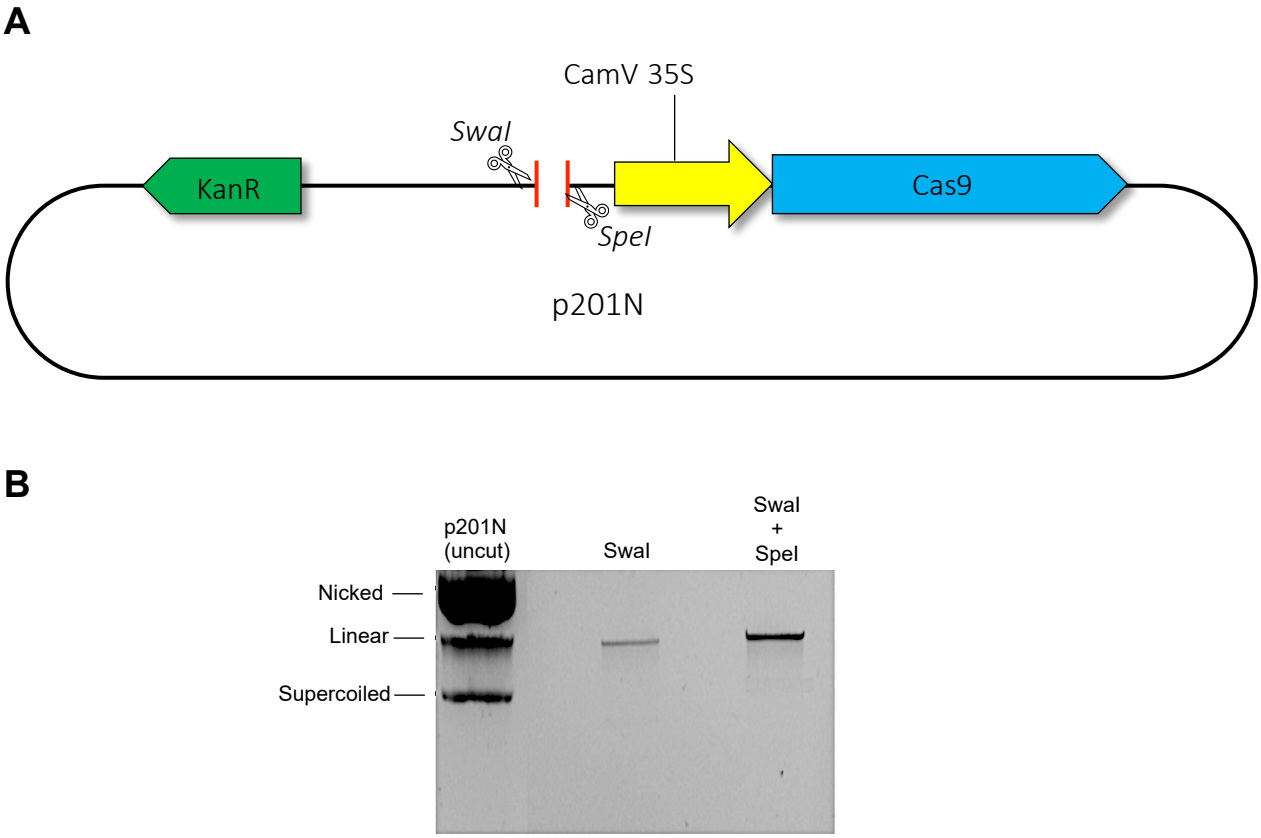


Figure 3.7 Linearizing p201N Cas9 by restriction enzyme digest.

(A) p201N Cas9 schematic, the vector was linearized by restriction enzyme digest with SwaI and SpeI. **(B)** Restriction digest products cut with SwaI and with both SwaI and SpeI were visualized and compared with uncut plasmid on a 1% (w/v) agarose gel.

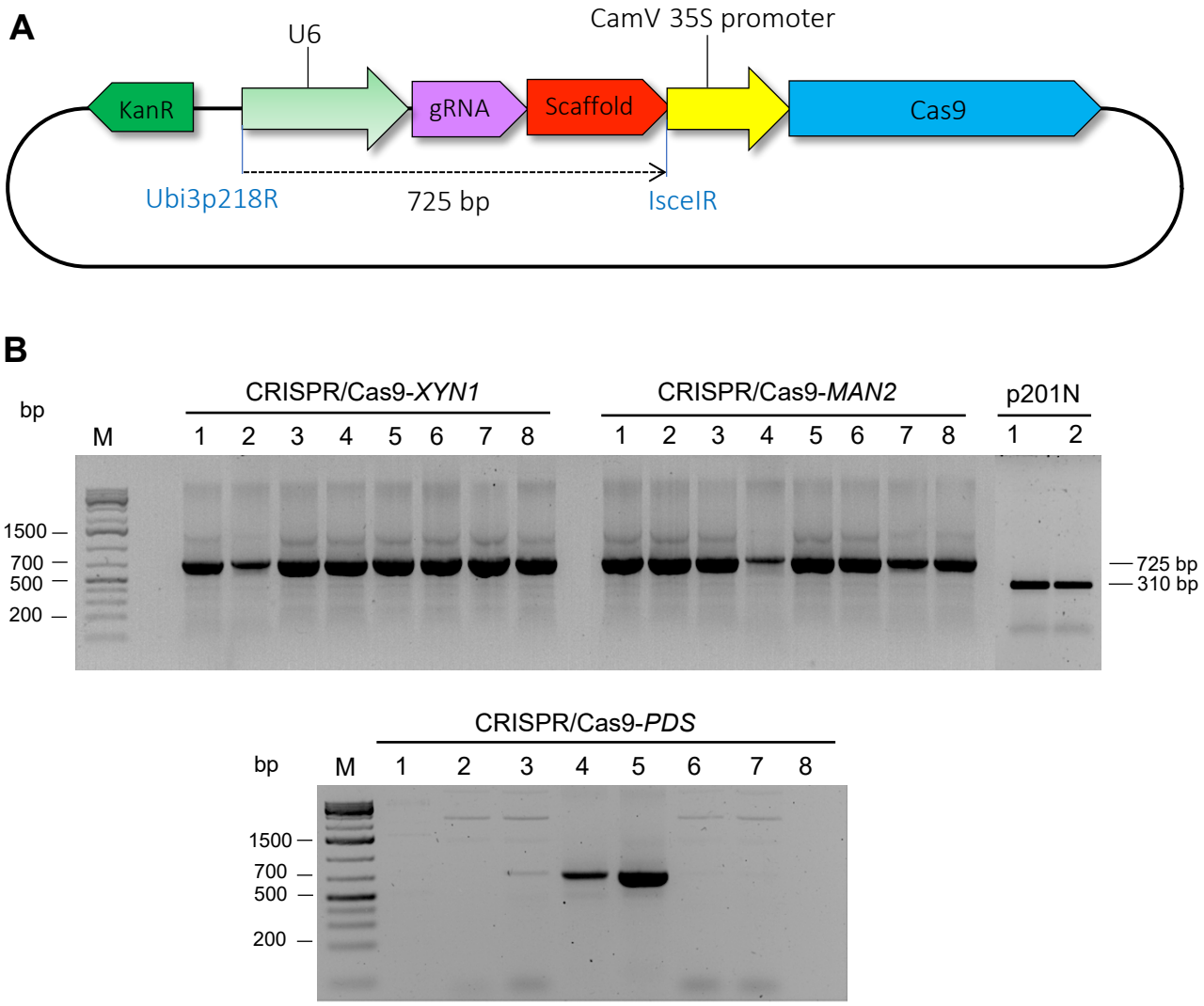


Figure 3.8 Screening CRISPR/Cas9 constructs for inserts through colony PCR.

(A) Schematic showing a portion of CRISPR/Cas9 construct with all the relevant inserts following successful Gibson assembly. **(B)** Confirmation of the presence of inserts through PCR with Ubi3p218R and IsceI R primers. Amplicons visualized and compared with linearized empty p201N vector on a 1% (w/v) agarose gel. Constructs were also Sanger sequenced to confirm that they contained all three inserts with no SNPs (**Figure S 3.2**).

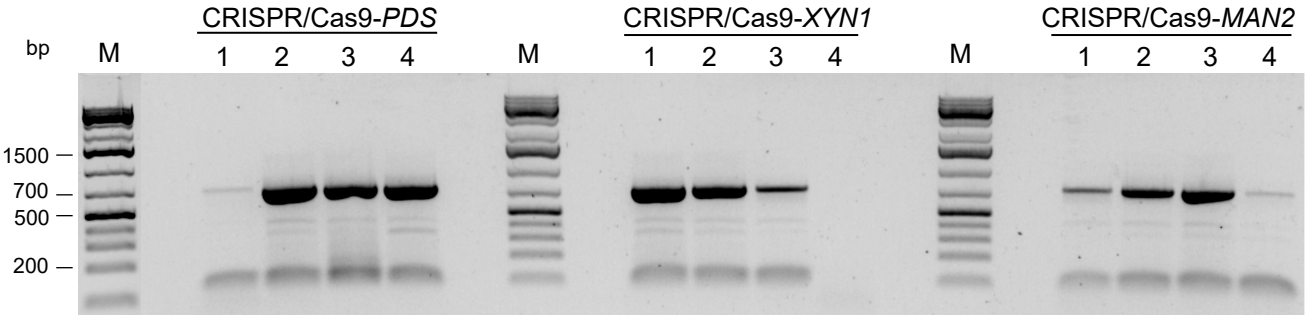


Figure 3.9 Comfirmation of the presence of inserts in transformed *Agrobacterium*.

Following transformation of *Agrobacterium* with CRISPR/Cas9 constructs, four randomly selected putative transformants were screened for inserts through colony PCR. Amplicons visualized in a 1% ethidium bromide-stained agarose gel.

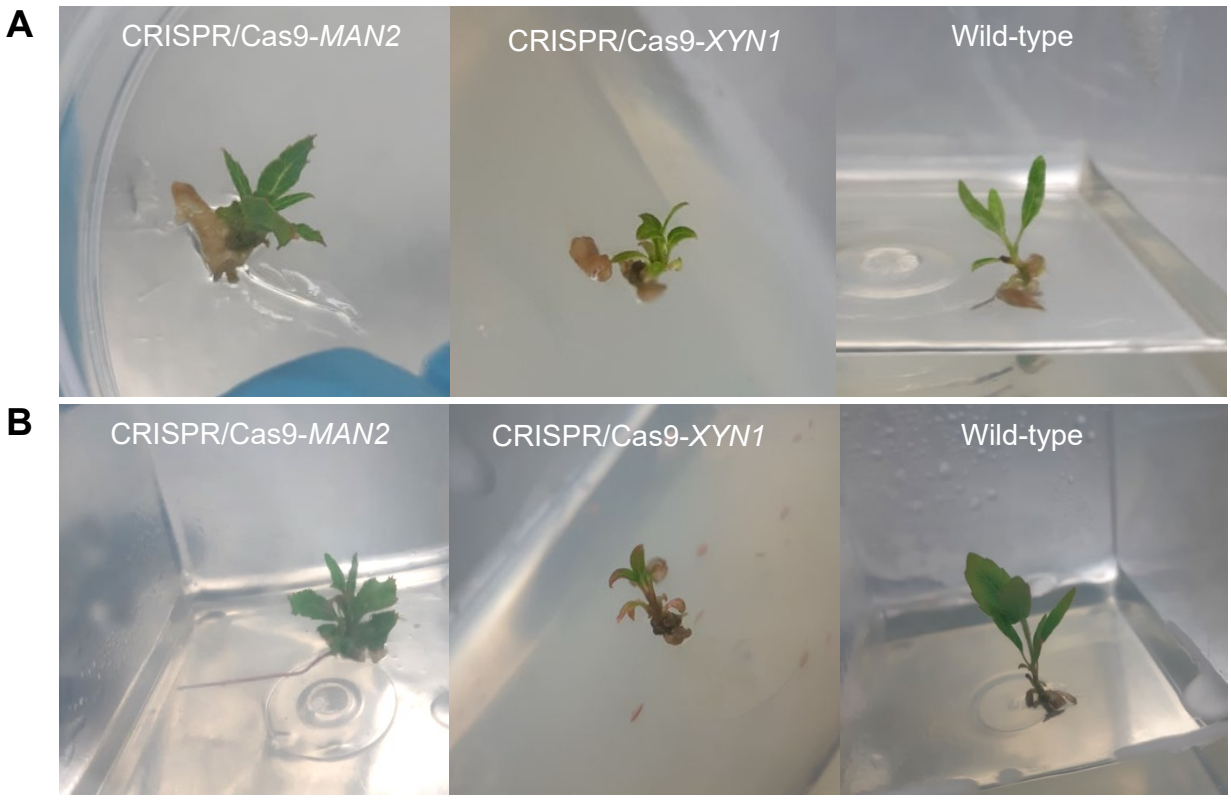


Figure 3.10 Preliminary growth phenotype analyses of *XYN1* and *MAN2* CRISPR lines.

(**A**) Representative shoots in shoot-inducing media (WPM+NAA/BA/TDZ/C containing 25 mg/l Kanamycin) after one month, no shoots were generated from CRISPR/Cas9-PDS line. (**B**) Shoots in pure WPM for three months, *XYN1* CRISPR line show no apparent growth.

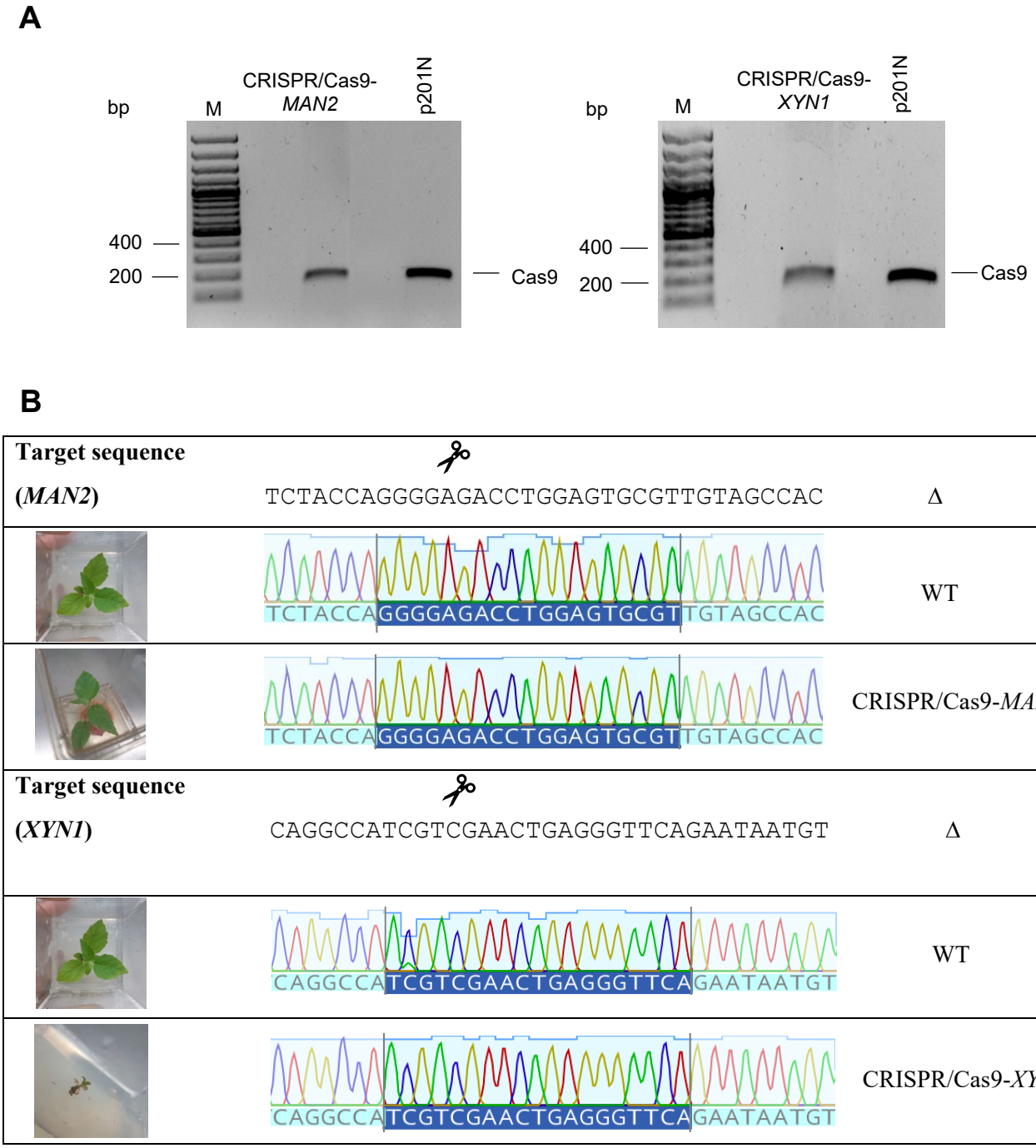


Figure 3.11 Cas9 transgene screening and mutation detection.

DNA was extracted from leaves of putative CRISPR/Cas9 mutants and wild-type. (**A**) The presence of Cas9 transgene was confirmed through PCR with Cas9-specific primers. p201N Cas9 empty vector was included here as a positive control. M is a Gene Ruler 100 bp DNA Ladder. (**B**) Sequence reads from Sanger sequencing for wild-type and putative mutants as visualized on Geneious for detection of gene modification. Target sequences are highlighted in blue and PAM is highlighted in red. Scissors show the region on the target site where a mutation is expected to occur.

3.8 Supplementary Tables and Figures

Table S3.1 Off-target sites determined from Aspen DB probe search online tool. Only potential off targets in the coding sequences are depicted. Mismatched nucleotide shown in red.

PDS	**ATTCTTTGAAATTCCCATT**	Number of mismatches
Tremula	ATTCTTTGAAATACCCTTT	2
Alba	ATTCTTTGAAATACCCTTT	
Tremula	ATTCTTTGAGTTTCCCATT	2
Alba	ATTCTTTGAGTTTCCCAAT	
XYN1	**TGAACCCTCAGTTCGACGA**	
Tremula	TGAACTCTCAGTTCTACAA	3
Alba	TGAACTCTCAGTTCTACAA	
Tremula	TGAATCCCCAGTTCGACCA	3
Alba	CGAATCCCCAGATCGACCA	
Tremula	TGAACCCTCAGATAGAAGG	4
Alba	TGAACCCTCAGATAGAAGG	
Tremula	AGAAACATCAGTTTGACGA	4
Alba	AGAAACATCAGTTTGACGA	
MAN2	**ACGCACTCCAGGTCTCCCC**	
Tremula	ATGCAATCCAGGACTCCCC	3
Alba	ATGCAATCCAGGACTCCCC	
Tremula	ACCAACACCAGGTCTCCCC	3
Alba	ACCAACACCAGGTCTCCCC	
Tremula	ACGCACTCCATCTCTCCAT	4
Alba	ACGCACTCCATCTCTCCAT	
Tremula	GCACTCTCCAGGTCACCCC	4
Alba	GCACTCTCCAGGTCACCCC	
Tremula	GCCCTCTCCAGATCTCCCC	4
Alba	GCCCTCTCCAGATCTCCCC	
Tremula	CCATACTCCAGTTCTCCCC	4
Alba	CCATACTCCAGTTCTCCCC	
Tremula	ACGCACTCCATGTCCCATC	4
Alba	ACGCACTCCATGTCCCATC	
Tremula	ACGCCCTCCAATTTTCCCC	4
Alba	ACGCCCTCCAATTTTCCCC	
Tremula	AAGCACTCCAGCACTACCC	4
Alba	AAGCACTCCAGCACTACCC	

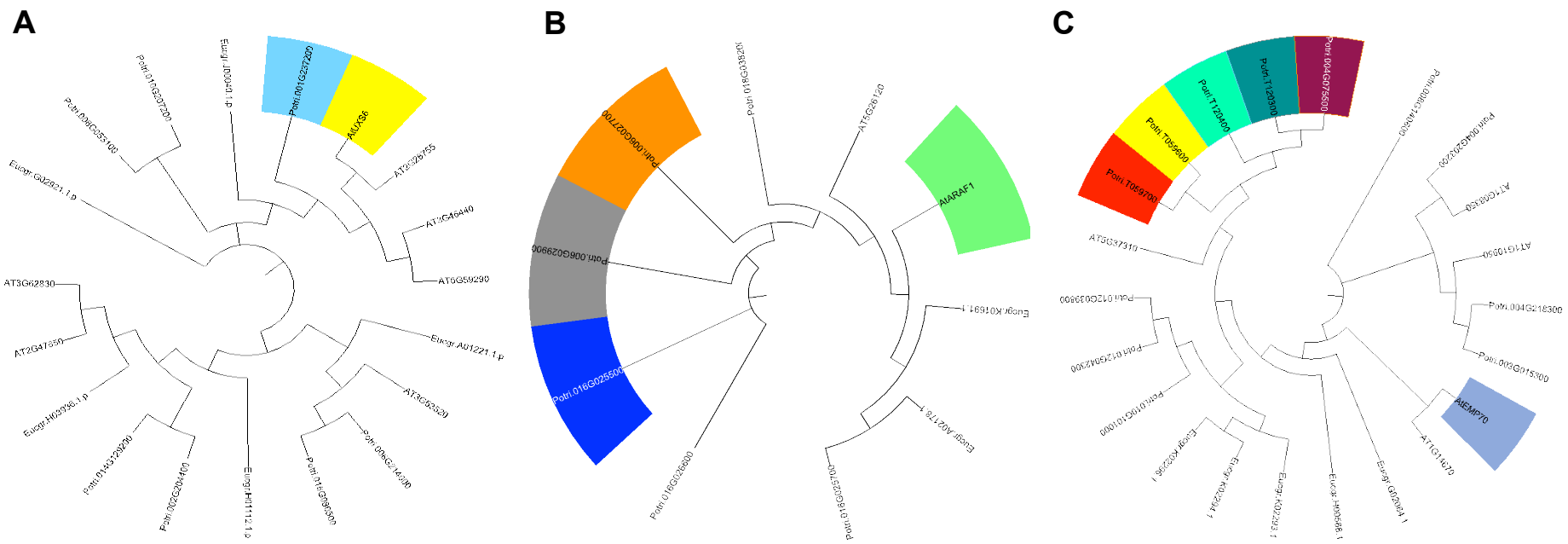


Figure S3.1 Phylogenetic trees representing homologs of UXS6, ARAF1 and EMP70.

UXS6 (Potri.001G237200, tree **A**) have no paralog, while ARAF1 (Potri.006G02770, tree **B**) have three paralogs, and EMP70 (Potri.T059600, tree **C**) have five paralogs. The lack of phenotype for USX6 in *Arabidopsis* disqualified it for CRISPR-editing in *Populus*. Multiple paralogs of ARAF1 and EMP70 demands several gRNAs to be designed for targeting, for feasibility of this study, *ARAF1* and *EMP70* were not selected for testing in *Populus*.

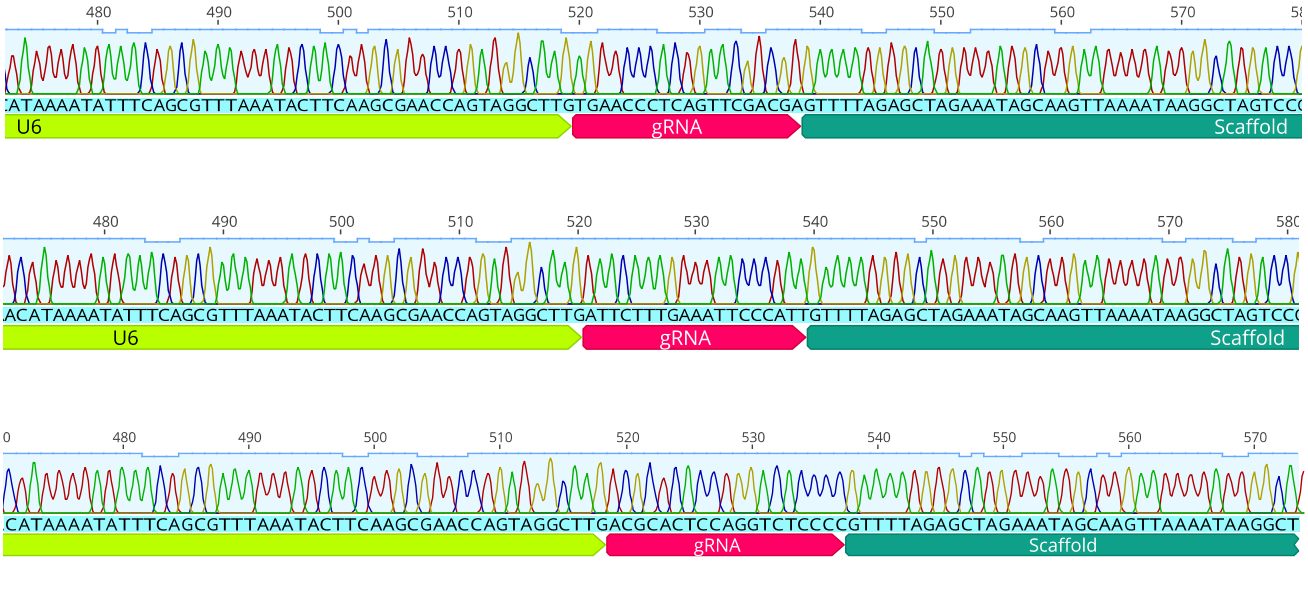


Figure S 3.2 Screening of CRISPR/Cas9 constructs for inserts through Sanger sequencing.

CRISPR constructs targeting *XYN1*, *PDS* and *MAN* (top to bottom) were sequenced with Ubi3p218R primer to confirm successful assembly of U6, sgRNA and scaffold into p201N Cas9 plasmid. Sequences were annotated in Geneious Prime® 2020.2.2.

Chapter 4

Concluding Remarks

4.1 Introduction

In the continued quest to eliminate the use of fossil fuels to alleviate challenges associated with climate change, it has become necessary to rely on lignocellulosic biomass as an alternative source for generating a wide range of value-added products. Such products include bioplastics, packaging materials, textiles, etc. Furthermore, with the advancements in bioprocessing approaches, lignocellulose has become significantly desirable for application in the pharmaceutical industry and the production of novel biomaterials and biofuels. Lignocellulosic biomass is primarily composed of the carbohydrate biopolymers, cellulose and hemicellulose, and the polyphenolic polymer, lignin, and represents the largest carbon sink not only in individual plants but also in the whole biosphere. In the biorefinery, lignocellulose-degrading enzymes, such as cellulases, hemicellulases, and lignin modifying enzymes are necessary for the bioconversion of lignocellulose into value-added products (Chandra and Madakka 2019). However, the lignocellulose structure, modifications thereof, and the tight interaction amongst the biopolymers contributes to biomass recalcitrance, which is the resistance of lignocellulose to bioprocessing (Himmel *et al.* 2007; Zeng *et al.* 2017). Such recalcitrance hampers the feasibility of large-scale production of bio-based products from lignocellulosic biomass. Accordingly, a comprehensive understanding of lignocellulose structure and how it precisely affects bioconversion is vital. Exploring the relationship between lignocellulose structure and bioprocessing starts with a good understanding of how lignocellulose biopolymers are synthesized, which in turn requires a substantial understanding of the precise biological roles of biosynthetic genes involved.

In this study, functional analysis of five xylan-associated genes was performed, focusing specifically on the impact of gene knock-out mutations on plant growth and secondary cell wall development. In this chapter, I discuss how the specific genes selected for analyses here were chosen by referring to the primary systems genetics study that identified these genes. I further discuss the implications of

growth phenotypes and cell wall morphology of mutants with reference to plant growth and physiology. More importantly, this being the very first study to analyze the growth phenotype of *EMP70* mutants in plants, we describe the growth advantage phenotype observed and suggest a functional model describing the possible role of *EMP70* in *Arabidopsis* seed germination. Finally, the functional analyses performed in *Arabidopsis* were further extended to poplar, since this is one of the major energy crops with substantially high biomass yield, and therefore more important for bioproducts production. Using the CRISPR/Cas9 gene editing tool, we generated transgenic poplar plants although they were not genetically modified at the target locus. Altogether, the work done here presents a starting point in understanding the exact functions of some of the essential cell wall biosynthetic genes. As previously mentioned, understanding the exact roles of such genes in wood formation is an essential step towards dissecting the impact of lignocellulose structure on bioprocessing efficiency and potentially reducing biomass recalcitrance.

4.2 Exploring the broad implications of candidate gene selection

Knowing which genes to target to achieve the desired bioprocessing traits is an important first step in any functional genetics study, this is essential, in part to evade unwanted growth penalties as observed in various studies (Xiong *et al.* 2013b; Petrik *et al.* 2020). Here, selecting candidate genes to be functionally tested relied on a previous systems genetics study which established a suite of putative xylan-associated genes in a *Eucalyptus* backcross population (Wierzbicki *et al.* 2019a). While relying on this data did not necessarily avoid the unwanted traits, it helped predict which genes were more likely to yield phenotypes desirable to bioprocessing. This is because, in addition to identifying the genes, the expression of these genes was correlated with traits that are known to affect the efficiency of bioprocessing (**Figure 2.1**; **Figure 2.2**). With this understanding, genes from biotechnology cluster 3 (namely, *XYN1*, *MAN2*, *UXS6*, *EMP70*, and *ARAF1*) were selected for functional analyses since their

expression profile was negatively correlated with xylan acetylation and positively associated with pulp yield (Wierzbicki *et al.* 2019a; **Figure 2.2**).

In addition to the gene expression profile and correlation at the population level, several other aspects at the molecular level ought to be considered to understand if targeting these genes can potentially affect cell wall development. For example, whether these genes are predominantly expressed in xylem vessels (which are composed of the secondary cell wall), and if they are, if their expression is xylem specific. Such discrepancy may exist most likely due to the possible mixed nature of the samples used for expression profiling in the systems genetics study. Through interrogation of literature, it was established that from the selected genes, only *AtARAF1*, *AtUXS6*, and *AtXYN1* are highly expressed in the xylem vessels, however, their expression is not restricted to xylem (Suzuki *et al.* 2002; Fulton and Cobbett 2003; Chavez Montes *et al.* 2008; Zhong *et al.* 2017). Therefore, an important question for all BC 3 genes would be why genes within this cluster all are transcriptionally coordinated with other secondary cell wall biosynthetic genes even though some are not xylem specific. Such questions emphasize the possibility that some complex relationships exist amongst the biosynthetic genes, which, for example, may point to the complexity of metabolism associated with secondary cell wall development. Furthermore, it is worth noting that the similarity of the gene expression profile of BC 3 genes at the population level may not necessarily represent the gene expression at the molecular level. Thus, the selected genes (and other genes in BC 3) may not be functioning in concert with other genes and therefore acting redundantly as reflected by the lack of strong and consistent mutant phenotypes of these genes in the current study (**Table S2.4**).

4.3 Lessons from functional gene testing in *Arabidopsis* and *Populus*

Arabidopsis XYN1 T-DNA insertional mutants showed a slow growth during early development as exhibited by the significantly smaller rosette leaves (**Figure 2.3**). We interrogated the underlying physiological aspect that could explain the poor growth phenotype of *Atxyn1* plants, and we showed that mutants have noticeable smaller roots (**Figure 2.6**), which may present limitations for efficient water and nutrient transport from roots to leaves. Supporting this possibility, overexpression of similar gene lead to increased transport from stem to leaves (Endo *et al.* 2019). Further evidence emphasize the possible involvement of *XYN1* in drought resistance (Goujon *et al.* 2003; Wang *et al.* 2013). Thus, it is suggested that *XYN1* may be playing an important role in the regulation of water transport from roots to the leaves, however, future analyses of the growth phenotype of mutants under different drought conditions are necessary to validate the possibility of this biological role. Notably, *xyn1* mutants were able to grow just like wild-type throughout the late vegetative growth phase, which is expected considering that several putative xylanases may compensate for the loss of XYN1 activity (Suzuki *et al.* 2002), therefore, sophisticated enzymatic studies that explore the exact biological role of XYN1-related putative xylanases will be beneficial to establish which xylanase plays the most prominent role in secondary cell wall biosynthesis.

Ultimate analyses of lignin content showed a significantly higher lignin content for *xyn1* mutants (**Figure 2.11**). This was a somewhat intriguing result considering that *xyn1* mutants did not show any notable increase in secondary cell development when xylem vessel morphology and monosaccharide content were examined (**Figure 2.8**; **Figure 2.9**). While there are several explanations for this observation, the most probable is that the accumulation of glucuronoxylan (which is expected in *XYN1* knock-outs) may lead to increased lignin content (Mortimer *et al.* 2010). Finally, with *man2*, *araf1* and *uxs6* mutant lines showing no obvious consistent growth phenotype or alteration of cell wall

morphology and chemistry, the sugar release analyses are necessary for these mutant lines. If mutation in any of these three genes is found to confer significant improvement in sugar release efficiency, genetically engineering cell walls through modification of these genes would be highly recommended for potential improvement of saccharification efficiency without any growth penalties.

4.4 Suggested model for the possible role of *EMP70* in *Arabidopsis*

Contrary to the slow growth observed in *xyn1* mutants, a far more interesting observation was made in *Arabidopsis emp70* mutants, which showed a consistent vegetative growth advantage relative to wild-type (**Table S2.4**). This growth advantage was marked by mutants germinating significantly earlier than wild-type, mutants having larger rosette leaves, more rosettes, and significantly taller inflorescence stem. These observations were not only interesting because they depicted a growth advantage but also because this was the very first study to identify a growth phenotype in *Arabidopsis EMP70* T-DNA insertion mutants. Before this study, the only available information pertaining to the EMP70 was the biochemical analyses that showed that members of this protein family are the most abundant in the Golgi, and more importantly, the *Arabidopsis* EMP70 members localize exclusively to the Golgi (Dunkley *et al.* 2006; Gao *et al.* 2012; Nikolovski *et al.* 2012, 2014). Both these aspects hinted that EMP70 may be important for cell wall biosynthetic processes that occur specifically in the Golgi, such as xylan biosynthesis. Another important aspect of EMP70 is the observation that heterologous expression of the *Arabidopsis* homolog (AT3G13772) in yeast leads to increased intracellular Cu content, which has been suggested to indicate that EMP70 may be a channel transporter involved in the regulation of Cu balance within cells (Hegelund *et al.* 2010). Considering that this is the first study to conduct an in-depth analysis of the growth phenotype and secondary cell wall chemistry of *EMP70* mutants, I assimilated suggestions from all previous studies and our observations to generate a comprehensive model suggesting the possible function of *EMP70* in

Arabidopsis germination (**Figure 4.1**). This model suggests that EMP70 acts as a Cu transporter in addition to some of the characterized Cu transporters such as COPT (Himelblau *et al.* 1998; Abdel-Ghany *et al.* 2005). In this role, EMP70 may be necessary for an extremely fine-tuned regulation of ABA and GA concentration, this may be achieved by regulation of ABA catabolism through modulating Cu concentration (**Figure 4.1**). In the absence of EMP70, intracellular Cu concentration may decrease to levels below the threshold necessary for normal regulation of ABA, leading to dysregulation of ABA concentration and activity, thereby increasing the concentration of GA, which promotes early seed germination (**Figure 4.1**).

It is important noting that this model is based on a very limited number of studies, and thus, it may not present the most accurate function of *EMP70*. Additional studies are necessary to establish the exact function and biological role of this gene family. Another limitation of this model is the inability to describe the possible interacting partners of EMP70, which may affect how EMP70 regulate Cu content and germination time course. Such limitation is due to the lack of prior EMP70 functional characterization information. In addressing these limitations, it is highly recommended that future studies measure the levels of Cu, ABA, and GA in the *Arabidopsis emp70* mutants to further support the model. Finally, it should be noted that the observation that alteration of SCW biosynthetic genes can alter plant physiological processes as predicted by our suggested model is not surprising. Similar relationships between SCW development and processes such as plants' response to abiotic stresses has recently been demonstrated (Coleman *et al.* 2021). Altogether, the phenotypic and cell wall morphology analyses observed in this study are extremely important for future work that aims to further expand on the functional characterization of this novel gene in plants.

4.5 Probing the efficiency of *Agrobacterium*-mediated transformation and CRISPR/Cas9-based gene editing

In the final research chapter, CRISPR/Cas9 approach was employed to generate mutations in *AtXYN1* and *AtMAN2* P717 orthologs. Following several rounds of *Agrobacterium*-mediated transformation, very few *XYN1* and *MAN2* CRISPR/Cas9 putative transformant shoots were recovered, which emphasized the known setback that transformation and gene editing efficiencies are low especially in woody plants. Such recalcitrance to transformation is attributed to several factors such as the transformation system, and the *Agrobacterium* strain used for transformation. For example, LBA4404 *Agrobacterium* has been shown to possess remarkably higher transformation rates compared to EHA105 in some plant species owing to its use of the super binary vector pTO233 with virulence genes derived from the super Ti plasmid pTiBo542 (Jeknic *et al.* 1999; Suzuki *et al.* 2001). However, using LBA4404 per se does not ensure an improved transformation rate because the protocols that employ LBA4404 or EHA105 are time-consuming, tedious, and inefficient, making successful functional characterization through gene editing very challenging. To overcome this limitation, a far more efficient and reliable protocol involving transformation of *Eucalyptus* hairy roots with *Agrobacterium rhizogenes* has been developed (Plasencia *et al.* 2016). Compared to the traditional *Agrobacterium*-mediated transformation approaches such as the one used for transformation of poplar here, plants transformed through the hairy root protocol can easily be detected through fluorescent markers with an the overall transformation efficiency of over 60% (Plasencia *et al.* 2016; Dai *et al.* 2020). Besides transformation protocol, the source and age of explants have also been shown to influence the transformation and regeneration success. Explants from wounded stems, callus, root crowns, and auxiliary meristematic tissue tend to offer greater transformation success compared to explants from mature leaves (Babu and Chawla 2000; Zhao *et al.* 2000; Blazquez *et al.* 2009; Sood *et al.* 2011; Baskaran and van Staden 2012). Therefore, combining the right transformation material and

protocol may improve both transformation efficiency and the potential to regenerate positive transformants.

The conventional approaches for improving transformation described above are not always effective as they are generally species- and genotype-specific, consequently, relying on such simple transformation customizations is not viable for large-scale genetic engineering applications. Accordingly, more elegant approaches are necessary, for example, studies focusing on manipulation of plant developmental genes (DEV genes) as an attempt to enhance transformation efficiency are currently underway (Matsuo *et al.* 2011; Yordanov *et al.* 2014; Nonaka *et al.* 2019). Furthermore, more reliable options for inoculation and regeneration are currently being tested (de Oliveira *et al.* 2011; Chin *et al.* 2018).

On the other extreme, attempts are being made to directly manipulate and improve the efficiency of CRISPR genome editing approach. For example, engineering gRNAs through extension or truncation of either the spacer or scaffold region has been shown to improve both target specificity and editing efficiency (Fu *et al.* 2014; Dahlman *et al.* 2015; Kiani *et al.* 2015). Similarly, a fusion of natural SpCas9 with chromatin-modulating peptides (CMPs) lead to a several-fold increase in gene editing efficiency (Ding *et al.* 2019). Furthermore, it has recently been demonstrated that pairing Cas9 with Ubiquitin 10 (UBQ10) promoter instead of U6 for gRNA expression can lead to a 95% and 70% increase in mutation efficiency in *Arabidopsis* and *Medicago truncatula*, respectively (Wolabu *et al.* 2020). More recently, a cutting-edge approach aimed at addressing most of CRISPR/Cas9-associated limitations, such as off-target activities and chromosomal translocations has been developed. This approach, called CRISPRoff, is unique and effective because it regulate gene function without modifying the underlying DNA sequence (Nuñez *et al.* 2021). Instead, CRISPRoff relies on modulation of epigenetic landscape to silence gene expression (Carlson-Stevermer *et al.* 2020; Nuñez

et al. 2021). Altogether, the potential successes from these ongoing attempts at improving both gene editing and transformation efficiency offer a promise for highly effective targeted gene editing. Coupling efficient gene editing tools with optimized transformation will be highly beneficial for large-scale gene editing-based studies aimed at improving bioprocessing efficiency.

4.6 Conclusions and future perspectives

Through the integration of growth phenotype analyses, in-depth secondary cell chemistry, and morphology analyses, this study established important preliminary functional information pertaining to some of the secondary cell wall-associated genes in *Arabidopsis*. It was shown that *AtXYN1* may be important during early vegetative growth, regulating water and possibly nutrient transport to ensure normal plant growth. More importantly, *AtXYN1* affects secondary cell wall chemistry as observed with increased lignin content in mutants. While the reduction in growth was subtle in these mutants, future work may consider tissue-specific gene knock-outs, to perhaps reduce the overall impact of the mutation on plant growth. In contrast to *xyn1'*s slightly reduced growth, mutant analyses of *EMP70*, a poorly characterized gene in *Arabidopsis,* revealed a growth advantage phenotype accompanied with no alteration in cell wall morphology or monosaccharide composition. Remarkably, further analyses showed that mutant germinates significantly earlier than wild-type, to incorporate these phenotype data, a simple functional model has been generated, which suggest that EMP70 may be involved in the regulation of seed germination by affecting the concentrations of Cu and consequently altering the crosstalk between ABA and GA. This model will assist with further functional characterization not only in *Arabidopsis* but other commercially important plants. More importantly, the consistent growth phenotypes observed in *Arabidopsis xyn1* and *emp70* mutants merit a further examination in a model tree. Finally, by employing the CRISPR/Cas9 approach to genetically modify *XYN1* and *MAN2* in hybrid poplar, the basic aspects of the CRISPR tool were mastered, however, failure to generate

CRISPR-mediated mutations in poplar points to the absolute necessity of optimizing both CRISPR/Cas9-based gene editing and the transformation approaches; these considerations will certainly be prioritized in future studies.

While phenotype analyses performed in this study have provided preliminary information that is important in the subsequent functional characterization of selected xylan-associated genes, sugar release assay is further required to ascertain whether mutation thereof has any impact on the improvement of saccharification yield. This is particularly important perhaps in *emp70* mutants which show a growth advantage, it will thus be important to confirm if this growth advantage is accompanied with improvement in the total saccharification yield. Conversely, the increase in lignin content observed in *xyn1* mutants may certainly impede sugar release, this is a characteristic worth confirming, as it will further indicate whether the increased lignin content observed is a true *xyn1* phenotype. Besides these yet-to-be-assessed aspects, this study suffers from specifically focusing on the impact of mutation only on secondary cell wall aspects; it is possible that some of the genes may be more important in primary wall development as already discussed. On the same note, it is known that several other aspects may affect secondary cell wall development and possibly alter recalcitrance, for example, GlcA modifications can alter xylem vessel morphology and it is suggested that the spacing of these modifications may further alter xylan-cellulose interactions (Mortimer *et al.* 2010; Lee *et al.* 2012a; Bromley *et al.* 2013) , this raises the question that, are mutations in any of the selected genes capable of affecting any of the cell wall substitutions and modifications, and ultimately the interaction of cell wall biopolymers. Answering this question can significantly aid in understanding the role of these genes from a structural perspective. Similarly, the impact of metabolism on cell wall development is important, and it has been demonstrated that desired traits can be obtained by altering the metabolic flux (Fan *et al.* 2017). Thus, another aspect that can be interesting to address is how a mutation in the

selected genes affects the metabolic flux, or the biosynthesis of any cell wall biosynthesis metabolic precursors, if at all.

With only mild growth phenotypes and essentially no impact of the mutation on secondary cell wall chemistry in *Arabidopsis*, it may be inferred that although important, the selected genes may not be indispensable in *Arabidopsis* plant growth and cell wall development. However, considering that the association of these genes with secondary cell wall biosynthesis was originally established from *Eucalyptus* (a far more significant carbon sink compared to *Arabidopsis*) data, it is possible that these genes play more secondary roles in *Eucalyptus* and possibly *Populus* wood formation processes. Indeed, the expression level of all five genes is markedly higher in *Eucalyptus* and *Populus* xylem compared to the *Arabidopsis* cambium as determined from online databases such as PopGenie. Thus, if these genes are knocked out in *Eucalyptus* or *Populus*, a more severe alteration of xylem vessel formation and cell wall chemistry can be expected. Moreover, the possibility of substantial functional involvement of these genes in other plant species besides *Arabidopsis, Eucalyptus,* and *Populus* cannot be ruled out.

Altogether, regardless of the few limitations of the current study and work that is still pending, the current preliminary results lay a strong groundwork for the subsequent functional characterization of selected genes. The potential phenotypes obtained here will be crucial in predicting the traits that can be obtained when homologs of these genes are genetically altered in other economically important plant species. More importantly, the novel transformation and CRISPR/Cas9-mediated gene editing customization approaches that are currently underway promise a substantial improvement of targeted gene editing, plant transformation efficiency, and mutant regeneration. Realizing the exact functional and biological roles of the uncharacterized genes in secondary cell wall biosynthesis will be beneficial in the definitive goal of improving bioprocessing efficiency.

4.7 References

Abdel-Ghany S. E., P. Müller-Moulé, K. K. Niyogi, M. Pilon, and T. Shikanai, 2005 Two P-type ATPases are required for copper delivery in *Arabidopsis thaliana* chloroplasts. Plant Cell 17: 1233–1251. https://doi.org/10.1105/tpc.104.030452

Babu P., and H. S. Chawla, 2000 In vitro regeneration and *Agrobacterium* mediated transformation in gladiolus. J. Hortic. Sci. Biotechnol. 75: 400–404. https://doi.org/10.1080/14620316.2000.11511258

Baskaran P., and J. van Staden, 2012 Somatic embryogenesis of Merwilla plumbea (Lindl.) Speta. Plant Cell. Tissue Organ Cult. 109: 517–524. https://doi.org/10.1007/s11240-012-0118-9

Blazquez S., E. Olmos, J. A. Hernández, N. Fernández-García, J. A. Fernández, *et al.*, 2009 Somatic embryogenesis in saffron (*Crocus sativus* L.). Histological differentiation and implication of some components of the antioxidant enzymatic system. Plant Cell. Tissue Organ Cult. 97: 49–57. https://doi.org/10.1007/s11240-009-9497-y

Bromley J. R., M. Busse-Wicher, T. Tryfona, J. C. Mortimer, Z. Zhang, *et al.*, 2013 GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. Plant J. 74: 423–434. https://doi.org/10.1111/tpj.12135

Carlson-Stevermer J., R. Kelso, A. Kadina, S. Joshi, N. Rossi, *et al.*, 2020 CRISPRoff enables spatio-temporal control of CRISPR editing. Nat. Commun. 11: 1–7. https://doi.org/10.1038/s41467-020-18853-3

Chandra M. R. G. S., and M. Madakka, 2019 Chapter 11 - Comparative Biochemistry and Kinetics of Microbial Lignocellulolytic Enzymes, pp. 147–159 in *Recent Developments in Applied Microbiology and Biochemistry*, edited by Buddolla V. Academic Press.

Chavez Montes R. A., P. Ranocha, Y. Martinez, Z. Minic, L. Jouanin, *et al.*, 2008 Cell wall modifications in Arabidopsis plants with altered α-L-Arabinofuranosidase Activity. Plant Cell Physiol. 147: 63–77. https://doi.org/10.1104/pp.107.110023

Chin D. P., I. Shiratori, A. Shimizu, K. Kato, M. Mii, *et al.*, 2018 Generation of brilliant green fluorescent petunia plants by using a new and potent fluorescent protein transgene. Sci. Rep. 8: 1–10. https://doi.org/10.1038/s41598-018-34837-2

Coleman H. D., A. M. Brunner, and C.-J. Tsai, 2021 Synergies and entanglement in secondary cell wall development and abiotic stress response in trees. Front. Plant Sci. 12. https://doi.org/10.3389/fpls.2021.639769

Dahlman J. E., O. O. Abudayyeh, J. Joung, J. S. Gootenberg, F. Zhang, *et al.*, 2015 Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. Nat. Biotechnol. 33: 1159–1161. https://doi.org/10.1038/nbt.3390

Dai Y., G. Hu, A. Dupas, L. Medina, N. Blandels, *et al.*, 2020 Implementing the CRISPR/CAS9 technology in eucalyptus hairy roots using wood-related genes. Int. J. Mol. Sci. 21. https://doi.org/10.3390/ijms21103408

Ding X., T. Seebeck, Y. Feng, Y. Jiang, G. D. Davis, *et al.*, 2019 Improving CRISPR-Cas9 Genome Editing Efficiency by Fusion with Chromatin-Modulating Peptides. Cris. J. 2: 51–63. https://doi.org/10.1089/crispr.2018.0036

Dunkley T. P. J., S. Hester, I. P. Shadforth, J. Runions, T. Weimar, *et al.*, 2006 Mapping the *Arabidopsis* organelle proteome. Proc. Natl. Acad. Sci. 103: 6518–6523. https://doi.org/10.1073/pnas.0506958103

Endo S., Y. Iwai, and H. Fukuda, 2019 Cargo-dependent and cell wall-associated xylem transport in *Arabidopsis*. New Phytol. 222: 159–170. https://doi.org/10.1111/nph.15540

Fan C., S. Feng, J. Huang, Y. Wang, L. Wu, *et al.*, 2017 *AtCesA8*-driven *OsSUS3* expression leads to largely enhanced biomass saccharification and lodging resistance by distinctively altering lignocellulose features in rice. Biotechnol. Biofuels 10: 1–12. https://doi.org/10.1186/s13068-017-0911-0

Fu Y., J. D. Sander, D. Reyon, V. M. Cascio, and J. K. Joung, 2014 Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. 32: 279–284. https://doi.org/10.1038/nbt.2808

Fulton L. M., and C. S. Cobbett, 2003 Two α-L-arabinofuranosidase genes in *Arabidopsis thaliana* are differentially expressed during vegetative growth and flower development. J. Exp. Bot. 54: 2467–2477. https://doi.org/10.1093/jxb/erg269

Gao C., K. Y. Li, S. Qu, M. W. Y. San, C. K. Y. Yu, *et al.*, 2012 The Golgi-localized *Arabidopsis* endomembrane protein12 contains both endoplasmic reticulum export and Golgi retention signals at its c terminus. Plant Cell 24: 2086–2104. https://doi.org/10.1105/tpc.112.096057

Goujon T., Z. Minic, A. El Amrani, O. Lerouxel, E. Aletti, *et al.*, 2003 *AtBXL1*, a novel higher plant (*Arabidopsis thaliana*) putative *BETA-XYLOSIDASE* gene, is involved in secondary cell wall metabolism and plant development. Plant J. 33: 677–690. https://doi.org/10.1046/j.1365-313X.2003.01654.x

Hegelund J. N., T. P. Jahn, L. Baekgaard, M. G. Palmgren, and J. K. Schjoerring, 2010 Transmembrane nine proteins in yeast and *Arabidopsis* affect cellular metal contents without changing vacuolar morphology. Physiol. Plant. 140: 355–367. https://doi.org/10.1111/j.1399-3054.2010.01404.x

Himelblau E., H. Mira, S. J. Lin, V. C. Culotta, L. Peñarrubia, *et al.*, 1998 Identification of a functional homolog of the yeast copper homeostasis gene *ATX1* from *Arabidopsis*. Plant Physiol. 117: 1227–1234. https://doi.org/10.1104/pp.117.4.1227

Himmel M. E., S. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, *et al.*, 2007 Biomass recalcitrance: engineering plants and enzymes for biofuels production. Nature 454: 804–807. https://doi.org/10.1126/science.1137016

Jeknic Z., S. P. Lee, J. Davis, R. C. Ernst, and T. H. H. Chen, 1999 Genetic transformation of *Iris germanica* mediated by *Agrobacterium tumefaciens*. J. Am. Soc. Hortic. Sci. 124: 575–580. https://doi.org/10.21273/jashs.124.6.575

Kiani S., A. Chavez, M. Tuttle, R. N. Hall, R. Chari, *et al.*, 2015 Cas9 gRNA engineering for genome editing, activation and repression. Nat. Methods 12: 1051–1054. https://doi.org/10.1038/nmeth.3580

Lee C., Q. Teng, R. Zhong, and Z. H. Ye, 2012 Arabidopsis GUX proteins are glucuronyltransferases responsible for the addition of glucuronic acid side chains onto xylan. Plant Cell Physiol. 53: 1204–1216. https://doi.org/10.1093/pcp/pcs064

Matsuo N., M. Makino, and H. Banno, 2011 *Arabidopsis ENHANCER OF SHOOT REGENERATION (ESR)1* and *ESR2* regulate in vitro shoot regeneration and their expressions are differentially regulated. Plant Sci. 181: 39–46. https://doi.org/10.1016/j.plantsci.2011.03.007

Mortimer J. C., G. P. Miles, D. M. Brown, Z. Zhang, M. P. Segura, *et al.*, 2010 Absence of branches from xylan in *Arabidopsis gux* mutants reveals potential for simplification of lignocellulosic biomass. Proc. Natl. Acad. Sci. 107: 17409–17414. https://doi.org/10.1073/pnas.1005456107

Nikolovski N., D. Rubtsov, M. P. Segura, G. P. Miles, T. J. Stevens, *et al.*, 2012 Putative glycosyltransferases and other plant Golgi apparatus proteins are revealed by LOPIT proteomics. Plant Physiol. 160: 1037–1051. https://doi.org/10.1104/pp.112.204263

Nikolovski N., P. V. Shliaha, L. Gatto, P. Dupree, and K. S. Lilley, 2014 Label-free protein quantification for plant Golgi protein localization and abundance. Plant Physiol. 166: 1033–1043. https://doi.org/10.1104/pp.114.245589

Nonaka S., T. Someya, Y. Kadota, K. Nakamura, and H. Ezura, 2019 Super-*Agrobacterium* ver. 4: Improving the transformation frequencies and genetic engineering possibilities for crop plants. Front. Plant Sci. 10: 1–12. https://doi.org/10.3389/fpls.2019.01204

Nuñez J. K., J. Chen, G. C. Pommier, J. Z. Cogan, J. M. Replogle, *et al.*, 2021 Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. Cell 1–17. https://doi.org/10.1016/j.cell.2021.03.025

Oliveira Y. de, L. G. Adamuchio, C. de Oliveira, J. Degenhardt-Goldbach, I. Gerhardt, *et al.*, 2011 Indirect organogenesis from leaf explants of *Eucalyptus benthamii* x *Eucalyptus dunnii* and shoot multiplication. BMC Proc. 5: P145. https://doi.org/10.1186/1753-6561-5-s7-p145

Petrik D. L., T. Tryfona, P. Dupree, and C. T. Anderson, 2020 *BdGT43B2* functions in xylan biosynthesis and is essential for seedling survival in *Brachypodium distachyon*. Plant Direct 4: 1–16. https://doi.org/10.1002/pld3.216

Plasencia A., M. Soler, A. Dupas, N. Ladouce, G. Silva-Martins, *et al.*, 2016 *Eucalyptus* hairy roots, a fast, efficient and versatile tool to explore function and expression of genes involved in wood formation. Plant Biotechnol. J. 14: 1381–1393. https://doi.org/10.1111/pbi.12502

Sood P., A. Bhattacharya, and A. Sood, 2011 Problems and possibilities of monocot transformation. Biol. Plant. 55: 1–15. https://doi.org/10.1007/s10535-011-0001-2

Suzuki S., K. Supaibulwatana, M. Mii, and M. Nakano, 2001 Production of transgenic plants of the Liliaceous ornamental plant *Agapanthus praecox* ssp. orientalis (Leighton) Leighton via *Agrobacterium*-mediated transformation of embryogenic calli. Plant Sci. 161: 89–97. https://doi.org/10.1016/S0168-9452(01)00393-4

Suzuki M., A. Kato, N. Nagata, and Y. Komeda, 2002 A xylanase, *AtXyn1*, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the *Arabidopsis thaliana* genome. Plant Cell Physiol. 43: 759–767. https://doi.org/10.1093/pcp/pcf088

Wang C. S., S. W. Hsu, and Y. F. Hsu, 2013 New insights into desiccation-associated gene regulation by lilium longiflorum asr during pollen maturation and in transgenic *Arabidopsis*. Elsevier. 301: 37-94. https://doi.org/10.1016/b978-0-12-407704-1.00002-6

Wierzbicki M. P., N. Christie, D. Pinard, S. D. Mansfield, E. Mizrachi, *et al.*, 2019 A systems genetics analysis in *Eucalyptus* reveals coordination of metabolic pathways associated with xylan modification in wood-forming tissues. New Phytol. 223: 1952–1972. https://doi.org/10.1111/nph.15972

Wolabu T. W., J. J. Park, M. Chen, L. Cong, Y. Ge, *et al.*, 2020 Improving the genome editing efficiency of CRISPR/Cas9 in *Arabidopsis* and *Medicago truncatula*. Planta 252: 1–14. https://doi.org/10.1007/s00425-020-03415-0

Xiong G., K. Cheng, and M. Pauly, 2013 Xylan O-acetylation impacts xylem development and enzymatic recalcitrance as indicated by the arabidopsis mutant tbl29. Mol. Plant 6: 1373–1375. https://doi.org/10.1093/mp/sst014

Yordanov Y. S., C. Ma, S. H. Strauss, and V. B. Busov, 2014 *EARLY BUD-BREAK 1* (*EBB1*) is a regulator of release from seasonal dormancy in poplar trees. Proc. Natl. Acad. Sci. 111: 10001–10006. https://doi.org/10.1073/pnas.1405621111

Zeng Y., M. E. Himmel, and S. Y. Ding, 2017 Visualizing chemical functionality in plant cell walls. Biotechnol. Biofuels 10: 1–16. https://doi.org/10.1186/s13068-017-0953-3

Zhao Z. Y., T. Cai, L. Tagliani, M. Miller, N. Wang, *et al.*, 2000 *Agrobacterium*-mediated sorghum transformation. Plant Mol. Biol. 44: 789–798. https://doi.org/10.1023/A:1026507517182

Zhong R., Q. Teng, M. Haghighat, Y. Yuan, S. T. Furey, et al., 2017 Cytosol-localized UDP-xylose synthases provide the major source of UDP-xylose for the biosynthesis of Xylan and xyloglucan. Plant Cell Physiol. 58: 156–174. https://doi.org/10.1093/pcp/pcw179

4.8 Figures

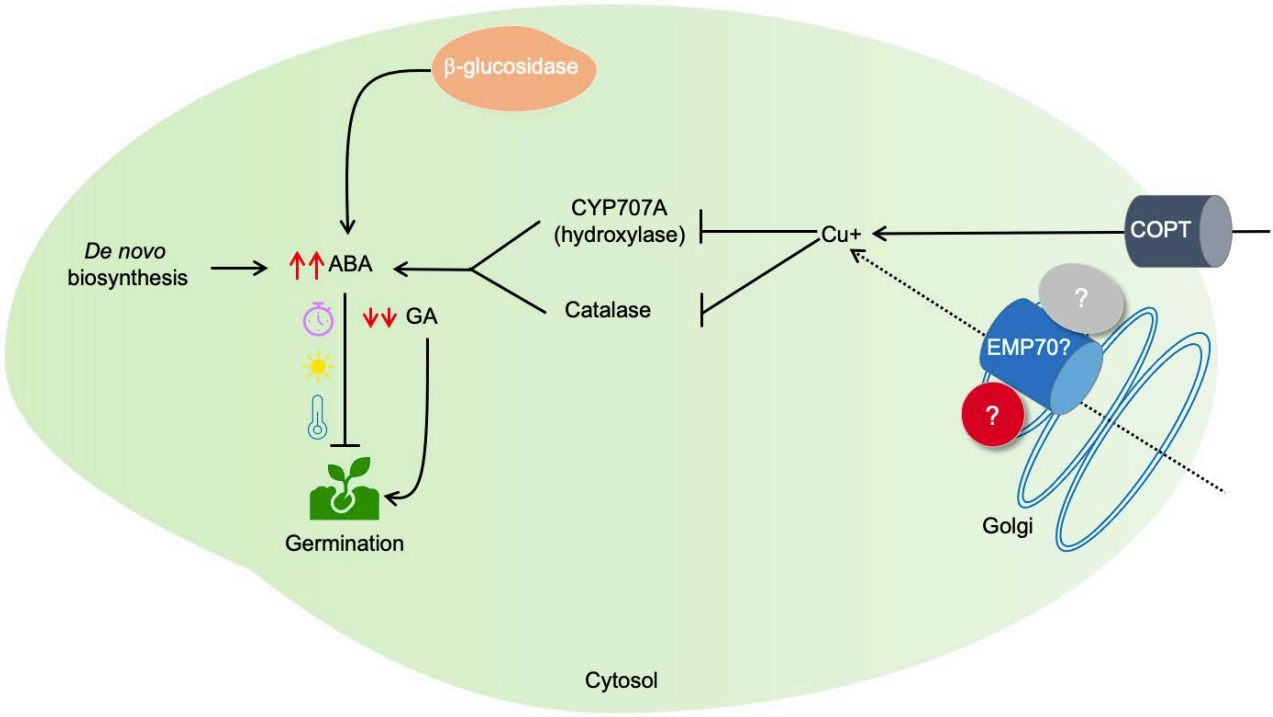


Figure 4.1 Proposed model for the role of EMP70 in seed germination.

Cu is acquired into plant cells through high-affinity Cu transporters such as members of the COPT protein family. EMP70 family may likely play a similar role, although such a role may require other unknown EMP70- and Golgi-associated proteins. ABA is generated through various means including hydrolysis of conjugated ABA by β-glucosidase, and through *de novo* biosynthesis. In normal physiological conditions, the net amount of Cu acquired through COPT and EMP70 transporters and other means allow normal regulation of ABA concentration through inhibition of ABA catabolism by hydroxylase (CYP707A) and catalase. The normally regulated concentration of ABA inhibits seed germination, however, at the appropriate time, and under specific environmental cues, GA counteracts the activity of ABA and promotes seed germination. When EMP70 is absent (such as in *emp70* knock-out mutants), the overall concentration of Cu will be below the threshold required for normal regulation of ABA, significantly lower levels of Cu may lessen the inhibitory effect on CYP707A and catalase, allowing ABA concentrations to decrease significantly, thereby preventing seed dormancy while encouraging early seed germination through GA. While there are several pathways involved in the uptake of Cu into the cell, the EMP70-acquired Cu may be specifically important in the regulation of seed germination through ABA and GA.