HYPER-LIGNIFIED ROOT SYSTEMS AS A CARBON SINK IN ARABIDOPSIS THALIANA

by

ADRIENNE JULIANA NYE

B. Sc., The University of Victoria, 2005

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Plant Science)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2009

Abstract

Lignified plant cell walls represent an immense carbon sink to offset rising atmospheric carbon dioxide (CO_2) due to the chemical stability and structural diversity of the bonds formed between lignin subunits, making it the slowest decomposing component of dead vegetation. This thesis explores the feasibility of increasing lignin deposition in roots through overexpression of SND1 (Secondary Wall Associated NAC Domain Protein 1), a key transcriptional activator involved in regulating secondary wall biosynthesis in fibres, under the control of two different putative root-specific promoters, 4-coumarate:CoA ligase 1 (4CL1) and glutathione S-transferase-*tau* class 19 (GSTU19). Transgenic plants were analyzed at: 1) the molecular level (transcription of lignin pathway genes and regulatory transcription factors (TFs) involved in cell wall biosynthesis), 2) the chemical level (lignin content) and 3) the plant growth and development level (phenotyping and microscopy). Results showed that: i) *SND1* was overexpressed in a tissue-specific manner in roots, ii) *SND1* overexpression caused an upregulation of its previously known direct downstream targets, iii) *SND1* overexpression did not result in a modification of indicative lignin biosynthetic pathway genes in roots, iv) plants overexpressing *SND1* in roots generally produced plants with decreased total lignin content, v) plants overexpressing *SND1* in roots generally showed an increase in lateral root density, and vi) seed traits, plant growth and development, plant height and lignin deposition patterns in roots remained unaltered. Misregulation of SND1 in roots did not result in the predicted increase in lignin deposition patterns in this organ.

Table of Contents

List of Tables

List of Figures

Abbreviations

4CL	4-coumarate:CoA ligase
4CL1	4-coumarate:CoA ligase 1
6xHis	hexameric histidine tag
ANAC012	Arabidopsis NAC domain containing protein 12
ARF	auxin response factor
ASF-1	nuclear activating sequence-1-binding factor
ATAF 1/2	arabidopsis transcription activation factor
bp	base pair
C	carbon
C3H	4-coumarate-3-hydroxylase
C4H	cinnamate-4-hydroxylase
CAD	cinnamyl alcohol dehydrogenase
CCR	cinnamoyl-CoA-reductase
CCoAOMT	caffeoyl-CoA 3-O-methyltransferase
cDNA	complementary deoxyribonucleic acid
CDS	coding sequence
CO_2	carbon dioxide
CoA	coenzyme A
COMT	caffeic acid/5-hydroxyconiferaldehyde O-methyltransferase
CUC2	cup-shaped cotyledon 2
DNA	deoxyribonucleic acid
F5H	ferulate-5-hydroxylase
G	guaiacyl (lignin)
GHGs	greenhouse gases
GPI	glycosylphosphatidylinositol
GSH	glutathione
GSTU19	glutathione S-transferase-*tau* class 19
Gt	gigatonnes
GUS	beta-glucuronidase
H	hydroxyphenyl (lignin)
HCT	p-hydroxycinnamoyl-CoA: quinate/shikimate 4-hydroxycinnamoyltransferase
KNAT	Knotted1-like TALE homeodomain protein
LB	Luria-Bertani
LRD	lateral root density
LRP	lateral root primordium
Mya	million years ago
mRNA	messenger RNA
MS	Murashige and Skoog
MYB	v-myb myeloblastosis viral oncogene homolog (avian)
NAC	NAM, ATAF1/2 and CUC2
NAM	no apical meristem
NST1	NAC secondary wall thickening promoting factor 1
NST2	NAC secondary wall thickening promoting factor 2
NST3	NAC secondary wall thickening promoting factor 3
OD	optical density
ORF	open reading frame
PAL	phenylalanine ammonia-lyase
PCR	polymerase chain reaction
RNA	ribonucleic acid
RT	reverse transcription
S	syringyl (lignin)
SND1	Secondary Wall Associated NAC Domain Protein 1
SOC	soil organic carbon
SIC	soil inorganic carbon
SURE	sulfur-responsive element
TF	transcription factor
UGT	UDP-glucosyltransferase
VND6	Vascular-related NAC-Domain 6
VND7	Vascular-related NAC-Domain 7

Acknowledgements

Research and material contributions

I would like to begin by expressing my sincerest gratitude to my supervisor, Dr. Brian Ellis, who has proven to be an exceptional mentor in guiding me through my master's degree. His warm and approachable nature, along with his exceedingly brilliant contributions, thoughtful insights and endless encouragement throughout my research, has inspired me to grow and develop on both a personal and academic level. This work would not have been possible without the guidance, ideas, suggestions and support offered by my committee members, Dr. Leonard Foster and Dr. James Kronstad, whose time and energy devoted to improving my research project is deeply appreciated. I acknowledge Dr. Patrick von Aderkas and the late Yousuf Ebrahim for recognizing my potential and encouraging me to pursue this master's degree. In addition, I would like to thank Margaret Ellis, our lovely lab manager, for her excellent ordering skills, organizational prowess and commitment to ensuring that lab affairs run smoothly. I am indebted to my student and lab colleagues, Hardy Hall, QingNing Zeng, Apurva Bhargava, Ankit Walia, Doris Vong, Jia Cheng, JinSuk Lee, Earl Alikpala, and Dr. Jun Chen for their endless patience, moral support and generous friendship offered while teaching me new techniques, answering my many questions and allowing me above all to be myself. Their contribution to stimulating conversations over coffee and shared meals, company during long hours poured over gels and petri dishes, priceless involvement in several lab adventures and eventful musical collaborations, made my time pass by here in somewhat of a bliss, providing me with memories to last a lifetime. I would

like to recognize the following people for their invaluable contributions to helping me complete my thesis research: Jim Guo of the Chen Lab for devoting his time to teach me how to use the microscope, Zorica Kotur of the Glass Lab for her help and guidance with the hydroponics, Julia Nowak of the Cronk Lab for aiding me with the hand sectioning, staining and microscopy of my transgenic plant lines, Dr. Sarah McKim of the Haughn Lab for assistance with the wax embedding protocol, microtome sectioning and staining of my transgenic plant lines, Lifang Zhao for providing me with the pPZP211 empty vector seed lines, Xue Feng Chang (Harry) of the Beatson Lab at BCIT for his immense involvement in the acetyl bromide lignin analysis, Vicki Maloney of the Mansfield Lab for her amazing contribution to the Starch and Klason lignin analyses, and to Paul Lee (undergraduate in the Ellis Lab) for his enthusiasm and painstaking commitment to helping me with seed phenotyping, germination assays and RNA extractions in the summer of 2009.

Support and guidance

A very heartfelt thanks goes out to my Mom and Dad for their immeasurable love, generosity, encouragement and financial support throughout my life. Their selfless acts of kindness and tireless ability to provide me with whatever I need, constantly remind me to remain humble and be grateful for everything that is given to me. Without providing me with the safe and loving environment needed for growth and self-discovery as I was growing up to the hour-long phone conversations and visits home throughout my degree, I would not be who I am today and for that am forever in their debt. I feel honoured to share my genes (and my birth date) with my sister, Gabrielle, whose love and support always seem to show up at exactly the right time.

Her belief in me and my ability to succeed throughout my graduate studies have given me the strength to pull through to the end. I would like to convey a very special message to all my friends, who have made my time here both memorable and delectably fun. To the ones who were there through all the ups and downs, ins and outs, who listened when I needed to vent, who lifted me up when I was down and who provided me with the courage I needed to deal with everyday life, I raise my glass to you!

Finally, I would like to dedicate this thesis to my late grandmother, Anyu, for always believing in me but more importantly showing me, by example, what it means to be a true Survivor. I am deeply grateful for the contribution she has made to my life as a caregiver, friend and kindred spirit. Her life has inspired me in more ways than I can express; she was the embodiment of strength, self-empowerment and above all grace. Her love, wisdom and spirituality have always taught me to be true to myself, a theme that has become the driving force behind all that I do, all that I am and is ultimately what propels me forward into the future, with no regrets.

1. Introduction

1.1 Global climate change and mitigating global carbon emissions

Our planet is habitable due to its proximity to the sun and to the layer of gases surrounding it, which we have come to know as our atmosphere (Karl & Trenberth, 2003). This natural greenhouse effect results from the presence of a particular combination of atmospheric gases including nitrogen, oxygen, argon, carbon dioxide (CO_2) and other trace gases (Karl & Trenberth 2003). These gases act as a sort of insulating blanket, trapping solar energy as heat and regulating average global surface temperatures within a range suitable for life to evolve (Li *et al.* 2009).

Life on Earth as we know it relies greatly on the balance of these important greenhouse gases (GHGs). Through both human endeavors and natural fluxes through the Earth system, it is now acknowledged that atmospheric CO_2 levels have increased about 35% since the early ages of industrialization (Karl & Trenberth 2003; Millard *et al.* 2007). What's more, is that roughly half of the CO_2 released since the Industrial Revolution remains in the atmosphere while the other half has been sequestered in the ocean as well as terrestrial ecosystems (Karl & Trenberth 2003; Millard *et al.* 2007; Raven & Karley 2006). It has been hypothesized that one result of this increase in GHGs (atmospheric CO_2, methane and nitrous oxide) is a pattern of climate change phenomena that includes but is not limited to: rises in global surface temperatures, increased occurrence of extreme weather events, increased incidence and intensity of wild fires, shifting of ecosystems, rising sea

levels and changes in disease transmission dynamics (Lal 2008). In addition, anthropogenic activities such as deforestation, fossil fuel combustion, altered land use through urbanization, wetland draining, soil cultivation and biomass burning have also had a large-scale impact on terrestrial surface characteristics and, by extension, on climate change (Lal 2008; Karl & Trenberth, 2003). As a result, there is growing interest in stabilizing increases in GHGs with the goal of mitigating the risks associated with global climate change (Lal 2008). An introduction to all of the mitigation strategies proposed for lowering GHG emissions is beyond the scope of this thesis; therefore, only those related to carbon (CO_2) emissions will be mentioned. The reason being that of all the GHGs contributing to global climate change, CO_2, next to water vapour, is considered one of the most important (Malhi *et al.* 2002).

There are currently three broad categories of mitigation strategies and practices for lowering CO_2 emissions: (i) to reduce global energy use, (ii) to reduce emissions, and (iii) to enhance removal via carbon sequestration (Lal 2008; Smith *et al.* 2008). The last of these approaches has garnered the attention of both scientists and politicians as an effective strategy for mitigating GHG emissions (Mondini & Sequi 2008). As a result of this interest, four main types of carbon sequestration have been proposed: ocean storage, geological storage, biomass storage and mineral carbonation (Oelkers & Cole 2008). In conclusion, carbon sequestration is by no means instantaneous and consideration must be given to the fact that strategies, practices and techniques take time to develop and implement.

1.1.1 The global carbon cycle and the role of plants as terrestrial carbon sinks

In order to develop workable strategies for mitigating global climate change we must first understand how the global carbon (C) cycle works (Lal 2008). The global C cycle is typically thought of as an interconnected flow of C through four principal reservoirs: the terrestrial biosphere, the oceans, fossil carbon and the atmosphere (Schimel 1995; Oelkers & Cole, 2008). In the terrestrial biosphere, a number of organisms (cyanobacteria, green algae and land plants) have specialized mechanisms that allow for absorption of CO_2 into their cells. With the addition of water and energy from solar radiation, they use photosynthesis to chemically convert CO_2 to carbohydrates (Black 1973). Conversely, CO_2 and energy can be released from terrestrial ecosystems by the process of respiration. This involves the metabolic breakdown of C-based organic molecules primarily into gaseous CO_2, among other byproducts. Every year, respiration returns almost half of the CO_2 that is absorbed by photosynthesis to the atmosphere (Falkowski *et al.* 2000; Black 1973). The movement of atmospheric C through photosynthesis, respiration, and back to the atmosphere is considerable, and this flux produces notable annual fluctuations in atmospheric CO_2 concentrations (Falkowski *et al.* 2000).

Photosynthetic organisms play a significant role in the global C cycle and over time, significant amounts of C can be stored or released from terrestrial biomes (Schimel 1995). For example, changes in land use can greatly contribute to carbon source/sink dynamics, as demonstrated by the accumulation of C in the living

tissues of new plant growth and within the soil of regenerating forests following the abandonment of agricultural lands, causing a net decrease in contributions to atmospheric CO_2 concentrations (Schimel 1995). In other words, as atmospheric CO_2 increases, terrestrial plants can become a potential sink for anthropogenic carbon (Falkowski *et al.* 2000). These net returns and losses (or fluxes) of C between the four previously mentioned C reservoirs are known as the global C budget (Schimel 1995). Although terrestrial ecosystems have the potential to mitigate rising atmospheric CO_2 levels in the coming decades, there is still considerable uncertainty surrounding how these ecosystems will respond to the combined effects of higher CO_2 concentrations, higher temperatures, and changes in soil dynamics (Falkowski *et al.* 2000). In order to predict how these sources and sinks will behave in the future, it is crucial that we enhance our understanding of how plants will respond to the foreseeable increases in human-derived CO_2 emissions (Lal 2008; Raven & Karley 2006).

1.1.2 Increasing agricultural soil carbon stocks through carbon sequestration

Since ratification of the Kyoto Protocol, most C mitigation strategies have focused on the use of C sinks (natural or manmade carbon reservoirs) as a form of C offset, and this focus has increased general awareness of carbon sink significance (Lal 2008). In attempting to balance the global C budget, future economic growth would be based on a 'carbon neutral' strategy, or rather a 'no net increase' in atmospheric C (Lal 2008). Therefore, of interest to those involved with the Kyoto Protocol is any mitigation practice that increases C input via photosynthesis or slows the return of stored C via respiration or fire (Smith *et al.* 2008). These strategies consequently

offer the potential to 'sequester' C or build C 'sinks' and will likely be a focal point for future approaches to mitigating climate change (Smith *et al.* 2008). In short, carbon sequestration is the process by which the terrestrial C sink generates a net removal of CO_2 from the atmosphere.

The capture, transport and final deposition of carbon, via carbon sequestration, are largely dependent upon a complex set of biochemical and chemical processes (Oelkers & Cole 2008). More specifically, sequestering C represents a metabolic dead end, inhibiting its reusability by terminating its physiological activity (Millard *et al*. 2007), and more securely storing it in other more long-lived C reservoirs (Lal 2004).

In the face of increasing carbon emissions, particular emphasis is being placed on this process of carbon sequestration (Lal 2008). For example, roughly a third of the terrestrial land surface is dominated by agricultural lands (crops or planted pastures) whose soils are capable of acting as either carbon sources or sinks (Smith *et al.* 2008). Of the ~2500 gigatons (Gt) of worldwide soil C, there is roughly 1550 Gt of soil organic carbon (SOC) and 950 Gt of soil inorganic carbon (carbonates) (Lal 2004). Furthermore, the global soil C pool is considerable compared to the atmospheric C pool of 760 Gt and the biotic pool of 560 Gt, (Lal 2004). However, soil carbon sequestration is a trickier long-term strategy for climate mitigation as opposed to reducing carbon emissions, given that it could be difficult to measure and verify the amount of carbon sequestered below ground (Mondini & Sequi 2008).

Nevertheless, strategies to improve soil carbon stocks are appealing as part of an integrated sustainability approach since enhanced agricultural management often brings with it an array of other desirable environmental and economic outcomes in addition to mitigating climate change (Mondini & Sequi 2008; Smith and Falloon, 2005). Lal (2008) summarizes these soil C sequestration benefits as including enhanced soil quality, improved soil productivity, decreased risk of soil erosion and sedimentation, and reduced water contamination and eutrophication. These potential outcomes also demonstrate that soil C sequestration could represent an approach to attain food security (Johnson *et al.* 2007). In addition to both environmental and economic benefits, C sequestration is attractive for one another reason: it is likely to be the most cost-effective and feasible method to lower atmospheric CO_2 levels within the first 20–30 years that it is implemented, thus effectively buying time while other technologies aimed directly at reducing GHG emissions are developed (Mondini & Sequi 2008). However, yearly increases in SOC can only be sustained perhaps for 50–100 years, at which point increases in SOC are likely to slow and ultimately cease as the soil reaches a new equilibrium. This emphasizes the point that C sequestration may even be a reversible process if suitable soil management practices are not maintained (Lal 2004; Mondini & Sequi 2008).

Given the sizeable amount of global carbon contained within agricultural soils, it is not surprising that the possibility of partially offsetting fossil-fuel emissions by sequestering excess atmospheric C within these soils is now being strongly

advocated (West & Marland 2002). Unfortunately, fossil-fuel emissions over the next 100 years are anticipated to greatly exceed even the maximum amount of carbon that could potentially be sequestered. Therefore, carbon sequestration should simply be seen as a modest contribution to a much larger mitigation plan and not as a replacement for the development of new energy supplies, improved energy use strategies and technological innovations required to stabilize concentrations of atmospheric CO_2 (Malhi *et al*. 2002).

1.1.3 Root-derived soil carbon

In terrestrial plants, the rhizosphere (the soil that immediately surrounds a plant root) encompasses the complex chemical, physical, and biological interactions between roots and their surrounding environment (Bais *et al.* 2006). Plant roots are actively involved in: soil-microbe interactions, the secretion of compounds required for pathogen defense and absorption of soil nutrients. Roots also play a role in protecting above ground tissues from acidic conditions, heavy metals and drought (Koyama *et al.* 2005). Studies have shown that soil C is predominantly composed of root C and that within the organic soil horizons, root-derived soil organic C generally decreases with depth (Jobbágy & Jackson 2001; Rasse *et al.* 2005). In natural ecosystems, root-derived SOC is almost entirely a result of materials released from the roots of natural vegetation or crops during growth, such as root exudates, sloughed off root tips and cells, mucilage and by decomposition of dead roots (Subedi *at al.* 2006). There is still considerable debate over the amount of plant root C that contributes to the total C pool in the terrestrial biosphere. According to Robinson (2007), the best approximation of the root carbon pool is

270–280 Pg of the total terrestrial biome C pool of 650 Pg (Subedi *at al.* 2006). A global root C reservoir this large has implications for land C sinks as a response to a rise in atmospheric CO_2. For instance, excess levels of CO_2 can stimulate photosynthesis leading to an estimated 20% increase in plant production, which in turn could enhance soil C input thus increasing soil C sequestration (De Graaff *et al.* 2007). Moreover, this increase in SOC could thereby counterbalance the rise in atmospheric CO_2 (De Graaff *et al.* 2007). Conversely, an increase in input of SOC due to increased rhizodeposition and root litter can have a profound influence on plant productivity and root growth (Subedi *at al.* 2006). It is worth noting that more in-depth measurement of the impacts of root-derived SOC from crop systems could make invaluable contributions to our study of C dynamics, the global C budget and C sequestration (Subedi *at al.* 2006).

1.1.4 *Arabidopsis thaliana* as a model organism

Arabidopsis thaliana, also known as thale cress or mouse-ear cress, is a small flowering plant widely used as a model organism in plant biology research. *Arabidopsis* is a member of the mustard family (Brassicaceae), which includes many familiar agricultural species such as broccoli, cabbage, turnip, rapeseed, cauliflower, brussels sprouts and radish. *Arabidopsis* itself is not of any major agricultural importance, but it is intensively used as a model organism for studies in genetics and molecular biology and is a close relative of canola, a major transgenic crop in Canadian agriculture. *Arabidopsis* can produce numerous self-progeny in a relatively short time period, and it has very limited growth space requirements, which means that large populations can be easily grown in a greenhouse or indoor

growth chamber. It has a relatively small, genetically tractable and sequenced genome that can be manipulated through genetic engineering more rapidly and easily than any other plant genome (About Arabidopsis 2008; Arabidopsis thaliana 2009).

1.2 Secondary cell walls and the importance of lignin in vascular plant biology

Plant cell walls have many important functions such as, providing mechanical strength, regulating cell expansion and cell cohesion, water conduction and pathogen defense (Knox 2008). The carbon-based polymers, cellulose, hemicellulose, pectin and lignin, are what form the strong, but flexible macromolecular complexes of the cell walls of higher plants (Weng *et al*. 2008). Cellulose, hemicellulose and pectin are the main carbohydrates comprising the growing primary cell wall, while cellulose, xylan, other hemicelluloses and lignin are the major contributors to secondary cell walls (Weng *et al*. 2008). These major cell wall components are variable in their composition and relative abundance, and the final combination in any given tissue often depends on the species, growing site, climate, age and part of the plant (Ko *et al.* 2009).

The composition of cell wall components can be distinguished based on the ground tissues that they are composed of: i.e. parenchyma, collenchyma, and sclerenchyma. Parenchyma and collenchyma cells, which possess primary cell walls, provide structural support in regions of the plant body that are still growing whereas sclerenchyma tissue has both primary cell walls and thickened secondary cell walls.

For example, specialized cells involved in structural support and water conduction, such as fibres, are composed primarily of sclerenchyma (Zhong *et al.* 2006; Burk *et al.* 2001; Rogers *et al.* 2005; Boerjan *et al.* 2003). The ability to resist the forces of gravity and/or tension associated with the pull of the water column due to transpiration (involved in transporting water and solutes over long distances) comes from the evolution of these specialized cells, which provide mechanical support to regions of the plant body that have ceased elongation (Rogers *et al.* 2005; Boerjan *et al.* 2003). A defining feature of these cells is the secondary cell wall, which is formed in a highly coordinated manner by successive encrustation and deposition of the various cell wall constituents (Ko *et al.* 2009). Lignin fills the spaces between cellulose and hemicellulose, where it is covalently linked to the hemicellulose and crosslinked to other plant polysaccharides (Weng *et al*. 2008). The secondary cell wall polysaccharides are highly hydrophilic and are easily permeable to water whereas lignin is more hydrophobic. Lignification of the secondary cell wall thus waterproofs the cell wall and facilitates the transport of water and solutes through the vascular system (Boerjan *et al.* 2003). In summary, lignified secondary cell walls are essential for the function of structurally supportive and conductive xylem tissues.

Cell wall lignification emerged in the plant kingdom about 430 million years ago (Mya) and is considered to be a relatively recent process in the evolution of photosynthetic organisms, which developed approximately 2000 Mya (Boerjan *et al.* 2003). The ability to produce lignin is thought to have been crucial for the

adaptation of aquatic plants to a terrestrial environment where they were likely to face critical new stresses including UV radiation, desiccation and attack by established and diverse communities of soil microbes (Emiliani *et al.* 2009). In fact, deposition of lignin or rather the synthesis of monolignols, has been shown to play an essential role in the assembly of cell wall appositions (CWAs), also known as papillae, which provide a primary means of defense against pathogens that are attempting to penetrate the cell wall (Bhuiyan *et al*. 2009).

The study of phenylpropanoid metabolism (the pathway responsible for the lignin biosynthesis as well as some other important secondary metabolic compounds) has been a central theme in plant biochemistry. In addition to lignin formation, the contributions to plant fitness of many phenylpropanoid pathway intermediates and end products such as antioxidants, ultra-violet protectants, phytoalexins, pigments, aroma compounds and antiherbivory compounds, emphasizes the importance of this metabolic system (Humphreys & Chapple 2002). Moreover, the phenylpropanoid pathway represents an essential and ubiquitous metabolic trait amongst land plants, since it supplies vital compounds such as lignin (essential for vascularization and xylem formation as well as structural support and stem rigidity out of water), and flavonoids (essential for reproductive biology and for protection against UV via pigment accumulation, for deterring microbial attack and for modulating symbiotic plant-microbe interactions by production of anti-microbial compounds such as phytoalexins, and signaling flavonoids) (Emiliani *et al.* 2009).

1.2.1 Lignin biosynthesis

The coordinated expression of numerous genes is required for the biosynthesis, assembly and deposition of both primary and secondary cell wall components, including the determining structural and chemical specificity of lignified secondary walls (Boudet *et al.* 2003). Lignin is a racemic aromatic polymer that results from the oxidative combination of three *p*-hydroxycinnamyl alcohol monomers known as monolignols (p-coumaryl, coniferyl and sinapyl alcohols) whose structure differ only in the number of methoxyl groups present in their aromatic rings (Fig. 1 (A)) (Goujon *et al.* 2003). While lignins tend to be dominated by these three monolignol components, there are several additional monomers that are sometimes found in lignin polymers. Many naturally occuring plant species contain lignins derived in part from these other monomers, in addition to trace amounts of units formed from incomplete or secondary reactions that occur during monolignol biosynthesis (Boerjan *et al.* 2003).

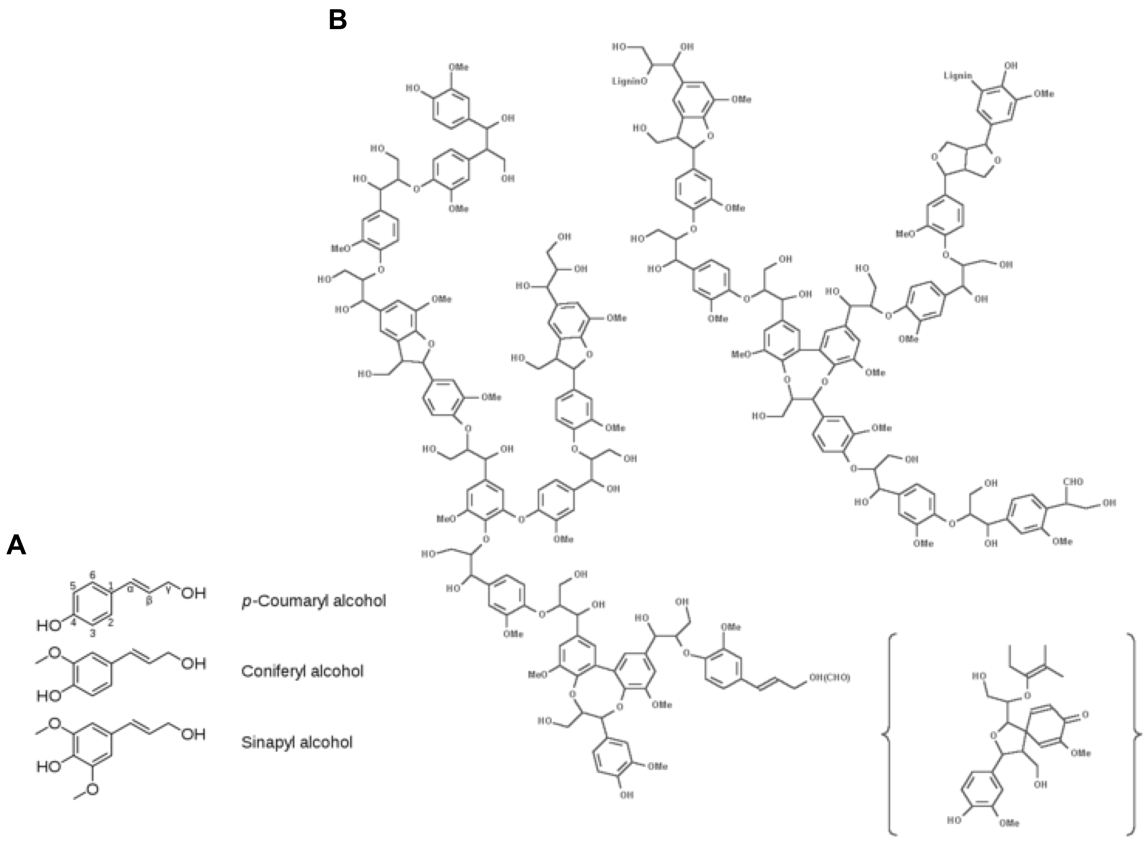


Figure 1. Molecular structures of the three main monolignols and of a putative lignin polymer. (A) Three traditional lignin precursors (*p*-coumaryl alcohol, coniferyl alcohol, sinapyl alcohol) (Monolignol 2008) and (B) a hypothetical lignin polymer (What Is Wood? 2009)

Initially, carbon flux is redirected from primary metabolism to phenylpropanoid biosynthesis through three enzyme-catalyzed reactions (PAL, C4H and 4CL; Fig. 2) which transform L-phenylalanine into *p*-coumaroyl CoA. The latter serves as the entry-point for the two main downstream branch pathways, monolignol and flavonoid biosynthesis (Ferrer *et al.* 2008). The synthesis of monolignols involves consecutive hydroxylations of the aromatic ring, phenolic O-methylation and side-chain carboxyl conversion to an alcohol group ultimately forming the p-coumaryl, coniferyl and sinapyl alcohols (Boerjan *et al.* 2003; Boudet A.-M. 2000). These

monolignols respectively give rise to *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) lignin residues within the lignin polymer (Fig. 2) (Grima-Pettenati & Goffner 1999; Vanholme *et al.* 2008). To produce the final intricate and interconnected lignin complex (Fig. 1 (B)), the monomeric residues are exported to the extracellular space (apoplast) where oxidative enzymes catalyze the formation of free radical derivativs of the monomers. The radicals are then coupled to the growing lignin polymer forming either carbon–carbon or ether bonds (Boudet A.-M. 2000; Grima-Pettenati & Goffner 1999; Vanholme *et al.* 2008).

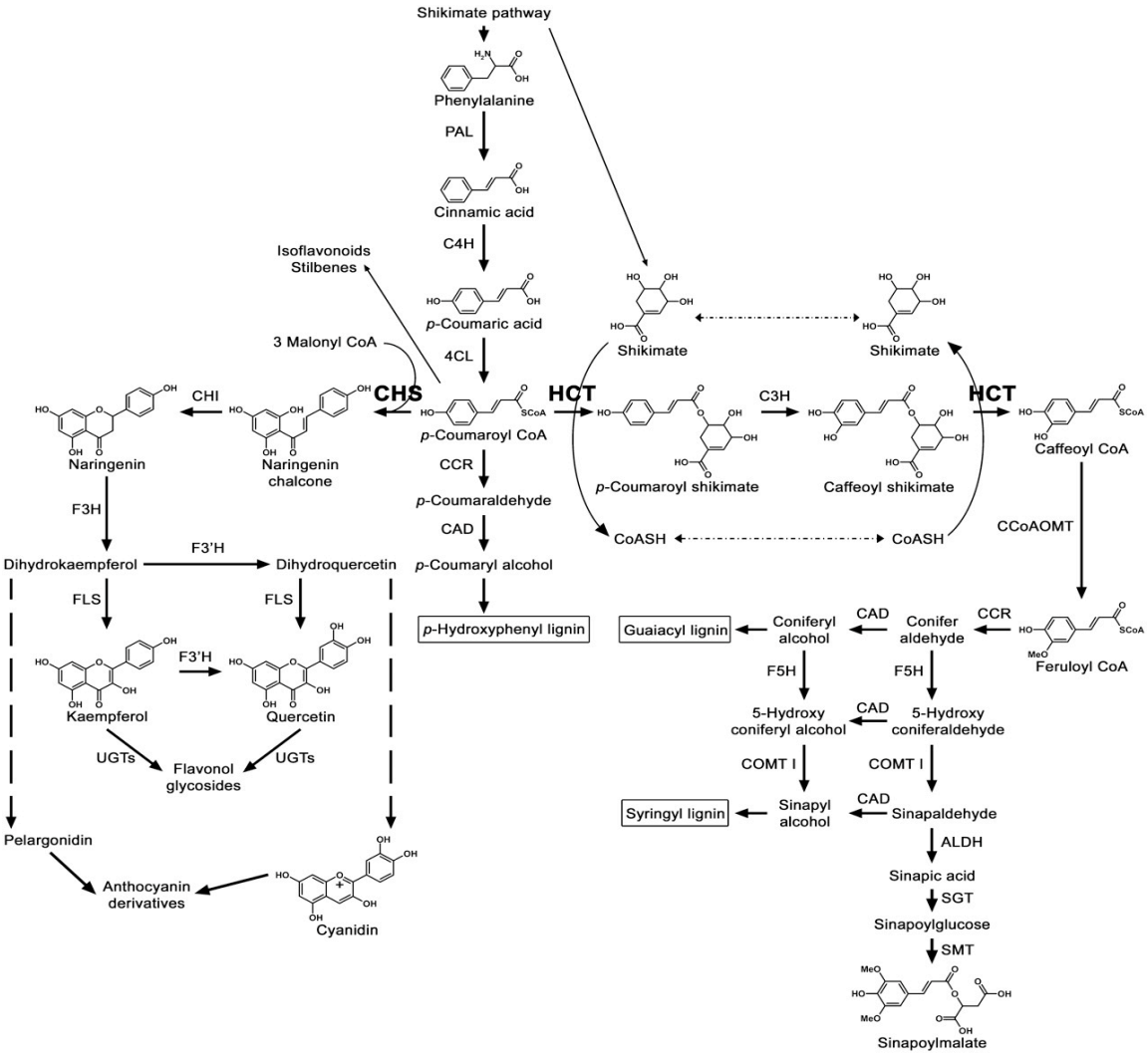


Figure 2. The Phenylpropanoid Pathway. PAL, phenylalanine ammonia-lyase; C4H, cinnamate-4-hydroxylase; C3H, 4-coumarate-3-hydroxylase; COMT, caffeic acid 3-O-methyltransferase; CCoAOMT, caffeic acid/5-hydroxyconiferaldehyde O-methyltransferase; F5H, ferulate-5-hydroxylase; 4CL, p-coumaroyl:CoA ligase; HCT, p-hydroxycinnamoyl-CoA: quinate shikimate p-hydroxycinnamoyltransferase; CCR, cinnamoyl-CoA-reductase; CAD, cinnamyl alcohol dehydrogenase; UGT, UDP-glucosyltransferase. (Besseau *et al.* 2007 (Figure); Vanholme *et al.* 2008 (Caption))

1.2.2 Lignin as a carbon sink

In addition to their many important biological functions, lignified plant cell walls represent a large proportion of plant biomass in the terrestrial biosphere and thus an immense carbon sink (Boudet *et al*. 2003). Next only to cellulose, lignin is the second most abundant biopolymer on earth (Boudet *et al.* 2003; Grima-Pettenati & Goffner 1999; Humphreys & Chapple 2002). Over $1.4x10^{12}$ kg of C is sequestered in terrestrial plant material each year (Battle *et al.* 2000) with lignin constituting about 30% of that total (Humphreys & Chapple 2002). Research interest in lignin biosynthesis and lignin deposition has been motivated by the multiple roles played by lignin in plant biology, including management of abiotic and biotic stress, water conduction, cell differentiation, and carbon partitioning, all of which have both industrial and agricultural importance (Boudet *et al.* 2003; Humphreys & Chapple 2002). An important aspect of lignin that impacts lignocellulosic biomass utilization, in both industry and agriculture, stems from the variable and stable cross-linking of the various cell wall components, which minimizes the accessibility of cellulose and hemicellulose to degradative enzymes (Bhuiyan *et al*. 2009). Not only is the capacity of lignin to resist degradation largely due to its unique polymeric structure, but this structure's distinct arrangement and representation of monomeric units varies widely among species, individuals and even within cell types of the same plant (Weng *et al.* 2008). In essence, the combination of chemical stability and structural diversity of the bonds formed between lignin subunits is sufficient to prevent complete degradation of the polymer by any single enzyme (Weng *et al.* 2008). This stability highlights the potential for lignin to act as a long-

lived C reservoir, and by extension, to serve as a vehicle increased carbon storage and sequestration.

1.2.3 Lignin modification via the monolignol biosynthetic pathway

The past twenty years of research has led to significant insight into lignin biosynthesis, particularly through the use of reverse genetics approaches in which expression of genes encoding individual monolignol and phenylpropanoid pathway enzymes has been altered (Vanholme *et al.* 2008). Generally speaking, in transgenic plants, the downregulation of PAL, C4H, 4CL, HCT, C3H, CCoAOMT, CCR, and, to a smaller degree, CAD, has been shown to have a major influence on lignin content as well as the ratios of H, G and S lignin, although these outcomes are often accompanied by other, undesirable pleiotropic impacts on plant growth, morphology or chemistry (Anterola & Lewis 2002; Vanholme *et al.* 2008).

1.3 Transcription factors as tools for metabolic engineering in plants

Transcriptional regulation is an important mechanism by which metabolic pathways and assembly of cell wall components in plants is controlled (Broun 2004; Zhong & Ye, 2007). Transcription factors (TFs) are regulatory proteins that modify the expression of specific sets of genes by interacting with the transcriptional machinery, including chromatin remodeling proteins and/or other transcription factors involved in transcription through sequence-specific DNA binding and protein–protein interactions (Broun 2004). In other words, these proteins are able to recognize and bind specific sequences in the promoter regions of their target genes, thereby subsequently activating or repressing entire metabolic or

developmental processes. This often occurs by mediation of either an increase or decrease of the encoded mRNA by acting as activators or repressors of gene expression (Broun 2004; Arce *et al.* 2008). The role of transcription factors in coordinated metabolic regulation is of great interest in metabolic engineering because of their ability to control both cellular processes and multiple pathway steps necessary for metabolite accumulation (Broun 2004; Petersen, 2007). Unlike alterations in single-enzyme expression, the use of TFs for metabolic engineering has the potential to generate more complex phenotypes in transgenic plants, as a result of simultaneous modification of different transcriptionally-regulated pathways (Tyo *et al.* 2007).

1.3.1 The role of transcription factors in the regulation and modification of lignin biosynthesis

Lignin synthesis and deposition requires strict spatial and temporal regulation of processes occurring during plant growth and development (Boudet A.-M. 2007). So far, numerous studies suggest that several features of cellular structure and metabolism, such as the cytoskeleton, phosphoinositide signaling, glycosylphosphatidylinositol (GPI)-anchored proteins, hormones, and the supply of sugar nucleotides, must all be integrated as part of the regulation of secondary cell wall biosynthesis and lignin deposition (Zhong & Ye 2007).

Although many of the genes encoding enzymes involved in lignin biosynthesis have been characterized, little is known about the molecular mechanisms underlying the coordinated expression of these genes (Weng *et al*. 2008). However, the study of

global patterns of gene expression by high-throughput technologies has recently revealed some additional features of the various regulatory networks through which this metabolic pathway is controlled (Broun 2004). For example, comparative transcriptome analyses in xylem cells of *Arabidopsis* plants undergoing secondary growth have identified a range of upregulated genes (specifically NAC and MYB TFs) involved in secondary cell wall formation, and these have provided an initial glimpse of the complex networks of TFs controlling this process (Ko *et al.* 2007; Weng *et al.* 2008; Zhong & Ye 2007; Zhong *et al.* 2008). A group of closely related NAC domain proteins in *Arabidopsis thaliana* (Fig. 3), including ANAC043/NST1 (NAC Secondary Wall Thickening Promoting Factor 1), ANAC066/NST2, ANAC012/NST3/SND1 (Secondary Wall Associated NAC Domain Protein 1), VND6 (Vascular-related NAC-Domain 6), and VND7 are now known to be major transcriptional regulators of secondary wall biosynthesis in various supporting cell types in plant tissues that have ceased elongation (Zhong *et al.* 2008).

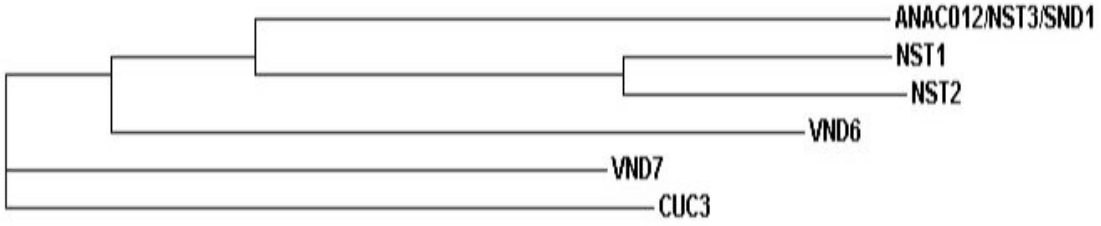


Figure 3. Phylogenetic analysis of five closely related NAC domain proteins in *Arabidopsis thaliana* involved in regulating secondary cell wall biosynthesis in various supporting cell types. The full-length coding sequences (CDS) were aligned using the CLUSTAL W program and the phylogenetic tree was constructed by neighbor-joining methods. The GenBank accession numbers for the used sequences are represented as follows: ANAC012/NST3/SND1 (NM_103011); NST1 (NM_130243); NST2 (NM_116056); VND6 (NM_125632); VND7 (NM_105851) and CUC3 (NM_106292).

SND1 and NST1 are proposed to function in a redundant manner to control development of secondary walls in fibres while VND6 and VND7, respectively, are proposed to regulate metaxylem and protoxylem differentiation in vessels (Zhong *et al.* 2008; Mitsuda *et al.* 2007). In anther endothecium cells, NST1 and NST2 were shown to function redundantly in regulating secondary wall thickening (Mitsuda *et al.* 2005). Overexpression of these NAC genes results in ectopic deposition of secondary walls in cells not normally reinforced with lignin, while inhibition of their functions via dominant repression or knockout results in secondary walls with reduced thickening in the mutant plants (Zhong *et al.* 2008). These secondary wall NACs are proposed to act through a cascade of downstream TFs, which in turn lead to the activation of secondary wall biosynthetic genes including SND2, SND3, MYB20, MYB42, MYB43, MYB46, MYB52, MYB54, MYB58, MYB63, MYB69, MYB85, MYB103, and KNAT7 (a Knotted1-like homeodomain protein), are regulated by SND1 (Zhong *et al.* 2006; Zhong *et al.* 2008; Zhong *et al.* 2007a; Zhong *et al.* 2007b; Zhou *et al.* 2009; Zhong & Ye 2007).

Previous studies by Zhong *et al.* (2006) have shown that *SND1* is expressed specifically in interfascicular fibres and xylary fibres of stems. Constitutive overexpression of *SND1* resulted in activation of the expression of secondary wall biosynthetic genes, leading to massive deposition of secondary walls in cells that are normally not lignified (Zhong *et al.* 2006). An activator is defined in the literature as a DNA-binding protein that regulates one or more genes by increasing the rate of transcription. Ko *et al.* (2007) showed that *SND1* gene expression was localized to

the procambium region of inflorescence stems and roots. They confirmed the function of SND1 as a transcriptional activator but also found that ectopic overexpression of 35S::SND1 plants in Arabidopsis noticeably suppressed secondary wall deposition in the xylary fibre. Moreover, they observed a slight increase in cell-wall thickness in xylem vessels which suggested that SND1 might act as a negative regulator of secondary wall thickening in xylary fibres. In contrast to activators, a negative regulator is defined in the literature as any regulator that acts to prevent transcription or translation. In addition to the elucidation of *SND1* as a major transcriptional activator of secondary wall biosynthesis, Zhong *et al.* (2007b) demonstrated that the *Arabidopsis thaliana* MYB46 transcription factor is a direct target of SND1. They showed that dominant repression of MYB46 caused a severe decrease in the secondary wall thickening of fibres and vessels while overexpression of this gene resulted in the activation of the cellulose, xylan, and lignin biosynthetic pathways, which concurrently led to ectopic deposition of secondary walls in cells not normally lignified. Overexpression of MYB46 caused an upregulation in gene expression among particular genes involved in the synthesis of all three major secondary cell wall components (Weng *et al.* 2008; Zhong *et al.* 2007b; Zhong *et al.* 2008). Furthermore, the expression of two secondary wall–associated transcription factors, MYB85 and KNAT7, was highly upregulated by MYB46 overexpression demonstrating that MYB46 is possibly another major player in the transcriptional network involved in regulating secondary wall biosynthesis in Arabidopsis (Zhong *et al.* 2007b). In addition, Zhou *et al.* (2009) demonstrated that overexpression of MYB58 and MYB63 resulted in specific activation of lignin biosynthetic genes and

simultaneous ectopic deposition of lignin in cells not normally lignified. MYB58 was able to directly activate the expression of lignin biosynthetic genes and a secondary wall–associated laccase (*LAC4*) gene. Furthermore, the SND1 homologs NST1, NST2, VND6, and VND7 as well as the SND1 downstream target, MYB46, were also shown to regulate the expression of MYB58 and MYB63. Their results suggest that MYB58 and MYB63 are transcriptional activators of lignin biosynthesis specifically within the SND1-mediated transcriptional network regulating secondary cell wall formation. Lastly, a recent high-throughput study using whole-transcriptome analyses by Ko *et al.* (2009) provided insight into the regulatory relationship of a group of transcription factors upregulated by MYB46, uncovering a speculative regulatory network with intricate cross communication.

Recently, another study identified a novel CCCH-type zinc finger protein, AtC3H14, as a potential master regulator of secondary wall biosynthesis operating downstream of MYB46 (Ko *et al.* 2009). These studies suggest that SND1, MYB46 and C3H14, act as key regulators of secondary cell wall deposition through their demonstrated ability to turn on the entire cellulose, xylan, and lignin biosynthetic pathways in transgenic plants (Zhong *et al.* 2008). In conclusion, this model of over-arching regulation of secondary cell wall biosynthesis by SND1, MYB46 and C3H14, along with the discovery of other TFs upregulated by these master regulator genes, has provided an initial glimpse into the regulatory networks controlling secondary cell wall formation (Zhong *et al.* 2007a; Zhong *et al.* 2007b; Zhou *et al.* 2009).

As mentioned earlier, the amount of global carbon contained within agricultural soils, offers the potential to partially offset fossil-fuel emissions by sequestering excess atmospheric C in the roots within these soils (West & Marland 2002; Subedi *et al.* 2006). Given the potential for lignin to act as a C sink in below-ground tissues, the recent identification of specific TFs involved in regulating lignin deposition is an important discovery. Single-enzyme modifications that have led to changes in lignin content and/or the ratios of H, G and S lignin (Anterola & Lewis 2002; Vanholme *et al.* 2008) that are generally unsuitable for metabolic engineering in current crop systems, due to their severe pleiotropic phenotypes. However, specific TFs that are involved in the regulation of lignin biosynthetic pathway genes may be important candidates for developing transgenic plants with enhanced levels of lignin in their roots for the purpose of improved soil carbon sequestration (Vijaybhaskar *et al.* 2008).

1.4 Root-specific and inducible gene expression systems

Identification of suitable tissue-specific and inducible promoter systems to drive target gene expression is another important step in developing plants that have the potential to increase below-ground carbon sticks. Normally, ectopic gene expression in plants is achieved by using a broadly active and constitutive promoter such as the Cauliflower Mosaic Virus (CaMV) 35S promoter (Brand *et al.* 2006). However, ubiquitous and constitutive gene expression can often be lethal or lead to severe defects if the gene being overexpressed is of vital importance to normal plant development. Therefore, the choice of promoter and inducible expression system often determines both the range of tissues and organs in which the gene can be

expressed, in addition to the specific developmental stage in which gene expression can be induced (Moore *et al.* 2006; Brand *et al.* 2006). Root-specific promoters, for example, would be of particular interest in plant biotechnology for genetically engineering improved tolerance to salt and water stress, resistance against root pathogens, improved uptake of nutrients and carbon sequestration (Vijaybhaskar *et al.* 2008; Maizel & Weigel 2004).

The organ and tissue types in higher plants, are both temporally and spatially controlled through the selective expression of specific parts of the genome, in different cells, over the organisms entire life cycle (Ma *et al.* 2005). With the development of high throughput technologies, such as DNA microarrays, there has been a substantial effort made in recent years to identify and determine the relative abundance of transcripts expressed within each organ or tissue type (Ma *et al.* 2005). The ability of microarrays to measure the individual transcript level, for tens of thousands of genes in parallel, provides a way to analyze gene expression levels among different cell types, tissues and even along developmental gradients (Ma *et al.* 2005; Birnbaum *et al.* 2003). Furthermore, a global map of gene expression patterns within an organ, such as the root, can identify genes whose expression is localized to particular areas, thus relating the activity of individual genes, or co-regulated sets of genes, to tissue specialization and even cell fate (Birnbaum *et al.* 2003). Birnbaum *et al.* (2003) mapped global gene expression to 15 different zones of the developing root corresponding to both cell types and tissues at progressive developmental stages. Their data, as well as additional publicly available

microarray data from experiments conducted in other plant organs, allow plant biologists to identify candidate genes involved in specific cell types within the root. By the same token, this data could reveal genes whose promoters may be useful in driving root-specific transgene expression.

The ability to turn on gene expression both spatially and temporally offers the ability to fine-tune ectopic gene expression without compromising the viability of the organism or the function of the organ being altered. However, since it may not be possible to easily identify genes whose expression is truly restricted to the time and place of interest, researchers have also sought "inducible" gene promoters; i.e. a promoter whose transcriptional activity is determined by the presence (or absence) of a specific chemical or physical induction stimulus. In principle, this allows expression of a transgene to be restricted to a given developmental stage for a specific duration. So far there have been several inducible-expression systems described in the literature, generally falling into three broad categories based on the nature of the "inducer": Chemical-inducible, hormone-inducible and temperature-inducible.

Since the early 1990s, several transactivated and chemical-inducible gene expression systems have been developed based on transcriptional de-repression, inactivation, and activation of the gene of interest, as reviewed in Moore *et al.* (2006). In the most popular hormone-inducible systems, the regulatory domains of the rat glucocorticoid receptor, the human estrogen receptor and an insect ecdysone

receptor have been used to construct chimeric transactivation systems whose gene expression activities are controlled by the use of specific hormones or chemically similar compounds (Zuo *et al.* 2001; Moore *et al.* 2006). Alternatively, the molecular responses to environmental temperature changes that have evolved throughout living systems has led to cold tolerance and heat shock phenomena. These phenomena have in turn contributed to the development of temperature-inducible gene regulation (TIGR) systems (Weber *et al.* 2003). Lastly, a further development towards a more stringent control of transgene expression is the use of inducible promoters, which are activated by the application of a specific chemical stimulus (Tang *et al.* 2004). Chemical-inducible systems are appealing compared to alternatives because they are generally dormant in the absence of the inducer, allowing a greater level of flexibility. This in combination with an appropriate tissue-specific promoter to control the chemically-responsive gene product can increase the specificity of target gene expression by restricting it to particular organs, tissues or cell types at a desired point in time (Tang *et al.* 2004). Chemicals that have been used to regulate transgene expression include the antibiotic tetracycline, the steroids dexamethasone (dex) and estradiol, copper, ethanol, benzothiadiazol (the inducer of pathogen-related proteins), the insecticide methoxyfenozide and herbicide safeners (Tang *et al.* 2004).

1.4.1 Herbicidal safeners as inducers of root-specific gene expression

Herbicidal safeners are chemicals that increase herbicide tolerance and protect monocot crops from herbicide burn (DeRidder & Goldsbrough 2006; De Veylder *et al.* 1997; DeRidder *et al.* 2002). Detoxification of these xenobiotics in plants is an

important process involving three enzyme-catalyzed phases. Phase one begins with the oxidation, reduction, or hydrolysis reactions catalyzed by cytochrome P450-dependent monooxygenases (De Veylder *et al.* 1997; DeRidder *et al.* 2002). Phase two involves the conjugation of the newly formed functional group with a hydrophilic substance such as sugars or the tripeptide glutathione (GSH). The GSH conjugation reaction is catalyzed by a class of enzymes known as glutathione S-transferases (GSTs), which essentially "tag" these molecules for excretion or storage. In the final phase, these conjugates are recognized by appropriate transporters (such as ATP-binding cassette transporters) and are then either excreted into the apoplast or sequestered in the vacuole (DeRidder *et al.* 2002).

In monocots, it was found that herbicide tolerance can be markedly enhanced using herbicide safeners, although this phenomenon is less effective dicotyledenous crops (DeRidder *et al.* 2002). Nevertheless, in *Arabidopsis*, a *tau*-class GST (*AtGSTU19*) was shown to respond to safeners in a manner similar to that observed in monocot plants, and to do so in a tissue-specific manner. In response to the safener benoxacor (and to a lesser extent fenclorim) *GSTU19* mRNA levels were increased 30-fold in roots compared to a relatively negligable 4-fold increase in shoots (DeRidder & Goldsbrough 2006).

1.5 Project rationale and thesis objectives

It is important that we learn how plants will respond to the anticipated increases in anthropogenic carbon emissions over the coming decades given their vital role in the global carbon cycle (Lal 2008). This information is critical to understanding the

effects of global climate change on our ecosystems and is required to assess the role of plant life in carbon sequestration (Raven & Karley 2006). Plants offer the potential to play a significant role in carbon sequestration, a process by which atmospheric CO_2 can be transferred to, and securely stored in more long-lived C reservoirs (Lal 2004; Millard *et al*. 2007).

The overall aim of my M.Sc. research was to design and engineer transgenic Arabidopsis plants with enhanced levels of lignin in their roots. If successful, these plants could then offer the potential to increase soil carbon stocks if implemented in crop systems such as canola or soybean.

Lignin is, after cellulose, the second most abundant terrestrial biopolymer and offers the potential to increase soil carbon stocks due to its ability to resist degradation (Humphreys & Chapple 2002; Weng *et al.* 2008). Lignin biosynthesis and accumulation is a highly localized and regulated process that requires strict spatial and temporal control of the processes occurring during normal plant growth and development. The past twenty years of research have led to the identification and characterization of many different lignin biosynthetic and regulatory genes involved in the biosynthesis of monolignols, control of the many genes involved in catalyzing the reactions of the lignin biosynthetic pathway, ultimately leading to secondary cell wall deposition (Anterola & Lewis 2002; Vanholme *et al.* 2008).

Specifically, the objectives of my project were:

1. To identify suitable genes for overexpression that would result in ectopic deposition of lignin
2. To identify suitable promoters needed to drive root-specific expression of the transgene
3. To identify inducible systems that may be used to turn on gene expression spatially and temporally
4. To engineer gene expression constructs designed to enhance lignin deposition in Arabidopsis roots
5. To analyze transgenic plants for relevant phenotypes

2. Materials and Methods

2.1 Organ-specific expression of candidate gene and promoters

Wild type *Arabidopsis thaliana* (Columbia ecotype) seeds were surface sterilized using 20% bleach solution and several washes of dH_2O, sown in (Sunshine Mix #5, Sun Gro Horticulture Canada Ltd., Seba Beach, Alberta, Canada) and grown in a chamber for a 16hr light/8hr dark photoperiod. Root, stem, leaf and flower tissue was harvested from four-week-old plants, frozen in liquid nitrogen and stored at -80˚C for later use. For semi-quantitative RT-PCR analyses of the *At4CL1, AtGSTU19* and *AtSND1* genes, total RNA (1μg) was extracted from frozen tissue using the RNeasy Plant Mini Kit (Qiagen) and the purified RNA treated with DNase I to remove any potential genomic DNA contamination before use for cDNA synthesis. RNA concentration was measured using a NanoDrop® ND-1000 Spectrophotometer at an OD of 260nm. cDNA was made via reverse transcription using qScript™ cDNA SuperMix (Quanta Biosciences), according to the specifications of the manufacturer. PCR (polymerase chain reaction) was performed in a 25μl reaction containing 10x PCR Buffer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.1μl *Taq* DNA polymerase, 0.5μl cDNA template and 0.5μl each of forward and reverse primers. The following program was used:

Step	Temperature	Time	Cycle
1	94°C	3 minutes	Step 4→2 x 35 cycles
2	94°C	30 seconds	
3	54°C	30 seconds	
4	72°C	1 minute	
5	72°C	10 minutes	
6	4°C	Pause	

RT-PCR was repeated three times on two biological replicates, and identical results were obtained. Primers designed to amplify fragments of these native genes can be found in Table 8 in Appendix C (1, 2, 3, 4, 21 and 22). The expression level of the β-*Actin8* gene was used as an internal control (Table 8; 29 and 30). Samples were visualized on 1% agarose gels stained with SYBR® Safe DNA gel stain (Invitrogen). Differentially expressed PCR products were analyzed using the Image J (1.42) (ImageJ: Image Processing and Analysis in Java) program to compare the expression levels of each transcript. The Image J program calculates the area and pixel value statistics of user-defined selections.

2.2 *Cis*-element analysis of candidate promoters

In order to investigate the promoter regions of the *4CL1* (At1g51680) and *GSTU19* (At1g78380) genes for common *cis*-acting root-specific regulatory elements, 500bp, 1000bp, 2000bp and 3000bp regions upstream of the transcription start sites were analyzed using the PLACE (Plant *Cis*-acting Regulatory DNA Elements) database (Higo *et al.* 1999). Putative regulatory elements that could contribute to root-specific expression were identified from previously published literature (Vijaybhaskar *et al.* 2008) and results for the 2000bp analysis is listed in Appendix B (Tables 6 and 7).

2.3 Preparation of the *4CL1pro-SND1* gene expression constructs and transgenic plants

A 1224bp fragment containing the *4CL1* (At1g51680) promoter was amplified via tailed-PCR from Arabidopsis (Columbia ecotype) wild type genomic DNA. The reaction was carried out in a 25μl reaction containing 10x HiFi PCR Buffer, 2mM

$MgCl_2$, 0.2mM dNTPs, 0.1µl HiFi *Taq* polymerase, 1.0µl wild type genomic DNA template and 0.5µl each of forward and reverse primers (Table 8 (Appendix C); 5 and 6) according to the following program:

Step	Temperature	Time	Cycle
1	94°C	5 minutes	Step 4→2 x 35 cycles
2	94°C	30 seconds	
3	54°C	30 seconds	
4	72°C	1 minute 20 seconds	
5	72°C	10 minutes	
6	4°C	Pause	

The forward primer (5'-GGGCAC**G˘AATTC**TTTTCGGTCTCTAATACCTCC-3') contained an *EcoRI* site (underlined and bolded) and the reverse primer (5'-CACGAG**G˘GATCCG˘GTNACCCCGC˘GG**CTGAAGGAAACAGGAGTTGTATC-3') contained restriction sites for *BamHI* (**G˘GATCC**), *BstEII* (**G˘GTNACC**) and *SacII* (**CCGC˘GG**) (underlined and bolded) respectively. Following enzyme digestion with *EcoRI* and *BamHI* the promoter fragment (*4CL1pro*) was ligated into the pPZP211 *Agrobacterium* binary vector (Hajdukiewicz *et al.* 1994). The *SND1* (At1g32770) open reading frame (ORF) was amplified from a pDG2 plasmid (obtained from Apurva Bhargava, Ellis lab) containing the SND1 cDNA using a forward primer (5'-GAGCTC**CCGC˘GG**ATGGCTGATAATAAGGTCAATCTTTCG-3') containing a *SacII* restriction enzyme site (underlined and bolded) and a reverse primer (5'-GGGTGT**G˘GATCCATGATGATGATGATGATG**TCATACAGATAAATGAAGAAGTGGGTC-3') containing a *BamHI* site (underlined and bolded) and a HIS x6 tag (bolded). PCR was carried out in a 25µl reaction containing 10x HiFi PCR Buffer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.1µl HiFi *Taq* polymerase, 0.5µl cDNA template and 0.5µl each of forward and reverse primers. Conditions for SND1 amplification were as follows:

Step	Temperature	Time	Cycle
1	94°C	5 minutes	
2	94°C	30 seconds	
3	58°C	30 seconds	
4	72°C	1 minute 18 seconds	Step 4→2 x 35 cycles
5	72°C	10 minutes	
6	4°C	Pause	

After digestion with *BamHI* and *SacII* the *SND1* ORF fragment was inserted into the pPZP211 vector (Hajdukiewicz *et al.* 1994) already containing the *4CL1pro* fragment. The recombinant plasmid (*4CL1pro-SND1*; Fig. 4) was sequenced (Applied Biosystems, NAPS Unit, UBC, Vancouver, Canada) using the standard M13 primers, transferred into *Agrobacterium tumefaciens* strain GV3101 by heat shock method and then used to transform Arabidopsis wild type plants via the floral dip method. The complete primary sequence of *4CL1pro-SND1* can be found in Appendix A.

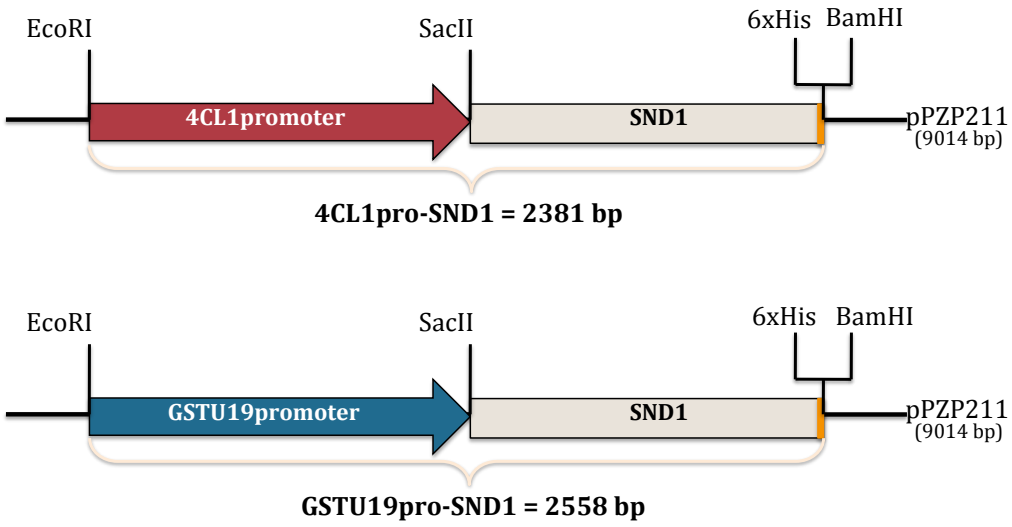


Figure 4. Schematic diagram of the *SND1* overexpression constructs in pPZP211. Separate SND1 overexpression constructs are driven by the 4CL1 and GSTU19 promoters, respectively (left to right the constructs are 5' to 3'). Both constructs contain *EcoRI*, *SacII* and *BamHI* restriction enzyme sites as well as a 6xHis tag at the 3' end (complete primary sequences may be found in Appendix A).

Genomic DNA was extracted from kanamycin-resistant (50μg/ml) T1 generation plants and PCR used to confirm the presence of the transgene. PCR was carried out

in a 25µl reaction containing 10x PCR Buffer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.1µl *Taq* DNA polymerase, 1.0µl cDNA template and 0.5µl each of *4CL1pro* forward (Table 8 (Appendix C); 5) and *SND1* reverse primers (Table 8 (Appendix C); 10) using the following program:

Step	Temperature	Time	Cycle
1	94°C	5 minutes	Step 4→2 x 35 cycles
2	94°C	30 seconds	
3	59.2°C	30 seconds	
4	72°C	2 minute 30 seconds	
5	72°C	10 minutes	
6	4°C	Pause	

T1 generation lines containing the transgene were harvested and T2 generation seeds screened on ½ Murashige and Skoog (MS) media plates containing 50µg/ml kanamycin. I selected 12 plants/line showing a 1:3 segregation ratio indicating a single insertion event and planted them in soil (Sunshine Mix #5, Sun Gro Horticulture Canada Ltd., Seba Beach, Alberta, Canada), where they were grown under 16hr light/8hr dark photoperiod. In addition, 12 plants/line were also transferred to ½ MS media and roots harvested at three weeks for analysis of *SND1* overexpression using RT-PCR. Total RNA (385ng and 1µg starting material) was extracted from frozen tissue using the RNeasy Plant Mini Kit (Qiagen) and the purified RNA treated with DNase I to remove any potential genomic DNA contamination before use for cDNA synthesis. RNA concentration was measured using a NanoDrop® ND-1000 Spectrophotometer at an OD of 260nm. cDNA was made via reverse transcription using SuperScript™ II RT (Invitrogen) and OligodT (Invitrogen), according to the specifications of the manufacturer. All PCR and RT-

PCR reactions were visualized on 1% agarose gels stained with SYBR® Safe DNA gel stain (Invitrogen).

Seeds from 8 lines showing *SND1* overexpression were harvested and screened for homozygosity on ½ MS media plates containing 50μg/ml kanamycin. Of the twelve T3 homozygous sub-lines identified, seven were planted in soil (Sunshine Mix #5, Sun Gro Horticulture Canada Ltd., Seba Beach, Alberta, Canada) and grown under 16hr light/8hr dark photoperiod. Seeds were harvested at approximately eight weeks and used for subsequent analyses.

2.4 Preparation of the *GSTU19pro-SND1* gene expression constructs and transgenic plants

A 1402bp fragment containing the *GSTU19* (At1g78380) promoter was amplified via tailed-PCR from Arabidopsis (Columbia ecotype) wild type genomic DNA. The reaction was carried out in a 25μl reaction containing 10x HiFi PCR Buffer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.1μl HiFi *Taq* polymerase, 1.0μl wild type genomic DNA template and 0.5μl each of forward and reverse primers according to the following program:

Step	Temperature	Time	Cycle
1	94°C	5 minutes	Step 4→2 x 35 cycles
2	94°C	30 seconds	
3	56°C	30 seconds	
4	72°C	1 minute 20 seconds	
5	72°C	10 minutes	
6	4°C	Pause	

The forward primer (5'-GGGTCT**G˘AATTC**GCTACGTGTCGTGAGATATCG-3') contained an *EcoRI* site (underlined and bolded) and the reverse primer (5'-

CACGAG**G˘GATCCG˘GTNACCCCGC˘GG**TGTTACGATCGCTAAAGCTCAC-3') contained restriction sites for *BamHI* (**G˘GATCC**), *BstEII* (**G˘GTNACC**) and *SacII* (**CCGC˘GG**) (underlined and bolded) respectively. Following enzyme digestion with *EcoRI* and *BamHI* the promoter fragment (*GSTU19pro*) was ligated into the pPZP211 Agrobacterium binary vector (Hajdukiewicz *et al.* 1994).

As previously described in section 2.3, the *SND1* amplicon was digested with *BamHI* and *SacII* and inserted into the pPZP211 vector (Hajdukiewicz *et al.* 1994) containing the *GSTU19pro* fragment. The recombinant plasmid (*GSTU19pro-SND1*; Fig. 4) was sequenced (Applied Biosystems, NAPS Unit, UBC) using the standard M13 primers, transferred into *Agrobacterium tumefaciens* strain GV3101 by heat shock method and then used to produce transgenic Arabidopsis plants via the floral dip method. The complete primary sequence of *GSTU19pro-SND1* can be found in Appendix A. Genomic DNA was extracted from kanamycin-resistant (50μg/ml) T1 generation plants and PCR used to confirm the presence of the transgene. PCR was carried out in a 25μl reaction containing 10x PCR Buffer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.1μl *Taq* DNA polymerase, 1.0μl cDNA template and 0.5μl each of *GSTU19pro* forward (Table 8 (Appendix C); 7) and SND1 reverse primers (Table 8 (Appendix C); 10). PCR conditions were as follows:

Step	Temperature	Time	Cycle
1	94°C	5 minutes	Step 4→2 x 35 cycles
2	94°C	30 seconds	
3	58°C	30 seconds	
4	72°C	2 minute 30 seconds	
5	72°C	10 minutes	
6	4°C	Pause	

T2 generation plants were screened on ½ MS media containing 50μg/ml kanamycin, treated with 100μM benoxacor for 24 hours and checked for SND1 overexpression using RT-PCR. I planted 12 plants/line, showing a 1:3 segregation ratio indicating one insertion event were planted, in soil (Sunshine Mix #5, Sun Gro Horticulture Canada Ltd., Seba Beach, Alberta, Canada) and grew them under 16hr light/8hr dark photoperiod. I transferred 10 plants per line to ½ MS media and roots harvested at three weeks for analysis of *SND1* overexpression using RT-PCR. Total RNA (1μg starting material) was extracted from frozen tissue using the RNeasy Plant Mini Kit (Qiagen) and the purified RNA treated with DNase I to remove any potential genomic DNA contamination before use for cDNA synthesis. RNA concentration was measured using a NanoDrop® ND-1000 Spectrophotometer at an OD of 260nm. cDNA was made via reverse transcription using SuperScript™ II RT (Invitrogen) and OligodT (Invitrogen), according to the specifications of the manufacturer. All PCR and RT-PCR reactions were visualized on 1% agarose gels stained with SYBR® Safe DNA gel stain (Invitrogen).

Seeds from eight lines showing *SND1* overexpression were harvested and screened for homozygosity on ½ MS media plates containing 50μg/ml kanamycin. Of the 22 T3 homozygous sub-lines identified, eight were planted in soil (Sunshine Mix #5, Sun Gro Horticulture Canada Ltd., Seba Beach, Alberta, Canada) and grown under 16hr light/8hr dark photoperiod at 22°C. Seeds were harvested at approximately eight weeks and used for subsequent analyses.

2.5 Molecular analysis of transgenic plants

2.5.1 Reverse transcription-PCR of direct downstream targets of SND1

Roots and shoots (aerial tissue in seedlings that does not include stems) from two-week-old plants grown on ½ MS media were harvested and frozen in liquid nitrogen from three different T3 lines for each construct as well as two different empty vector control lines. Total RNA (1μg starting material) was extracted from frozen tissue using the RNeasy Plant Mini Kit (Qiagen) and the purified RNA treated with Dnase I to remove any potential genomic DNA contamination before use for cDNA synthesis. RNA concentration was measured using a NanoDrop® ND-1000 Spectrophotometer at an OD of 260nm. I made cDNA via reverse transcription using SuperScript™ II RT (Invitrogen) and OligodT (Invitrogen), according to the specifications of the manufacturer. PCR was performed in order to amplify four known downstream targets of SND1 (SND3, MYB46, MYB103 and KNAT7) as well as SND1 itself. Primers used can be found in Table 8 of Appendix C (13, 14, 15, 16, 17, 18, 19, 20, 21 and 22) and the PCR reaction carried out in a Biometra Tpersonal thermocycler. The reaction was 25μl and contained 10x PCR Buffer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.1μl *Taq* polymerase, 0.5μl cDNA template and 0.5μl each of the appropriate forward and reverse primers according to the following program:

Step	Temperature	Time	Cycle
1	94°C	5 minutes	
2	94°C	30 seconds	
3	54°C	30 seconds	
4	72°C	50 seconds	Step 4→2 x 35 cycles
5	72°C	10 minutes	
6	4°C	Pause	

All PCR and RT-PCR reactions were visualized on 1% agarose gels stained with SYBR® Safe DNA gel stain (Invitrogen). The PCR reaction was repeated three times yielding similar results.

2.5.2 Reverse transcription-PCR of lignin biosynthetic pathway enzymes

Roots and shoots (aerial tissue in seedlings that does not include stems) from two-week-old plants were harvested and frozen in liquid nitrogen from three different T3 lines for each construct as well as two different empty vector control lines grown on ½ MS media. Total RNA (1μg starting material) was extracted from frozen tissue using the Rneasy Plant Mini Kit (Qiagen) and the purified RNA treated with Dnase I to remove any potential genomic DNA contamination before use for cDNA synthesis. RNA concentration was measured using a NanoDrop® ND-1000 Spectrophotometer at an OD of 260nm. I made cDNA via reverse transcription using SuperScript™ II RT (Invitrogen) and OligodT (Invitrogen), according to the specifications of the manufacturer. PCR was performed to amplify *4CL1* (At1g51680), *CCR1* (At1g15950) and *COMT1* (At5g54160); specific enzymes involved in the lignin biosynthetic pathway. The primers for these enzymes along with the *Actin8* control can be found in Table 8 (Appendix C; 23-30). The PCR reaction was carried out in a Biometra Tpersonal thermocycler. The reaction was 20μl and contained 2x MangoMix (Bioline), 0.5μl cDNA template and 0.5μl each of the appropriate forward and reverse primers according to the following program:

Step	Temperature	Time	Cycle
1	94°C	5 minutes	Step 4→2 x 35 cycles
2	94°C	30 seconds	
3	54°C	30 seconds	
4	72°C	30 seconds	
5	72°C	10 minutes	
6	4°C	Pause	

All PCR and RT-PCR reactions were visualized on 1% agarose gels stained with SYBR® Safe DNA gel stain (Invitrogen). The PCR reaction was repeated in triplicate yielding similar results.

2.6 Determination of lignin content in transgenic plants overexpressing *SND1*

2.6.1 Plant growth conditions

T3 generation transgenic and empty vector lines were grown hydroponically in an open-top liquid culture system. Plastic cylinders that were 1.5-cm in diameter were cut from the tops of disposable 10mL pipette tips were lined with wire mesh, filled with coarse sand, topped off with fine sand and placed in a 0.64-cm-thick Styrofoam platform specifically cut and fitted to float on 7L of hydroponic nutrient medium in an 8L plastic basin. Each platform contained 25 holes (diameter 1.6 cm), into which were fitted the plastic cylinders. Two to four seeds were sown in each cylinder and germinated in dH2O for the first ten days, then transferred to aerated complete nutrient solution at pH 6.1 (1/10 Johnson; see Appendix D). Nutrient solutions were replaced weekly, light was provided from fluorescent tubes (150 E m^{-2} s^{-1}) and the walk-in environment chamber was maintained under the following conditions: light/dark, 8/16 h; 24/20°C; relative humidity = 70%; photon flux of 150 to 200 uE m^{-2} s^{-1}. Roots from both constructs were harvested at eight weeks,

GSTU19pro-SND1 lines treated for 24hrs with 100μM Benoxacor and tissue was stored at -80°C for later use.

2.6.2 Rapid, micro scale, acetyl bromide-based method for lignin content analysis

Lignin content was measured using a modified acetyl bromide method to enable the rapid microscale determination of lignin content in *Arabidopsis* as outlined in Chang *et al.* (2008). Samples (roots from ~10-20 plants) were dried overnight in a 40°C oven and ground using a microball mill at 80-mesh then transferred to vials, placed in a vacuum drying oven at 40°C for 48hrs and then into a P_2O_5 desiccator overnight. Approximately 0.10g (±0.01g) of oven-dried sample was weighed and transferred to a large test tube by adding water. Tubes (containing the sample plus water) were then placed in a 65°C water bath for 30 minutes and vortexed at 10 minute intervals. Samples were then hot filtrated using a Millipore filter with preweighed D47mm (0.45μm) nylon membrane. Samples were washed roughly 25 times with 2mL dH_2O using a glass pipette. Subsequent washes entailed: 25x1mL of ethanol, 25x1mL acetone and 25x1mL of diethyl ether. Membranes were removed carefully and transferred to preweighed aluminum pans and placed into a vacuum drying oven at 40°C for 48hrs and then into the P_2O_5 desiccator overnight. Weights were recorded and difference for extracted weights obtained. Samples were then transferred to new vials. Approximately 5.00mg (±1.00mg) of oven-dried extracted sample (times three replicates per line) was weighed and transferred to a sealable glass test tube. Samples were digested with 1.0mL of 25% acetyl bromide in acetic acid. Tubes were capped and placed in a 70°C water bath for 30 minutes, vortexing

every 10 minutes. Samples were then cooled and stored on ice for a minimum of five minutes up to two hours. Acetic acid (5mL) was added to the tubes containing the samples, vortexed and centrifuged to spin down any precipitate. Subsequently, 300μL of sample mixture was transferred to a quartz cuvette followed by 400μL of 1.5M NaOH, 300μL of 0.5M $H_2NOH{\cdot}HCL$ and 1.5mL of acetic acid for a total volume of 2.5mL. Absorbance was measured at 280nm against a blank and recorded.

2.6.3 Klason lignin or 72% (v/v) H_2SO_4 acid procedure and carbohydrate analysis

Samples were dried at 40°C overnight and ground using a microball mill at 80-mesh then transferred to vials and stored in the desiccator until used. Approximately 0.2g of sample was weighed into a test tube and its mass recorded. The separation reaction was carried out by adding 3mL of 72% (w/w) H_2SO_4 to the weighed samples and mixing with a glass rod every 10 minutes for two hours. Contents of tubes were completely transferred to serum bottles and sealed with septa. Samples were then autoclaved along with the **sugar control** (Appendix D) for one hour at 121°C.

For the insoluble lignin analysis, bottles were allowed to cool before filtering through a pre-weighed Medium Coarseness (M) sintered-glass crucible. The crucible solids were washed by filtering through 200mL warm deionized water followed by drying overnight at 105°C. To complete the retentate analysis, after filtration, crucibles containing the insoluble lignin were weighed and recorded. In order to determine the final weight (dry mass) of insoluble lignin, total crucible

weight (crucible and insoluble lignin) was subtracted from the weight of the pre-weighed empty crucible. For the acid soluble lignin filtrate analysis the absorbance at 205 nm was determined using a quartz cuvette.

For the carbohydrate analysis used to determine hemicellulose content, the filtrate from the autoclaved samples was retained. The sugar analysis of the filtrate required the preparation of a 1mL sample for HPLC by weight using ~950 mg hydrolysate + 50 mg of **fucose standard** (Appendix D).

2.7 Starch analysis

Roughly 25-50mg of dried ground tissue per sample (in duplicate) (see Klason analysis protocol for drying and grinding protocol) was weighed into a 10mL glass culture tube. Following this, 5mL of 4% H_2SO_4 was added to each tube, gently vortexed, then autoclaved for 3½ minutes. Samples were cooled and gently spun for five minutes at 500rpm to pellet the insoluble matter. The supernatant containing the glucose fraction was retained and the pellet discarded. Samples were prepared for HPLC by adding fucose and filtered. Using the glucose standards (Appendix D) and regression analysis, the amount of glucose in the HPLC vial was calculated and then back calculated to determine how much glucose the entire sample released. The glucose content was used to determine the relative cellulose composition of the samples analyzed.

2.8 Phenotypic analysis of transgenic plants

2.8.1 Seed phenotyping

The average weight per seed was determined by weighing six samples of 100 seeds per line and the average seed number per silique was measured by counting the number of seeds in each of 30 siliques. Silique length was determined by measuring 30 siliques for each transgenic line. For the germination assay, 28-36 seeds from two transgenic lines per construct and two empty vector controls were surface sterilized using 70% and 95% ethanol, dried and then sown on ½ MS media. Plates were kept in the dark at 4°C for four days then placed in a walk in growth chamber under 16hr light and 8hr dark. Germinants were counted 24 hours later and every 12 hours after that up to 48 hours. A one-way Analysis of Variance (ANOVA) was performed in the statistical environment 'R' (http://www.bioconductor.org/) using the function 'aov'** with the balanced linear model function 'lm', and contrasts made upon 8 levels for seed weight (A-7, B-5, D-2, F-5, F-7, G-8, EV40, EV41) and 5 levels for lateral root density (A-7, B-5, F-7, G-7, EV40) (see section 2.8.2 below) (Chambers *et al.* 2002).

2.8.2 Root growth and lateral root density

After cold treatment for two days at 4°C, surface sterilized seeds were individually pipetted out in a single row at a seed density of 15 seeds per plate at the top of petri dishes containing 1.2% agar in ½ MS media. Plants were grown vertically in a walk in growth chamber at 16hrs light/8hrs dark for 14 days. I measured 20 seedlings of similar length (to account for different germination times) per genotype and

recorded both the root length and number of lateral roots. *GSTU19pro-SND1* lines were treated for 24hrs with 100μM benoxacor.

2.8.3 Plant growth and height

Transgenic and empty vector control seeds were surface sterilized with 20% bleach solution for 20 minutes and rinsed several times with distilled water then germinated on ½ MS plates then transferred to soil (Sunshine Mix #5, Sun Gro Horticulture Canada Ltd., Seba Beach, Alberta, Canada) and placed in a growth chamber at 16hrs light/8hrs dark photoperiod. Plants were photographed weekly with a Nikon Coolpix E3200 digital camera to track plant height over a six week period.

2.8.4 Microscopy

Fresh sections of the lower and mid part of the stem as well as a 5mm section of the root-hypocotyl (portion of the hypocotyl below the soil surface), from both transgenic and empty vector lines, grown as above, were obtained using a fine razor blade and stained with Phloroglucinol-HCl. Sectioned were placed in water on a slide and visualized using a Leica DM 6000B fitted with a Leica DFC350 Fx camera.

In addition, 5mm sections of root-hypocotyl from both transgenic and empty vector lines were fixed in 20mL vials using a mix of ethanol, acetic acid, formaldehyde and water (Appendix D) then dehydrated with 50%, 60%, 70%, 85%, 95% and 100% ethanol. Tissues were then cleared to allow for paraffin permeation with 100% ethanol and then 25% xylene:75% ethanol, 50% xylene:50% ethanol, 75%

xylene:25% ethanol and 100% xylene. Infiltration was achieved slowly in order to preserve the morphology of the tissue by incubating overnight with a mixture of 100% xylene and Paraplast® plus (Sigma) embedding chips. The vials were incubated at 42°C for one hour to melt the Paraplast® chips and then incubated at 60°C for at least four hours. The xylene/wax mixture was then replaced with 100% molten Paraplast® embedding media and exchanged twice a day for three days (total of six wax changes). Wax moulds were made by pouring the hot wax and tissue into petri dishes, which were then stored at 4°C for later use. Paraffin wax embedded tissues were individually mounted on wooden blocks and sectioned using a rotary microtome (Microm HM 325). The 10 µm sections were heat fixed to glass slides, used for phloroglucinol-HCl staining and lignin autofluorescence (UV 360±40nm) and visualized using a Leica DM 6000B microscope fitted with a Leica DFC350 Fx camera.

3. Results

3.1 Organ-specific expression of candidate gene and promoters

Genes involved in regulating cell fate in all major root tissues have been previously described in *Arabidopsis* (Birnbaum & Benfey, 2004). Birnbaum *et al.* (2003) developed a method that measured high-resolution spatial and temporal gene expression profiles for more than 22,000 genes throughout the *Arabidopsis* root. Using an Affymetrix ATH1 GeneChip, they mapped gene expression in 15 different root zones (endodermis, endodermis and cortex, epidermal atrichoblasts and lateral root cap) that relate to cell types and tissues at progressive developmental stages (stage 1, 2 and 3) (Birnbaum *et al.* 2003).

To identify candidate root-specific genes, I mined the Birnbaum microarray gene expression data set for genes expressed in either the stele or endo-cortex, whose relative probe intensity values were between 1500 and 5000 for those two cell types. Based on this gene expression data, suitable candidate genes were selected whose promoters had the potential to drive root-specific transgene expression, as summarized in Table 1.

Genes found within these specified parameters were then checked via Genevestigator (Genevestigator, 2008) for their relative expression in root compared to other plant organs and tissues (Fig. 5).

Table 1. Candidate genes whose promoters have the potential to drive root-specific transgene expression. These values are based on microarray hybridization signals, which have no units. Values for each of the 40 candidate genes expressed in two cell types (stele and endo-cortex) along three stages of development are summarized.

Gene ID	stele stage1	stele stage2	stele stage3	cortex-endo stage1	cortex-endo stage2	cortex-endo stage3
AT5G11740	2328.20	4072.65	4187.95	1985.63	3473.40	3571.74
AT1G02500	1533.40	2485.74	3911.02	1909.86	3096.00	4871.20
AT5G08690	2447.19	2722.15	2262.39	2381.70	2649.31	2201.85
AT5G19760	2130.62	2934.41	1991.17	2273.26	3130.84	2124.47
AT5G64350	2036.62	2192.35	1815.69	2179.20	2345.83	1942.81
AT5G64400	1951.80	2439.48	1970.78	1963.30	2453.85	1982.39
AT5G44340	1679.18	2069.35	1915.47	2039.91	2513.89	2326.97
AT5G42980	2388.33	2173.63	2080.07	1943.89	1769.15	1693.00
AT3G62290	2491.67	3133.28	2949.11	2569.18	3230.76	3040.86
AT3G55440	3077.48	3992.49	3306.27	2927.79	3798.29	3145.46
AT3G48140	2206.00	2961.10	2688.84	1991.44	2673.10	2427.32
AT4G37830	1950.50	2691.84	1996.96	1677.77	2315.44	1717.73
AT4G33865	4566.83	4278.17	1712.47	4659.56	4365.03	1747.24
AT4G27960	2935.53	3639.66	2929.31	3228.73	4003.20	3221.90
AT4G11150	2053.33	2758.95	2240.15	1862.41	2502.42	2031.87
AT4G09000	2233.24	3067.99	2910.17	2006.88	2757.02	2615.20
AT4G05320	3069.37	4034.32	3232.28	2209.25	2903.80	2326.51
AT4G01850	2554.28	3461.86	2174.89	2819.90	3821.87	2401.06
AT1G18080	3571.20	2940.98	1540.72	3955.22	3257.23	1706.39
AT3G52300	1663.77	1834.42	1629.61	1775.90	1958.06	1739.45
AT3G17390	2674.72	3894.37	2358.48	2787.03	4057.89	2457.51
AT3G09820	1851.13	2265.78	1842.36	1797.13	2199.68	1788.62
AT3G02230	2841.93	4425.77	3271.57	2374.53	3697.89	2733.51
AT1G13440	3409.25	4162.22	4092.73	4073.49	4973.17	4890.14
AT1G78380	3422.04	3875.63	3367.03	2955.32	3347.04	2907.81
AT1G49140	2004.04	2631.03	2008.53	2042.08	2680.97	2046.65
AT1G07890	2681.04	3714.33	3476.60	2993.99	4147.88	3882.40
AT1G65930	2429.96	3191.61	2690.26	2577.25	3385.06	2853.33
AT1G56075	4228.00	4219.88	2731.18	4343.62	4335.28	2805.87
AT1G78040	2328.73	3661.73	2509.98	2305.95	3625.91	2485.43
AT1G79550	1946.76	2503.74	1927.94	2496.27	3210.46	2472.13
AT1G04410	3069.00	3539.86	3280.20	2943.20	3394.76	3145.74
AT2G36530	3452.06	3965.86	2905.62	3172.00	3644.12	2669.89
AT1G09640	2944.01	2990.13	1917.17	3190.15	3240.13	2077.46
AT1G22840	2644.99	3147.32	2588.76	2461.02	2928.41	2408.70
AT1G08830	2355.48	2318.31	2170.68	2290.24	2254.10	2110.56
AT2G16850	3075.62	3860.32	3456.93	2804.01	3519.41	3151.65
AT2G47110	3795.65	3816.35	2060.36	4195.35	4218.23	2277.33
AT2G30870	2107.32	3218.08	4080.73	1603.56	2448.79	3105.22
AT2G33040	2228.69	2377.96	2174.81	2017.78	2152.92	1969.00

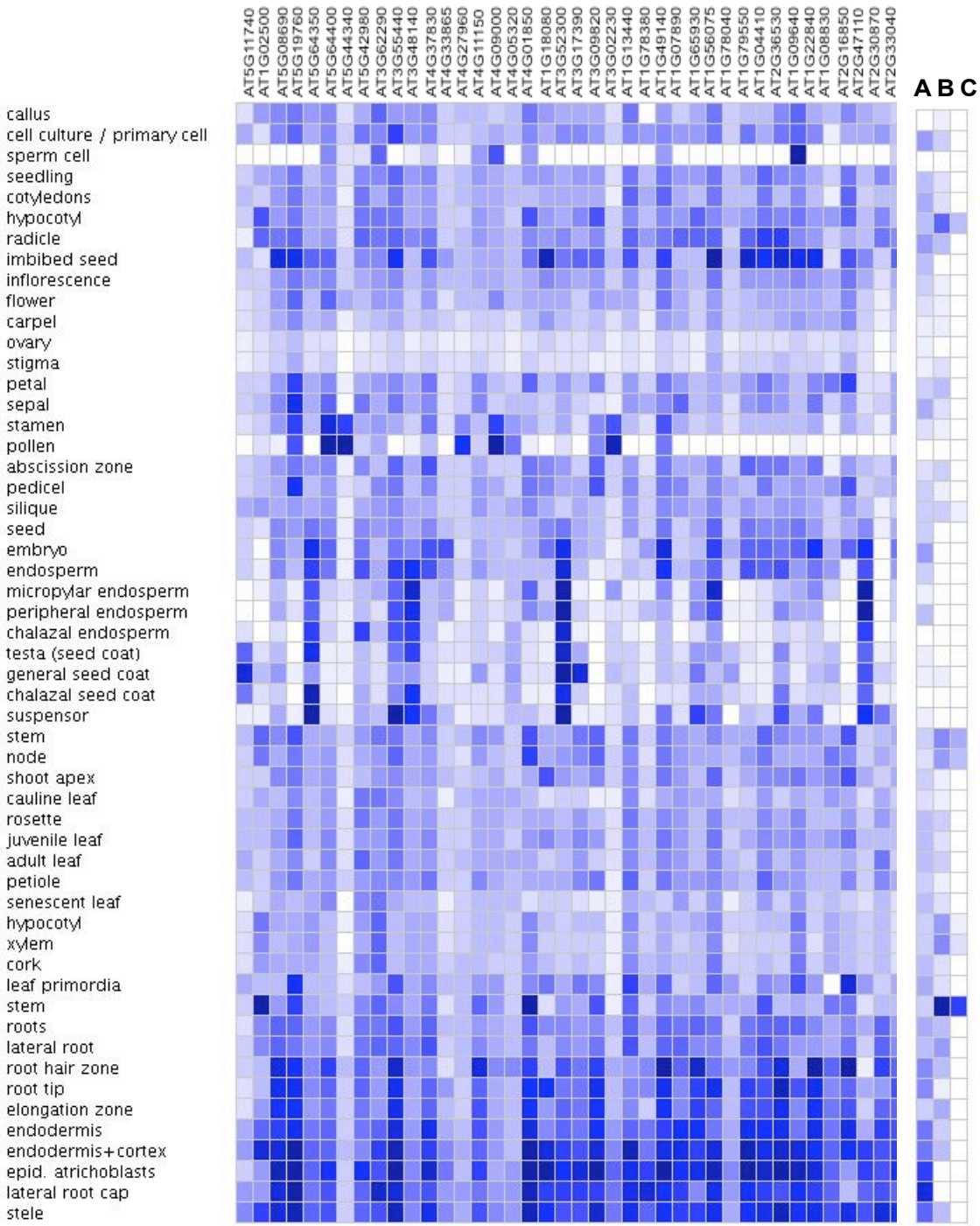


Figure 5. Genevestigator heat map of candidate genes whose promoters have the potential to drive root-specific transgene expression. The diagram represents a global expression map depicting major patterns of gene activity among candidate genes listed in Table 1, in different plant organs and tissues (Genevestigator, 2009). Columns on the right represent two candidate promoters (**A**=*GSTU19* and **B**=*4CL1*) and one candidate gene (**C**=*SND1*) for engineering gene expression constructs to enhance levels of lignin in the roots of transgenic plants.

Based on these results, one candidate gene, *GSTU19,* was selected for further analysis. For the second candidate gene, *4CL1*, previous studies have showed high levels of *4CL1* gene expression in seedling roots, as demonstrated by analysis of transgenic *Arabidopsis* plants containing the *4CL1* or *4CL2* promoter fused to the *beta*-glucuronidase (GUS) reporter gene. These GUS reporter plants show developmentally regulated GUS expression in the xylem tissues of both the root and shoot, although, *At4CL1::GUS* lines showed root-specific expression in seedlings (Soltani *et al.* 2006). In order to confirm these results and validate the potential of these candidates to drive root-specific expression, the activity of both candidate promoters were checked using semi-quantitative reverse transcription (RT)-PCR in flower, leaf, stem and roots of four-week-old plants (Fig. 6). Results confirmed that *GSTU19* is, in fact, expressed at a noticeably higher level in roots compared to other plant organs. However, *4CL1,* showed only a negligible increase in expression in the roots of four-week-old plants as compared to other tissues. Although these results showed *4CL1* to be less promising for root-specific transgene expression, it was retained as a candidate, based on the earlier published data. In addition, *SND1* showed expression in stems but no detectable expression in other organs (Figs. 5 & 6). Along with previous publications on the role of *SND1* in regulating lignin biosynthesis, the combined data shown supports the use of these candidate promoters in producing transgenic plants with higher levels of lignin in their roots. In addition to the endogenous root-specificity of the *GSTU19* gene, the previous studies in *Arabidopsis* showing the increased root-specific expression of *GSTU19* in response to the herbicide safeners, benoxacor and fenclorim (DeRidder &

Goldsbrough 2006), suggested that the *GSTU19* promoter could be useful as a chemical-inducible root-specific gene expression system.

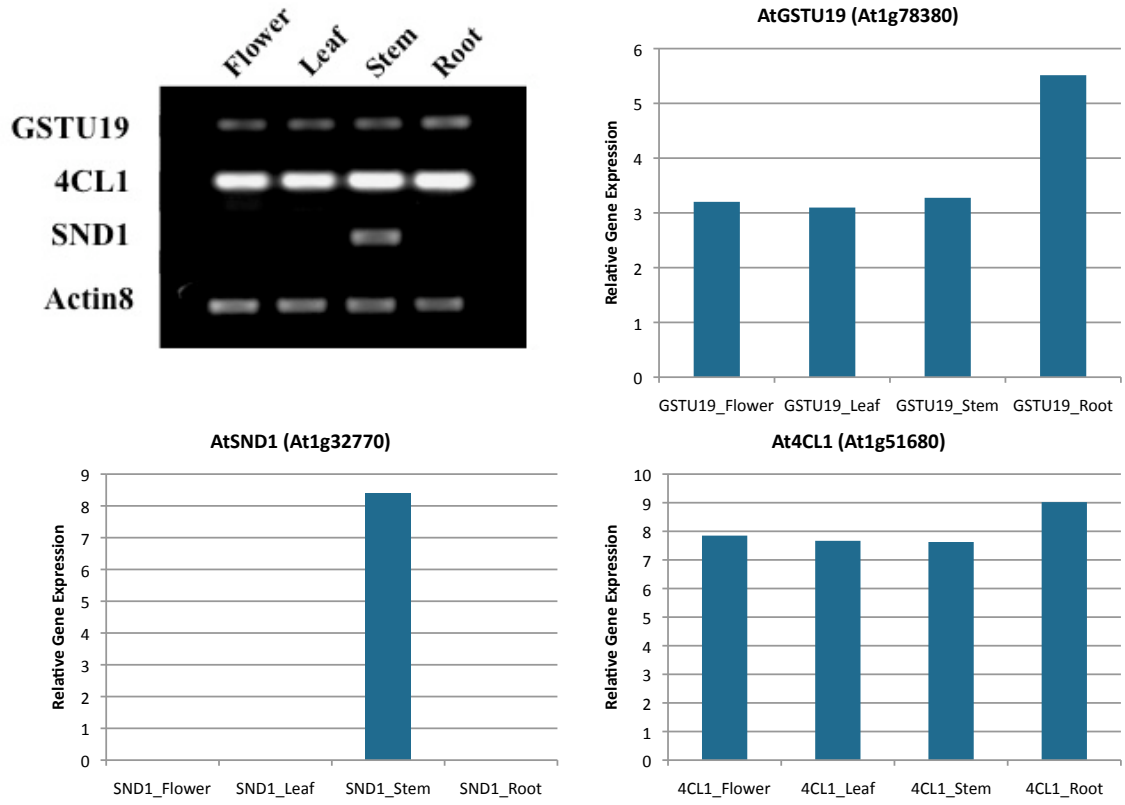


Figure 6. Organ-specific gene expression of candidate gene and root-specific promoters from four-week-old *Arabidopsis* plants. Semi-quantitative reverse transcription (RT)-PCR analysis showing the relative gene expression of *SND1*, *GSTU19* and *4CL1* in flower, leaf, stem and root tissues. Expression of the *Act8* gene was used as both an internal control and loading control. RT-PCR was carried out in triplicate on two biological replicates. Differentially expressed PCR products were analyzed using the Image J (1.42) (ImageJ: Image Processing and Analysis in Java) program to compare the expression levels of each transcript.

3.2 *Cis*-regulatory element analysis of candidate promoters

Several tissue-specific *cis*-acting regulatory elements have been previously described; ACGTROOT1 (Salinas *et al.* 1992), ROOTMOTIFTAPOX1 (Elmayan & Tepfer 1995), WUSATAg (Kamiya *et al.* 2003), OSE1ROOTNODULE (Vieweg *et al.*

2004), OSE2ROOTNODULE (Vieweg *et al.* 2004), RAV1AAT (Kagaya *et al.* 1999), ASF1MOTIFCAMV (Klinedinst *et al.* 2000), SURECOREATSULTR11 (Maruyama-Nakashita *et al.* 2005), SP8BFIBSP8BIB (Ishiguro & Nakamura 1992), ARFAT (Inukai *et al.* 2005), TELO (Tremousaygue *et al.* 1999) and SORLIP1AT (Jiao *et al.* 2005). To investigate possible root-specific elements in the promoters of my candidate genes, 2kb regions of the *4CL1* and *GSTU19* promoters were analyzed using the PLACE (Plant *Cis*-acting Regulatory DNA Elements) database (Higo *et al.* 1999). In addition to the TATA-box and CAAT-box (core promoter sequences required to properly initiate transcription), this analysis revealed the presence of many elements that could possibly be related to root-specific expression. The *cis*-regulatory elements for *4CL1* are summarized in Table 2 and include all of the root expression-associated motifs mentioned above, with the exception of the ACGTROOT1, TELO and SORLIP1AT elements. Similarly, as shown in Table 3, the *GSTU19* promoter contained all the previously described root expression-associated motifs with the exception of the ACGTROOT1 and TELO elements. It should be noted that the frequency of any given *cis*-regulatory motif sequence occurring in the promoter region by random chance may be calculated based on the nucleotide frequency that could occur within a 2kb promoter region, assuming that nucleotides are arranged at random. The elements that were of doubtful statistical significance in the *in silico GSTU19* promoter analysis, are demarcated by an asterisk (Table 3). It is important to note that the sizes of the promoter fragments that were amplified for the transgenic constructs (*4CL1pro* (1224bp) and *GSTU19pro* (1402bp)), were slightly less then the 2kb regions analyzed in PLACE but contained at the very least

one of each of the root expression-associated elements found in the 2kb fragments analyzed.

Table 2. Cis-acting DNA regulatory elements located 2000 bp upstream of the transcription start site of *At4CL1* (At1g51680). The high frequency regulatory elements are shown first as well as the number of times the element appears on both the (+) and (-) strands (actual frequency). The third column represents the number of times that a motif could occur at random assuming all four nucleotides are represented equally, given the number of base pairs in the sequence (i.e. $1:4^x$, where x is the number of base pairs in the motif sequence), in the 2kb promoter region analyzed. This number gives an indication of the number of elements that would need to appear in the promoter (on a single strand) in order for the over-represented motif to be statistically significant, based on the statistical frequency of occurrence of that sequence.

Putative root motif	Sequence	Statistical frequency of occurrence in the 2kb promoter fragment analyzed	Actual frequency
ROOTMOTIFTAPOX1	ATATT	1.95:2000	13+; 16-
RAV1AAT	CAACA	1.95:2000	6+; 1-
ASF1MOTIFCAMV*	TGACG	1.95:2000	2+; 2-
OSE2ROOTNODULE	CTCTT	1.95:2000	4+
OSE1ROOTNODULE	AAAGAT	0.488:2000	2+; 1-
SURECOREATSULTR11	GAGAC	1.95:2000	3-
SP8BFIBSP8BIB	TACTATT	0.122:2000	2-
ARFAT	TGTCTC	0.5:2000	1+
WUSATAg	TTAATAG	0.122:2000	1-

*Sequence of doubtful statistical significance

Table 3. Cis-acting DNA regulatory elements located 2000 bp upstream of the transcription start site of *AtGSTU19* (At1g78380). The high frequency regulatory elements are shown first as well as the number of times the element appears on both the (+) and (-) strands (actual frequency). The third column represents the number of times that a motif could occur at random assuming all four nucleotides are represented equally, given the number of base pairs in the sequence (i.e. $1:4^x$, where x is the number of base pairs in the motif sequence), in the 2kb promoter region analyzed. This number gives an indication of the number of elements that would need to appear in the promoter (on a single strand) in order for the over-represented motif to be statistically significant, based on the statistical frequency of occurrence of that sequence.

Putative root motif	Sequence	Statistical frequency of occurrence in the 2kb promoter fragment analyzed	Actual frequency
ROOTMOTIFTAPOX1	ATATT	1.95:2000	6+; 8-
OSE1ROOTNODULE	AAAGAT	0.488:2000	1+; 5-
OSE2ROOTNODULE	CTCTT	1.95:2000	4+; 2-
ASF1MOTIFCAMV	TGACG	1.95:2000	3+; 2-
RAV1AAT*	CAACA	1.95:2000	2+; 2-
SORLIP1AT	GCCAC	1.95:2000	4+
SURECOREATSULTR11*	GAGAC	1.95:2000	1+; 1-
ARFAT	TGTCTC	0.488:2000	1+
SP8BFIBSP8BIB	TACTATT	0.122:2000	1-
WUSATAg	TTAATAG	0.122:2000	1+

*Sequence of doubtful statistical significance

3.3 *SND1* overexpression in transgenic plants

Two gene expression constructs (*GSTU19pro-SND1* and *4CL1pro-SND1*) were engineered by PCR amplification and ligation of the *GSTU19* and *4CL1* promoters and *SND1* ORF with the pPZP211 *Agrobacterium* binary vector. These constructs were then introduced into *Arabidopsis* plants using *Agrobacterium*-mediated transformation. PCR analysis of genomic DNA was used to select T1 generation kanamycin-resistant transgenic lines by confirming the presence of the transgene. Roots from three-week-old T2 generation kanamycin-resistant transgenic lines were subsequently analyzed using RT-PCR to determine whether the *SND1* transgene was being overexpressed. The RT-PCR analysis detected overexpression of *SND1*, compared to wild type, in ~90% of the lines analyzed for both constructs, as shown in Figure 7.

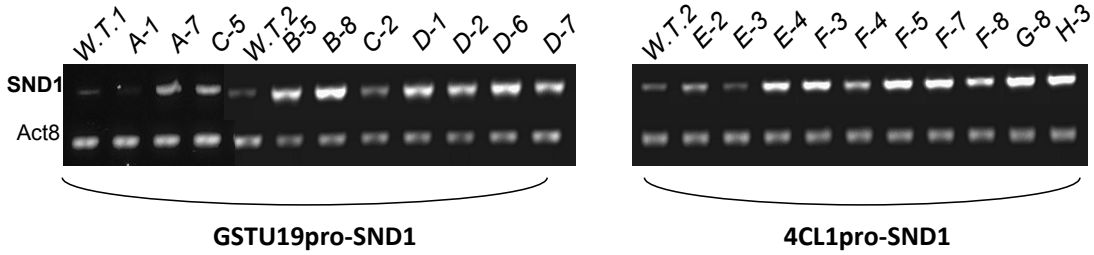


Figure 7. Transcriptional analysis of T2 generation plants overexpressing *SND1* using RT-PCR. Total RNAs were isolated from three-week–old root tissue of 10 independent transgenic plant lines from each construct as well as wild type control plants. *GSTU19pro-SND1* lines were induced with 100μM benoxacor on ½ MS solid media for twenty-four hours prior to RNA extraction. *Actin8* was used as an internal and loading control as shown by comparable expression levels.

These lines represent a mixture of both homozygous and heterozygous individuals; therefore, among the T2 generation lines showing overexpression, 12 sub-lines

were screened for homozygosity. Twenty-two kanamycin-resistant homozygous sub-lines were identified for *GSTU19pro-SND1* and twelve for *4CL1pro-SND1*. These T3 generation transgenic lines, homozygous for a single active T-DNA insert, were used for further experiments to determine the possible effects of *SND1* overexpression.

3.4 Molecular analysis of transgenic plants overexpressing *SND1*

Given the recent identification of SND1 as a master transcriptional switch activating the developmental program of secondary cell wall biosynthesis and as an activator of several transcription factors that are involved in that process (Zhong *et al.* 2006; Zhong *et al.* 2008), I predicted that *SND1* overexpression would result in an increase in expression of direct targets of SND1, such as MYB46, SND3, MYB103 and KNAT7. Reverse transcription PCR analysis of these direct targets was conducted for two reasons: 1) to determine whether the secondary cell wall gene regulatory networks previously described were present and functional in roots, and 2) to investigate the root-specificity of the constructs. As shown in Figure 8, *SND1* was found to be upregulated in both roots and shoots (aerial tissue in seedlings that does not include stems) compared to empty vector control lines. In contrast, the other transcription factors (TFs) analyzed showed negligible changes in gene expression in shoots but showed a more noticeable increase in gene expression in roots. Given that these TFs are normally preferentially expressed in stems (Zhong *et al.* 2006; Zhong *et al.* 2008), this data provides evidence that the *SND1* overexpression constructs are behaving in a root-preferential manner and that *SND1* overexpression results in an increase in gene expression of its direct targets.

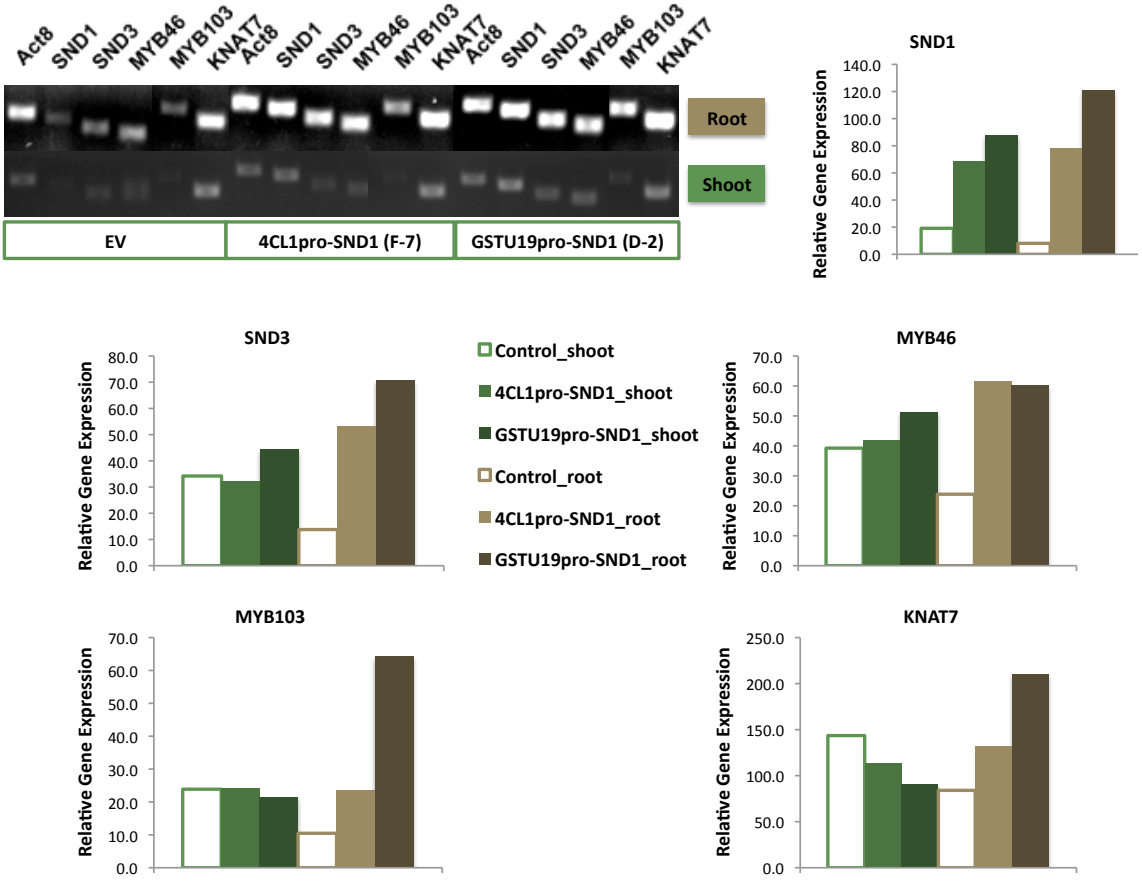


Figure 8. Transcriptional analysis of transcription factors known to be direct targets of SND1. Three-week-old T3 generation Arabidopsis seedlings grown on ½ MS solid medium and *GSTU19pro-SND1* lines treated for 24 hours with benoxacor (100 µM). Total RNA was extracted from roots (R) and shoots (S) of transgenic and empty vector lines. Transcription factors were analyzed using RT-PCR. *Actin8* was used as an internal and loading control as shown by comparable levels. Differentially expressed PCR products were analyzed using the Image J (1.42) (ImageJ: Image Processing and Analysis in Java) program to compare the expression levels of each transcript relative to the *Actin8* control. SND1 (At1g32770); SND3 (At1g28470); MYB46 (At5g12870); MYB103 (At1g63910); KNAT7 (At1g62990).

To determine whether the result of the ectopic gene expression of these TFs specifically influences lignin biosynthesis in roots, RT-PCR analysis was also performed on genes encoding three indicative lignin biosynthetic pathway enzymes (4CL1, CCR and COMT), as shown in Figure 9.

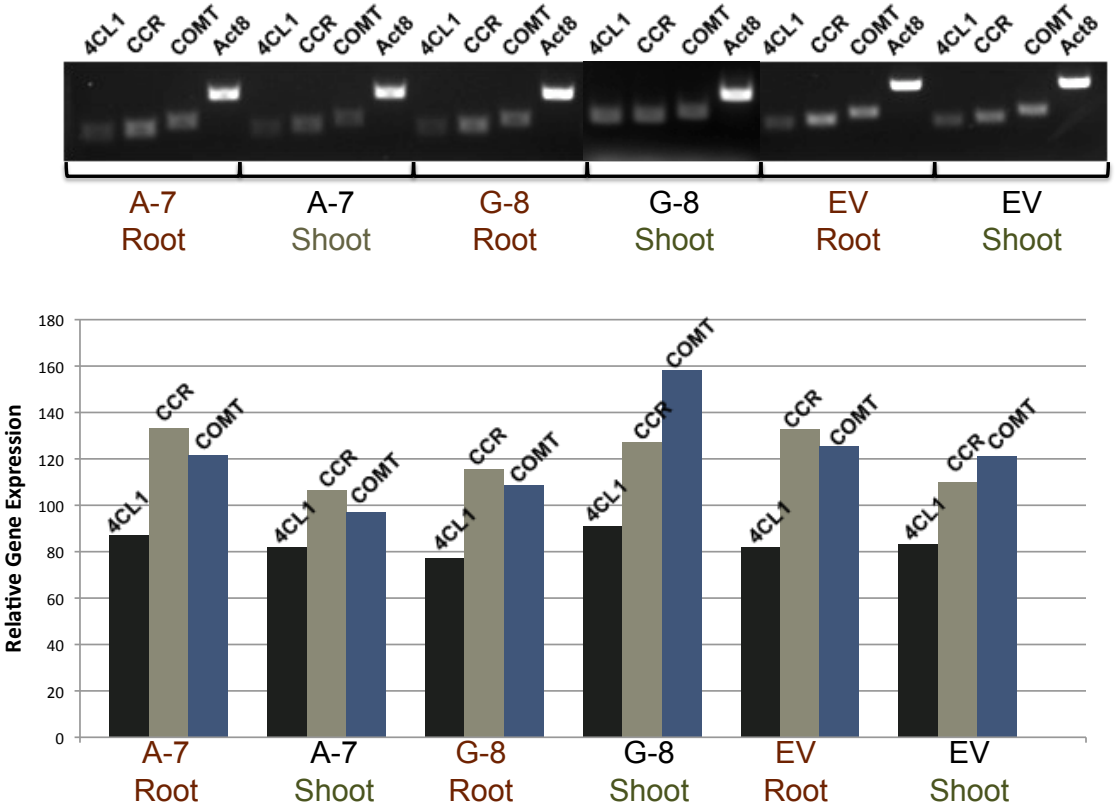


Figure 9. **Reverse transcription PCR analysis of genes involved in lignin biosynthesis.** Three-week-old T3 generation Arabidopsis seedlings grown on ½ MS solid medium and *GSTU19pro-SND1* lines treated for 24 hours with benoxacor (100 µM). Total RNA was extracted from roots (R) and shoots (S) of transgenic and empty vector lines. Act8 was used as an internal and loading control as shown by comparable levels. Differentially expressed PCR products were analyzed using the Image J (1.42) (ImageJ: Image Processing and Analysis in Java) program to compare the expression levels of each transcript relative to the Act8 control.

In contrast to the results for expression of the secondary cell wall-related TFs, I observed no difference in gene expression among the lignin biosynthetic genes or among tissue types compared to empty vector controls. This data suggests that unlike the previously described constitutive overexpression of *SND1*, which showed ectopic deposition of lignified secondary walls in normally non-sclerenchymatous cells of flowers, leaves and stems (Zhong *et al.* 2006), overexpression of *SND1* in

roots had no influence on the expression of genes encoding certain key lignin biosynthetic enzymes.

3.5 Determination of lignin content in transgenic plants overexpressing *SND1*

3.5.1 Determination of lignin content in transgenic plants overexpressing *SND1* by rapid micro-scale acetyl bromide method

To determine total lignin content, several methods and techniques have been developed and adapted in order to quantitatively determine total lignin content and composition in plant tissues (Hatfield & Fukushima 2005). To analyze total lignin content (w/w) in the roots of transgenic plants overexpressing *SND1*, I first used a rapid micro-scale method as outlined in Chang *et al.* (2008). This acetyl bromide-based lignin micro-scale assay was primarily developed to provide a rapid yet sensitive method of determining lignin concentration, using small amounts of plant material. This method is useful for small samples whose size is unsuitable for procedures that rely on the production and gravimetric measurement of an insoluble lignin residue, such as the Klason lignin analysis.

Based on the previous studies that had shown SND1 to be a master transcriptional switch activating the developmental program of secondary cell wall biosynthesis in fibres, I predicted that the overexpression of *SND1* and its direct target genes in roots would cause an increase in total lignin content (Zhong *et al.* 2006). Unexpectedly, my analysis of roots of the transgenic *SND1* overexpression lines, showed a 47% and 40% decrease in total lignin content in both *GSTU19pro-SND1* overexpression lines (A-7 and B-5 respectively) and a 46% decrease in lignin

content in one of the two *4CL1pro-SND1* overexpression lines (G-8) (Fig. 10), compared to the roots of empty vector control plants. The second *4CL1pro-SND1* overexpression line analyzed (F-5) showed no obvious change in lignin content (1% decrease) compared to the empty vector control. This result appears to be correlated with the lack of altered gene expression among the lignin biosynthetic genes observed in these same genotypes (Fig. 9).

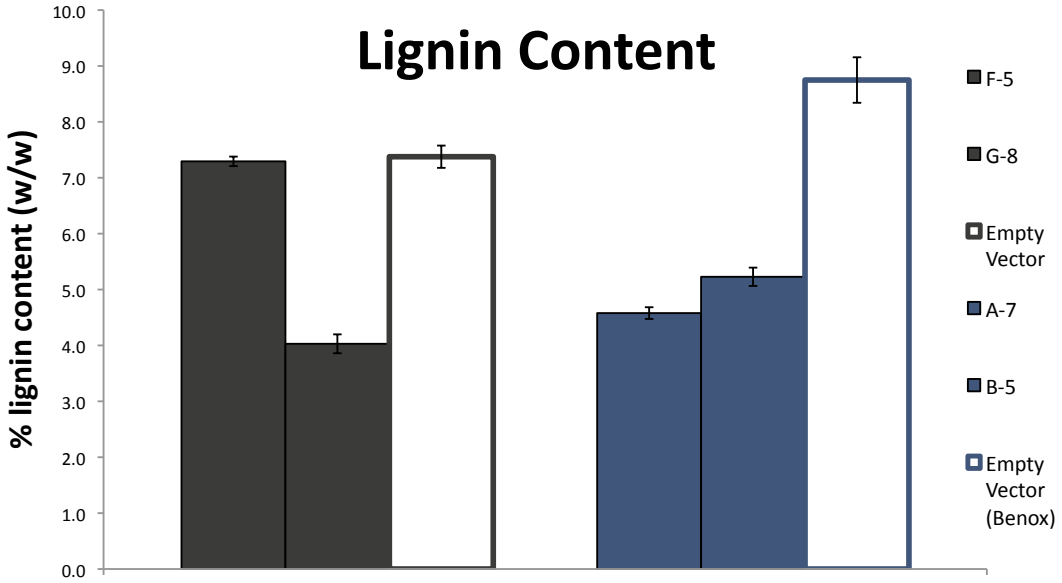


Figure 10. Lignin content in transgenic *Arabidopsis* plants overexpressing SND1. Percent lignin content (w/w), determined by the rapid microscale acetyl bromide method, in empty vector control and transgenic Arabidopsis plants expressing the *4CL1pro-SND1* (Grey) and *GSTU19pro-SND1* (Red) constructs. Control plants contain pPZP211. Error bars indicate standard error from three technical replicates (control and transgenic lines are T3 generation).

3.5.2 Cellulose, starch and Klason lignin analysis

Because SND1 has been shown to be a master transcriptional switch activating the developmental program of overall secondary cell wall biosynthesis in fibres, as opposed to just lignin biosynthesis, I reasoned that the decrease in lignin content

and lack of change in expression of genes encoding lignin biosynthetic enzymes, could be a result of carbon being reallocated to a different area of carbon metabolism.

Plants use photosynthesis to chemically convert CO_2 to carbohydrates, such as cellulose and starch. Cellulose is an important component of the cell walls of higher plants and the world's most abundant organic polymer, serving as another major carbon sink in plants (similar to lignin) (Delmer & Haigler 2002). One other major plant carbon sink is the other major glucan, starch (α-1,4-glucan with α-1,6 branches). As leaves (sources that export carbon) and storage organs (sinks that import carbon) expand, they enlarge and deposit their cellulose in their primary walls before the developmental transition that leads to starch deposition (Delmer & Haigler 2002). While the ratio of cellulose to other cell wall polymers can change considerably, until recently it was not clear from the publicly available literature whether carbon flux in plants with altered lignin biosynthetic pathways directly altered other carbon-polymer synthetic pathways (Delmer & Haigler 2002). Studies have now shown that alterations in lignin deposition can cause relative cellulose content to increase, as a result of these perturbations (Coleman *et al*. 2008).

To test my carbon reallocation hypothesis, I analyzed both cellulose and starch content (in addition to insoluble lignin content). The carbohydrate analysis provided an indirect measure of the cellulose (quantified as glucose monomers) and other wall pollysacharides (pectin and hemicellulose), (quantified as other sugars

such as rhamnose, fucose, arabinose, xylose, mannose and galactose monomers) content of transgenic lines overexpressing *SND1* (Table 4). These results showed a 29% decrease in cellulose content and 26% decrease in hemicellulose content in the *GSTU19pro-SND1* line (D-2) compared to the empty vector control. The *4CL1pro-SND1* line (G-8) showed a slight decrease of 3.5% in cellulose content and a negligible 1.5% decrease in hemicellulose content. Furthermore, the Klason lignin analysis revealed a 23% decrease in lignin content in the *GSTU19pro-SND1* overexpression line (D-2) compared to the control, which was consistent with the decrease in lignin content found for *GSTU19pro-SND1* lines analyzed using the acetyl bromide-based method. Conversely, the *4CL1pro-SND1* line (G-8) showed less then a 0.1% increase in lignin content compared to the empty vector control, a nominal amount. This line, when analyzed by the acetyl bromide based method, showed a 46% decrease in lignin content as described in the previous section. Although the *4CL1pro-SND1* (G-8) line showed different results when analyzed using two different methods, the results shown here using the Klason procedure are similar to the other *4CL1pro-SND1* line (F-5) analyzed using the acetyl-bromide based method.

Table 4. Cell wall composition of roots from empty vector and transgenic lines overexpressing *SND1*. Numbers represent milligrams of cellulose, hemicellulose and lignin per 100 milligrams of initial dry weight. Absolute values shown are from a single biological replicate.

Line	Cellulose (mg/100mg)	Hemicellulose (mg/100mg)	Lignin (mg/100mg)
Empty Vector	30.47	22.97	22.1
GSTU19pro-SND1 (D-2)	21.47	17.08	17.1
4CL1pro-SND1 (G-8)	29.39	22.63	23.8

3.6 Phenotypic analysis of transgenic plants overexpressing *SND1*

3.6.1 Seed phenotyping

It is desirable to avoid pleiotropic effects that might result from constitutive overexpression of target genes in agricultural systems, which is why the ability to drive transgene expression in a location-specific and controlled manner is important. I wished to determine whether the transgenic plants overexpressing *SND1,* displayed any phenotypes that might reflect an impact of transgene expression on normal plant growth and development. As one measure of overall growth and productivity, I decided to analyze seed-related traits. My transgenic plants overexpressing *SND1* did not show significant deviations from control plants (empty vector lines) in terms of the average number of seeds per silique, average silique length, or average germination rate when compared to empty vector control lines (Fig. 11; A, B and D). Average seed weights for all lines fell into a range of 18-25μg per seed. The results for average weight per seed showed significant differences (Fig. 11; C) as represented by the lack of overlap in the error bars, but significant variation was also seen for both control lines as well. A one-way ANOVA test of the overall model was done to determine if there was a statistically significant difference in means (with respect to seed weight) between genotypes. In this case, the p-value was small $P < 0.001$ (Appendix E), therefore there was a statistically significant difference in seed weight among genotypes.

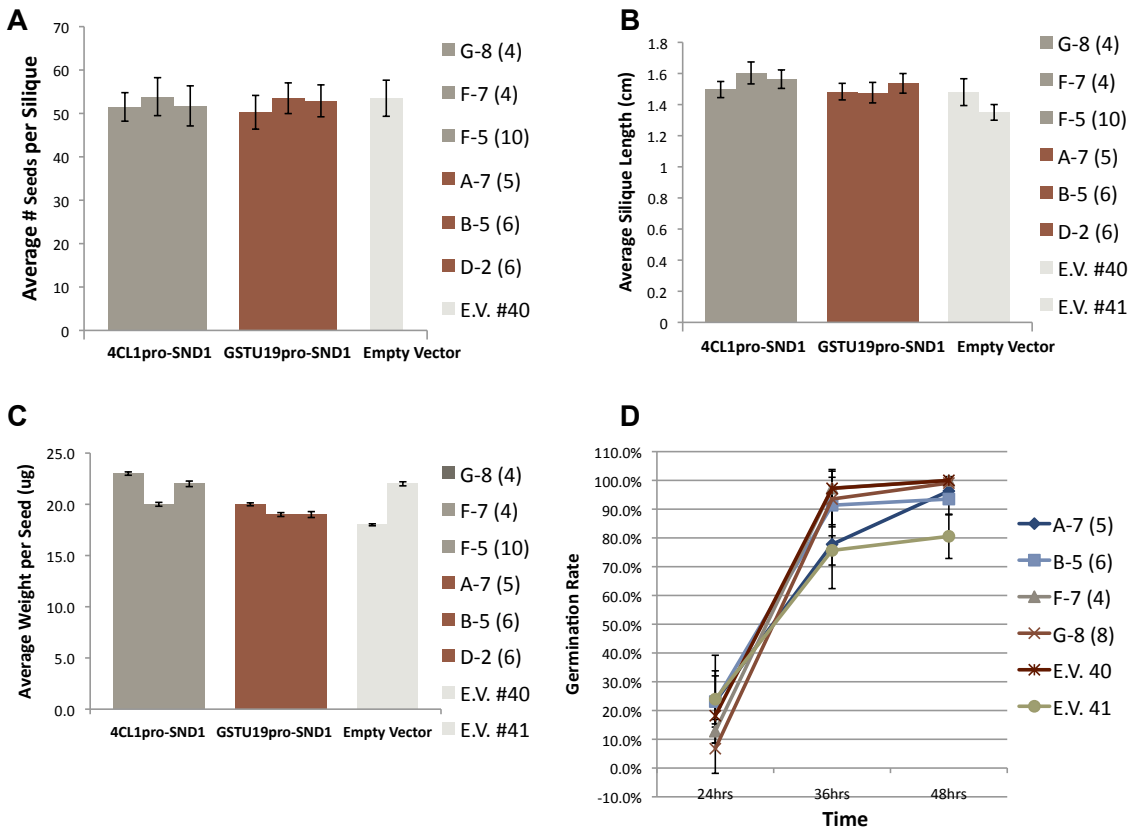


Figure 11. Seed-related phenotypes of T3 generation seeds from transgenic and empty vector constructs. Average number of seeds per silique (A), average silique length (B), average weight per seed (C) and average germination rate (D). Error bars represent the 95% confidence interval of an average of 30 samples per genotype for (A) and (B), 6 samples per genotype for (C) and 3 samples per genotype for (D). *GSTU19pro-SND1* lines were treated for 24 hours with 100μM benoxacor in water, at 4-weeks.

3.6.2 Root growth and lateral root density

Another facet of plant development in *Arabidopsis* is the production of a highly branched root system. Plant roots are important tissues involved in many processes such as uptake of water, interactions with soil-microbes, the secretion of compounds required for defense against pathogens and absorption of soil nutrients. Furthermore, they protect the aboveground tissues against the effects of acidic

conditions or heavy metals in the soil, and against desiccation (Koyama *et al.* 2005). Since I was overexpressing a transcription factor in the roots, which is not a tissue in which it is usually highly expressed, I asked whether the overexpression of *SND1* in the roots of my transgenic plants had altered their root development. To assess this, I examined root growth and architecture by measuring: 1) the primary root extention among 14-day-old seedlings at a similar stage of developmental (i.e. similar primary root length) (Fig. 12A) and 2) the number of lateral roots forming on these primary extentions (Fig. 12B). Lateral root density (LRD) was then determined by dividing the average number of lateral roots counted, by the average length of the primary root (Fig. 12C).

Both of the *GSTU19pro-SND1* transgenic lines (A-7 and B-5) showed an increase in LRD compared to the empty vector control. In comparison, one of the two *4CL1pro-SND1* transgenic lines analyzed, (G-8), showed higher lateral root density, whereas the other line (F-7) did not show any difference compared to the empty vector control line. A one-way ANOVA test of the overall model was done to determine if there was a statistically significant difference in means (with respect to LRD) between genotypes. In this case, the p-value was small P = 0.000 (Appendix E), therefore there was a statistically significant difference in lateral root density among genotypes.

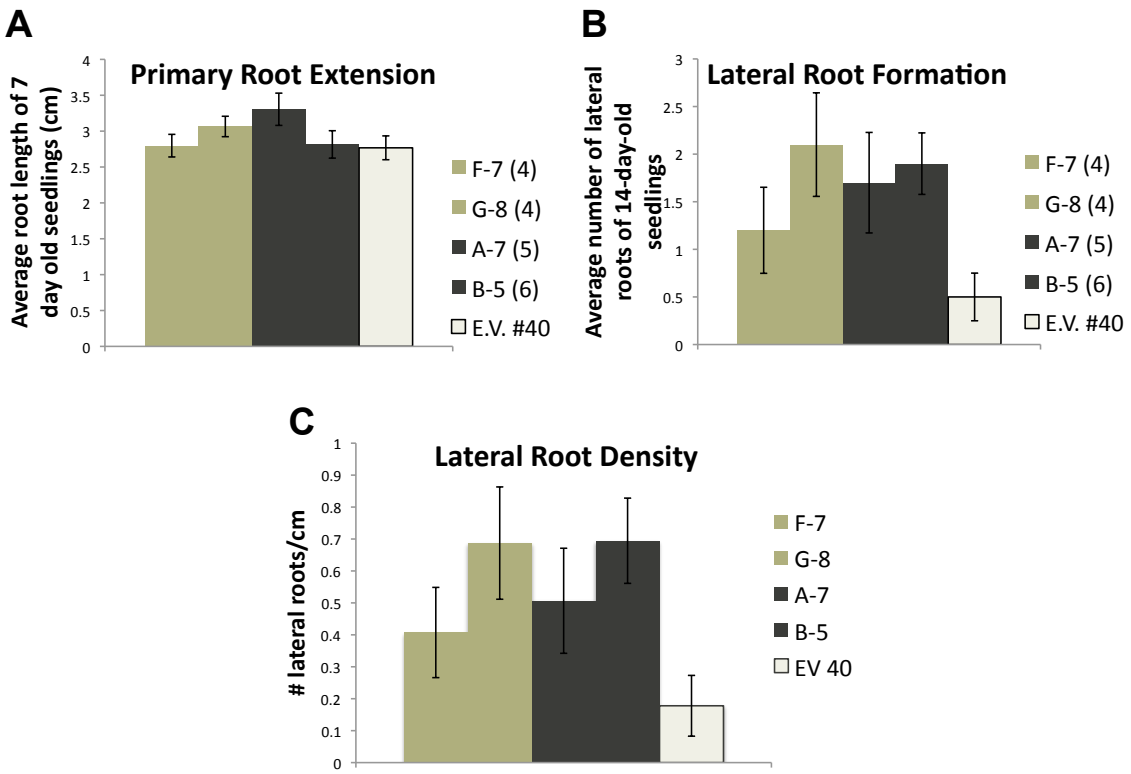


Figure 12. Primary root extention, lateral root formation and number of lateral roots per cm (lateral root density) of 14-day-old seedlings. Left to right, *4CL1pro-SND1* (red); *GSTU19pro-SND1* (grey); empty vector (neutral). Lateral root density was calculated by dividing the number of lateral roots by the length of primary root (cm). Lateral roots that had emerged at least 1.0 mm from the root surface were counted. Error bars represent the 95% confidence interval of an average of 20 samples per genotype.

3.6.3 Plant growth and height

Plant growth and development are controlled by the combined action of many different signaling pathways, which integrate information from the environment with metabolic and developmental signals. If these normal developmental pathways, such as the phenylpropanoid pathway, are disrupted or altered, severe consequences to overall plant growth and function could ensue. To investigate the effects of *SND1* overexpression in roots on general plant growth and development, I

examined transgenic lines over a six-week period to look for any obvious phenotypic differences in normal plant growth and development, such as flowering time, overall plant height and shape, and leaf morphology. Transgenic plants did not show any visible phenotypic differences as compared to empty vector control lines as shown in Figure 13.

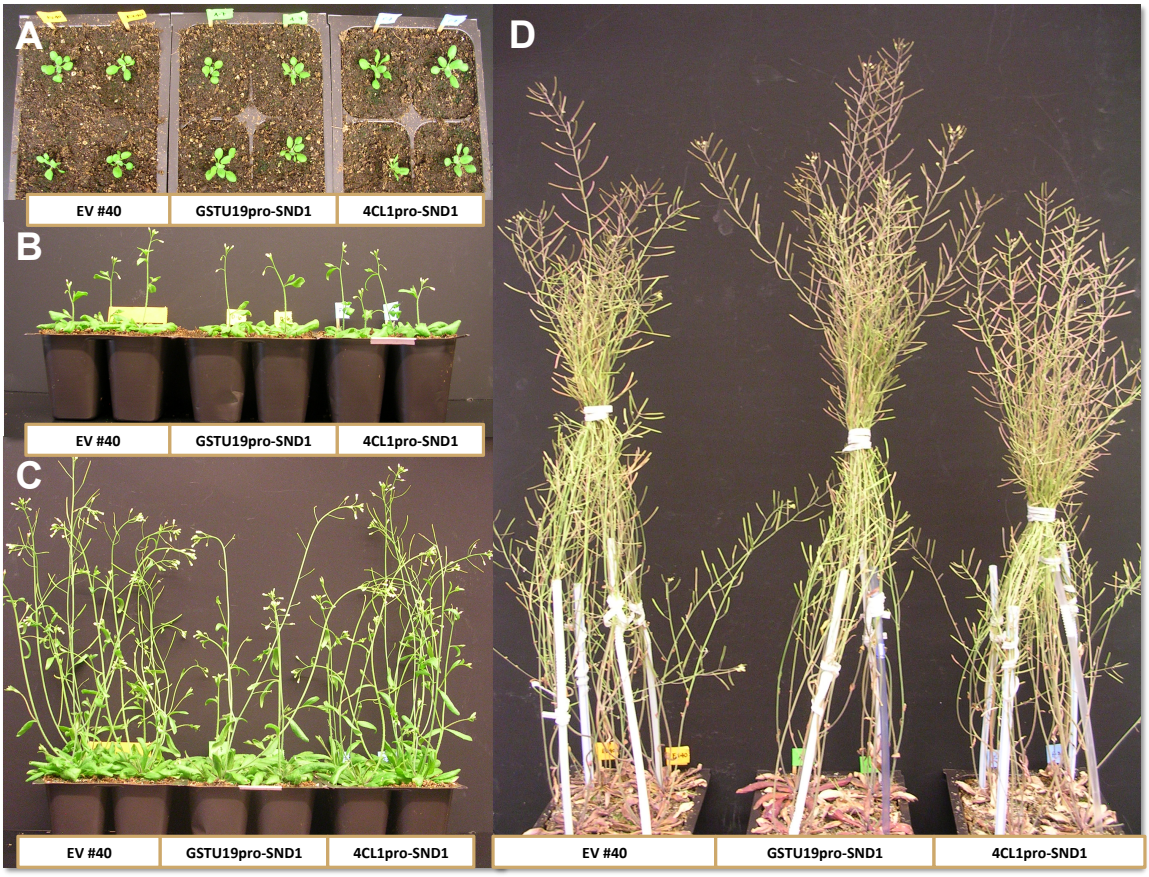


Figure 13. Plant growth and height time-course experiment for transgenic plants overexpressing *SND1* and empty vector lines. Photographs are detailing plant growth at (A) 3 weeks, (B) 4 weeks, (C) 5 weeks and (D) 6 weeks. *GSTU19pro-SND1* lines were treated at 4 weeks for 24 hours with 100μm benoxacor.

3.6.4 Microscopy

In SND1 overexpression plants where the gene was under the control of the CaM35S promoter, ectopic secondary wall thickening was not always observed in root cell types but was occasionally seen in the epidermal cells of hypocotyls and cortical cells (Zhong *et al.* 2006). To investigate lignin deposition patterns in my transgenic lines overexpressing *SND1* under the control of more root-specific promoters, I conducted a histochemical analysis of the root-hypocotyl in the various different transgenic plant genotypes I had developed. Root-hypocotyls were fixed, embedded in paraffin wax and sectioned. For visualization of lignified secondary walls, the sections were stained with phloroglucinol-HCl reagent (Pomar *et al.* 2002). Phloroglucinol–HCl staining should identify cell walls that have lignin deposition, by staining them red. However, my transgenic plants did not show any visible differences in lignin deposition when compared with empty vector controls (Fig. 14). There was notable variation in lignin content along the 5mm sections of hypocotyl analyzed, however, which made it difficult to establish developmental equivalencies. Nevertheless, the histochemical results suggest that *SND1* overexpression in root tissue had produced no observable difference in lignin deposition patterns in the tissues analyzed.

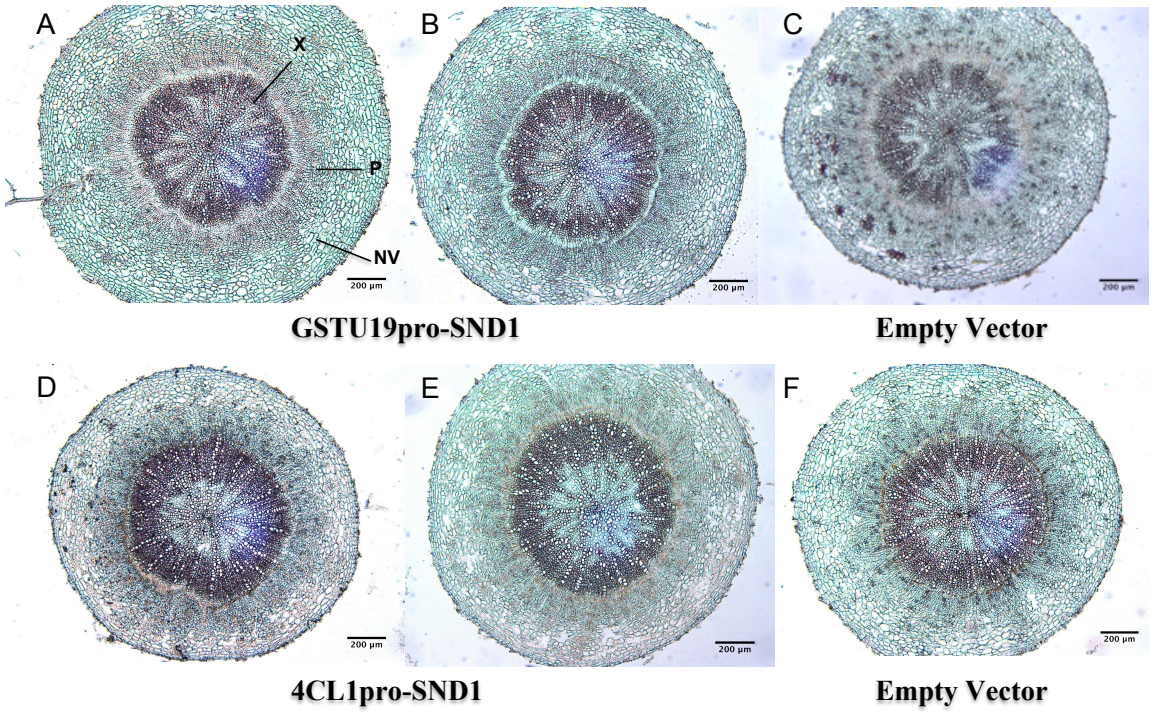


Figure 14. Wax-embedded root-hypocotyl cross-sections of SND1 overexpressors and empty vector control lines. The 10 µm sections were stained with phloroglucinol-HCl to show lignified walls. **X**=xylem, **P**=phloem and **NV**=non-vascular. Scale bars represent 200µm at 5x magnification. A=A-7; B=D-6; C=EV (safener treated); D=F-7; E=G-8; F=EV. A, B and C were treated with 100µM benoxacor for 24 hours.

To further analyze lignin deposition patterns, lignin autofluorescence was monitored in tissues irradiated with UV light at 360±40nm (Fig. 15). Autofluorescence at this irradiation wavelength allows an assessment of the overall localization of lignin in tissues that are lignified (Tao *et al.* 2009). Observations from low (5x) to high (40x) magnification (data not shown) revealed no apparent differences in lignin location or architecture. Again, there was some variation in the observed fluorescence along the 5mm developmental gradient. However, as seen at 20x magnification (Fig. 15) there was no substantial difference in cell wall thickness or organization among sections and tissues analyzed.

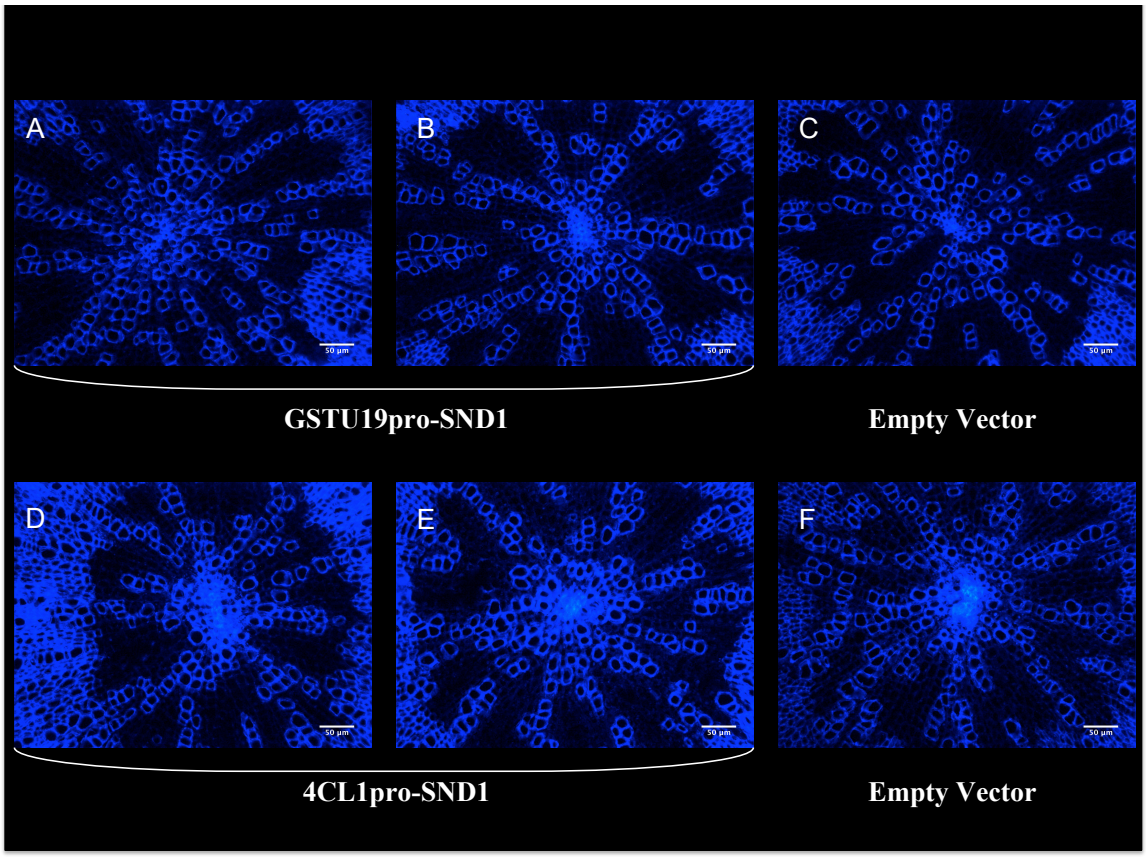


Figure 15. Auto-fluorescence of lignin in root-hypocotyl cross-sections. UV fluorescence microscopy (UV 360±40nm) of 10μm wax-embedded root cross-sections visualized at 20x magnification. Bars=50μm. A=A-7; B=D-6; C=EV (safener treated); D=F-7; E=G-8; F=EV.

4. Discussion

Soils represent the main carbon pool of the global carbon cycle. Photosynthesis enables plants to convert atmospheric CO_2 into carbohydrates (such as starch and cellulose) or into other more stable organic carbon forms such as lignin (Zibilske & Bradford 2007). Next to cellulose, lignin is the second most abundant carbon biopolymer on earth, accounting for an estimated 30% of the organic carbon (C) in the biosphere (Dungait *et al.* 2008). It is known that the abundance, tissue distribution and composition of this important plant cell wall polymer can have an important effect on plant health, as well as agro-industrial processing and carbon sequestration potential (Saballos *et al.* 2009). In fact, the decomposition of lignin in roots and plant residues in soils used for agriculture, forestry and land reclamation has been recognized as a potential option to sequester carbon and mitigate global change by trapping carbon into longer-lived pools (Kumar *et al.* 2006). Furthermore, a high content of polyphenolic compounds, such as lignin, in plant residues can prolong the retention of C in soils (Zibilske & Bradford 2007). Soil organic carbon is an essential component of healthy soils and has been reported to increase the water-holding capability of sandy soil and to improve the structural stability of clay loam soils by helping to form particle aggregates (Zibilske & Bradford 2007). Soil organic carbon is an effective medium for sequestration of inorganic nutrients; it can bind both cations and trace elements that can affect crop growth and yield. This yield enhancement can involve either the direct supply of nutrients to plant root systems, or indirectly alter the physical properties of the soil,

thus improving the root environment and stimulating plant growth (Hati *et al.* 2007). The sequestration capacity of organic carbon in soils is advantageous to plants when it comes to plant stress because roots serve as the proverbial foot soldiers in the plant's battle to survive in an often hostile environment. Roots are the first and most critical plant organ to experience nutrient deficiency, drought, osmotic and ionic stress, soil salinization, heavy metal accumulation and pathogen interactions. In response to these various stresses, plants undergo physiological and metabolic changes underpinned by alterations in gene expression that produce, among other things, complex mixtures of biologically active secondary metabolites involved in important processes such as cellular protection and ion homeostasis (Jones *et al.* 2008). For example, the production of secondary metabolites via the phenylpropanoid pathway provides intermediates for the synthesis of UV protectants (flavonols), defense compounds (isoflavonoids), pigments (anthocyanins/flavonols), nodulation inducers (flavones) and lignins (monolignols) (Kumar *et al.* 2006; Nessler 1994).

As a model for engineering increases in soil carbon stocks (if implemented in a widely planted crop system), I proposed to create transgenic *Arabidopsis* plants with the ability to produce enhanced levels of lignin in their roots. To engineer transgenic plants with a desired phenotype, such as enhanced root lignin, the choice of promoter is a crucial factor. Strong promoters are needed for effective transgene expression in plant cells, but regrettably, most of the widely used constitutive gene expression systems, like the 35S promoter from the Cauliflower Mosaic Virus

(CaMV35S), can produce undesirable pleiotropic effects due to spatially and/or temporally inappropriate ectopic gene expression patterns (Yoshida & Shinmyo 2000). For my project, it was desirable to restrict transgene expression exclusively to root tissues. So far, only a handful of root-specific gene promoters have been identified in plant species such as *Arabidopsis*, rice, tomato and tobacco. However, these promoters are often limited in their applicability due to: a) restricted activity in specific developmental stages, regions or tissues within the root structure, b) to undesirable effects of biotic and abiotic factors on their regulation, or c) to a requirement for specific growth conditions (Jones *et al.* 2008). Genes controlling cell fate in *Arabidopsis* in 15 different root zones (endodermis, endodermis and cortex, epidermal atrichoblasts and lateral root cap) that relate to cell types and tissues at progressive developmental stages (stage 1, 2 and 3) have been previously described (Birnbaum & Benfey, 2004). Data mining of the complete microarray gene expression data set from these studies enabled me to develop my own list of candidate genes whose promoters could be used to drive *SND1* gene expression. These candidate genes were then examined within the Genevestigator microarray database (Hruz *et al.* 2008, https://www.genevestigator.com) for relative expression in roots compared to other plant organs and tissues. Based on these results, one candidate gene, *GSTU19,* was selected for further analysis.

The second candidate gene, *4CL1*, was selected based on previous studies that reported high levels of *4CL1* gene expression in seedling roots. Specifically, transgenic *Arabidopsis* plants containing the *4CL1* or *4CL2* promoter fused to the

beta-glucuronidase (GUS) reporter gene showed developmentally regulated GUS expression in the xylem tissues of both the root and shoot, with *At4CL1::GUS* lines showing its highest levels of gene expression in seedling roots (Soltani *et al.* 2006). In order to confirm these *4CL1* results and validate the potential of these candidate promoters to drive transgene expression in the roots, the expression of both candidate genes was checked using semi-quantitative reverse transcription (RT)-PCR in flower, leaf, stem and roots of four-week-oldplants. *4CL1* showed only a negligible increase in gene expression in the roots compared to other tissues but was retained as a candidate based on earlier studies of the organ-specific expression pattern of *4CL1*, which detected the highest *4CL1* mRNA levels in 3-day-old seedling roots and in bolting stems of mature plants (Soltani *et al.* 2006; Ehlting *et al.* 2002). This difference in gene expression patterns among plant organs, and among these organs at different stages of development, suggests that *4CL1* may exhibit some root-specificity but only at a given point in the plant's growth cycle. As a side note, the *4CL1* promoter was an attractive candidate due to its active involvement in channeling carbon flow into branch pathways of phenylpropanoid metabolism.

This RT-PCR analysis also demonstrated that *GSTU19* is in fact expressed at a noticeably higher level in roots compared to other organs. These results were consistent with the Genevestigator heat map profile as well as with previous studies showing that, under control conditions, expression of *GSTU19* mRNA was higher in roots than in shoots (DeRidder & Goldsbrough 2006). In summary, the data shown for the expression of *4CL1* and *GSTU19* in different plant organs along with evidence

from the previous studies mentioned, suggests that the promoters from these genes could both be good candidates to drive *SND1* transgene expression in roots, but for different reasons. Based on my results from the RT-PCR analysis, the *4CL1* promoter does not seem to be a good candidate for driving root-specific expression but conversely, *GSTU19* does seem to have the potential to drive root-specific expression, which is further supported by this promoters ability to be induced in a root specific manner when treated with the herbicide safener benoxacor.

The ability to turn on gene expression both spatially and temporally offers the opportunity to fine-tune ectopic gene expression without compromising the viability of the organism or the function of the organ being altered, in this case the roots. Plant promoters that impart root-specific expression are of interest for improving tolerance to abiotic stresses such as drought and salinity, for engineering pathogen resistance and for improving nutrient uptake (Vijaybhaskar *et al.* 2008), due to their potential to express recombinant proteins, such as the Cry toxins, in the root (Nitz *et al.* 2001; Vijaybhaskar *et al.* 2008; Maizel & Weigel 2004). Interestingly, a considerable number of root-specific promoters have been characterized, including: *Pyk10* from *Arabidopsis thaliana* (Nitz *et al.* 2001), a glycosyltransferase gene (At1g73160) from *Arabidopsis thaliana* (Vijaybhaskar *et al.* 2008), the *PHT1* gene from *Arabidopsis thaliana* (Koyama *et al.* 2005), the mannopine synthase 2' (mas2') promoter from *Agrobacterium tumefacians* (Ni *et al.* 1996), the iron deficiency specific clone no. 2 (IDS2) promoter from barley (Kobayashi *et al.* 2003), putrescine N-methyltransferase (*PMT*) gene (Mizusaki *et al.*

1971) and *TobRB7* promoter from tobacco (Yamamoto *et al.* 1991) and *SlREO* gene from *Solanum lycopersicum*. Despite these examples, strong root-specific promoters (i.e. promoters that provide for a high level of gene expression) that can be used for various crop improvements are still thought to be limited (Cai *et al.* 2007). Indeed, when I examined the expression level of the so-called "root-specific" *Arabidopsis* promoters (mentioned above) within the Birnbaum *et al.* (2003) data set used to identify *GSTU19*, I found that their relative probe intensity values within the stele and endo-cortex, fell below my chosen cutoff of 1500-5000. They were therefore excluded from this project, but that does not mean they should be rejected as candidate promoters to drive root-specific transgenes in general. Further studies could test the strength of these promoters experimentally by quantifying the GUS activity expressed in promoter::GUS transgenic lines. Moreover, when considering the use of these promoters for genetic engineering, it may be important to determine (via the data set in Birnbaum *et al.* (2003) or by promoter::GUS expression patterns) in which tissues these promoters are predominantly expressed, so that their usefulness to drive transgene expression can be assessed in the context of particular biological questions and objectives.

The identification of the afore-mentioned root-specific promoters from the primary literature, along with the various other candidates that I screened, raises an interesting question: What makes a promoter root-specific? The answer to this question remains somewhat inconclusive, but there is some evidence suggesting that gene expression is determined, at least in part, by motifs or *cis*-elements, within

the promoter sequence of regulated genes (Cai *et al.* 2007). In plants, distinct *cis*-regulatory elements have been linked to specific responses to various treatments, and analysis of the associated DNA sequence motifs has resulted in the elucidation of a number of promoter sequence motifs related to stress responses, developmental and organ-specific regulation (Ma & Bohnert 2007). The characteristics of some of these root-specific *cis*-acting regulatory DNA elements are summarized in Table 5. In my *in silico* analysis of the *4CL1* promoter using the PLACE (Plant Cis-acting Regulatory DNA Elements) database, almost all of the *cis*-regulatory motifs mentioned in Table 5 were present, with the exception of the ACGTROOT1, TELO and SORLIP1AT elements. Similarly, my analysis of the *GSTU19* promoter showed that it contained almost all the motifs with the exception of the ACGTROOT1 and TELO elements. These findings suggest that these elements could play a role in conferring the root-specificity previously described for these genes, albeit at different stages of plant growth and development. For this reason, I chose the largest possible promoter region sequences for my constructs that excluded any upstream genes, yet included as many of the putative root-specific regulatory elements as possible.

Table 5. Summary of *cis*-acting regulatory DNA elements associated with root-specific gene expression.

Putative root-specific element	Sequence	Description	Reference
ARFAT	TGTCTC	ARF binding site found in the promoters of primary/early auxin response genes of Arabidopsis thaliana.	(Inukai et al. 2005)
ASF1MOTIFCAMV	TGACG	A xenobiotic stress-activated transcription factor that binds to the TGACG motif and is expressed preferentially in root apical meristems.	(Klinedinst *et al.* 2000),
OSE1ROOTNODULE	AAAGAT	A consensus sequence motif of organ-specific elements characteristic of activated promoters found in the infected cells of root nodules.	(Vieweg *et al.* 2004)
OSE2ROOTNODULE	CTCTT	A consensus sequence motif of organ-specific elements characteristic of activated promoters found in the infected cells of root nodules.	(Vieweg *et al.* 2004)
RAV1AAT	CAACA	Binds specifically to DNA with bipartite motifs of RAV1-A (CAACA) and RAV1-B (CACCTG). Expression levels of RAV1 were reported to be high in rosette leaves and roots.	(Kagaya *et al.* 1999)
ROOTMOTIFTAPOX1	ATATT	Motif found in rolD promoters. The rol A, B, C and D genes have been identified as the main determinants of the hairy root disease caused on dicots by *Agrobacterium rhizogenes* (Bettini *et al.* 2003).	(Elmayan & Tepfer 1995)
SORLIP1AT	GCCAC	Sequences Over-Represented in Light-Induced Promoters (SORLIPs) in Arabidopsis. Over-represented in light-induced cotyledon and root common genes and root-specific genes.	(Jiao *et al.* 2005)
SP8BFIBSP8BIB	TACTATT	A nuclear factor that binds to the 5′ upstream regions of three different genes coding for major proteins of sweet potato tuberous roots.	(Ishiguro & Nakamura 1992)
SURECOREATSULTR11	GAGAC	Core of sulfur-responsive element (SURE) found in the promoter of SULTR1;1 high-affinity sulfate transporter gene in Arabidopsis. SURE contains auxin response factor (ARF) binding sequence (GAGACA)	(Maruyama-Nakashita *et al.* 2005)
TELO	AAACCCTAA	Found in the *Arabidopsis* eEF1A A1gene promoter as well as in the 5′ region of genes encoding components of the translational apparatus. Implicated in the activation of gene expression in root primordia and root meristems.	(Tremousaygue *et al.* 1999)
WUSATAg	TTAATAG	Target sequence of WUS in the intron of AGAMOUS gene in Arabidopsis. WUSCHEL-type homoebox gene that is specifically expressed in the central cells of a quiescent center in the root apical meristem.	(Kamiya *et al.* 2003)

I chose to use a chemical-inducible system to turn on gene expression of *SND1* at a specific time point in order to avoid the possible negative effects of constitutive gene expression. The benoxacor-inducible system used to induce *SND1* expression from the *GSTU19pro-SND1* construct, offers an advantage over other available chemical-inducible gene expression systems. My results showed that the *GSTU19* promoter was already root-specific in its expression and that this expression could be further induced by the herbicide safener causing an additional increase in gene expression within that organ. These results were confirmed by the transcriptional analysis of *SND1* and its downstream targets in T3 generation transgenic plants, which caused a marked increase in gene expression preferentially in roots.

These results show that when driven by the *GSTU19* promoter, benoxacor may in fact be an excellent inducer of transgene expression but there are some important points to consider (such as induction time, concentration and application methods) when examining the potential of this safener-induction system to be used in root-specific crop biotechnology applications. The use of herbicidal safeners as chemical-inducible gene expression systems in *Arabidopsis*, was previously examined by De Veylder *et al.* (1997) who expressed the *In2-2* promoter from maize in *Arabidopsis* and induced its expression by treatment with benzenesulfonamide herbicide safeners. Similar to later studies done on the induction of GSTs in *Arabidopsis* by herbicide safeners (DeRidder & Goldsbrough 2006; DeRidder *et al.* 2002), GUS staining of the *In2-2* transgenic lines was visible exclusively in the root as soon as 24 hours after induction. In addition, the authors conducted a time-course experiment

on two-week-old *In2-2* transgenic plants containing the GUS reporter gene by transferring seedlings from safener-free media to media containing safener (50mg/L). After transfer of the plants back to safener-free medium, they found that GUS staining disappeared within three days, indicating a strong correlation between the presence of the safener and In2-2 expression. They also found that prolonged induction by safeners (at a concentration of 50mg/L) resulted in inhibition of root growth, indicating that the amount of time the plant was exposed to the chemical at that concentration was critical. Therefore, the majority of studies involving herbicide safeners use an induction time of 24 hours. It was not immediately clear in the literature why the standard induction concentration now used among most research groups for herbicide safeners is 100μM but it appears that this concentration is thought to serve as an "antidotally effective amount" (Mccutchen *et al.* 2008) that is the amount that should be added to an herbicide formulation in order to eliminate or reduce the phytotoxic effects of the herbicide to certain crops.

Although studies have suggested that herbicide safeners could be potentially useful as a tissue-specific transient expression system where inducible transcription in the root is required, there have been no studies reported where this system had been optimized with respect to safener concentration and time of induction within the context of driving transgene expression. In addition to the time of induction and concentration of reagent, the type of application method may be an important component of a safener-inducible gene expression system. For example, previous studies have shown that adding the safeners to hydroponically grown plants

resulted in consistent induction patterns among all safeners tested, whereas, foliar application did not induce any GUS activity (De Veylder *et al.* 1997). Later studies using three-week-old Arabidopsis plants treated with safeners (100μM) by foliar application required treatment once per day for four consecutive days to achieve the desired level of gene induction (DeRidder & Goldsbrough 2006). These results provide some insight into the efficacy of a particular application method with respect to the time of induction of the inducible promoter. Absorption of the safener via the roots seems to result in a much faster and more direct induction whereas to achieve similar results via foliar application longer exposure to the inducer at similar concentrations is required. Further studies are needed to optimize this system if safeners are to be more widely used as root-specific chemical induction systems.

The reverse transcription PCR analysis in flower, leaf, stem and roots also detected *AtSND1* expression exclusively in stems of four-week-oldplants. Given that the lignin biosynthetic pathway seems to be regulated by a network of TFs, such as *SND1,* it is important to consider the implications of introducing a regulatory gene into an environment in which it is usually not expressed. Previous studies have shown that, in roots, the expression level of a cohort of TF genes working downstream of *SND1*, as well as of *SND1* itself, was largely restricted to the developing secondary xylem but this expression was at very low levels compared to their expression in stems (Zhong 2008). At the outset of this project, it was not known how root-specific overexpression of *SND1* might affect secondary cell wall

thickening in roots, or if the regulatory network activated by *SND1* would function the same way in this organ as it does in stems. It is possible that transcriptional activators, such as SND1 and its downstream targets, might be able to regulate secondary cell wall formation in non-sclerenchymatous tissues of the growing plant by acting as repressors of gene expression in order to prevent any pleiotropic effects associated with the ectopic expression of genes controlling and involved in lignin biosynthesis. However, only a limited number of expression repressors have been identified in plants thus far.

Secondary wall formation is a highly coordinated process that results from the subsequent deposition of cellulose, hemicelluloses and lignin as soon as primary cell growth has ceased. The proportion of each of these major components is highly variable depending on the climate, geographic location, species, age and part of the plant. Knowledge of how the coordinated regulation of genes leading to secondary cell wall formation and how this regulation leads to the relative composition of the main constituents, is still growing (Ko *et al.* 2009). However, there are still some gaps in our understanding and as a result, it was difficult to predict how *SND1* overexpression would influence lignin deposition in roots, a tissue in which only low levels of the TFs involved in regulating secondary cell wall formation have been previously described (Zhong *et al.* 2008). I created two different root-specific overexpression constructs (*4CL1pro-SND1* and *GSTU19pro-SND1*) in *Arabidopsis* and results from the transcriptional analysis of *SND1* gene expression in T2 generation plants confirmed that *SND1* was indeed overexpressed in the roots in almost all of

the lines analyzed within each overexpression construct. Transcriptional analysis of *SND1* in T3 generation transgenic plant lines, however, showed overexpression in both roots and shoots compared to empty vector control lines, indicating transgene expression was observed in both tissues and that expression in the roots was only slightly higher in shoots. The promoters selected to drive transgene expression (*4CL1pro* and *GSTU19pro*) are not necessarily "root-specific" in the sense that their native expression pattern indicate that they are expressed elsewhere in the plant, which may be why *SND1* was seen to be overexpressed in shoots as well as roots in transgenic plants. On the other hand, given that the native expression analysis in different plant organs in addition to the data obtained from Genevestigator, showed that *SND1* was expressed somewhat exclusively in stems, the fact that overexpression of *SND1* was seen in roots of transgenic lines indicates that the promoters are functioning in their ability to drive expression of the transgene in roots, albeit not in a comparatively restricted manner.

SND1 has been previously shown to upregulate the expression of several transcription factors that are highly expressed in fibres during secondary cell wall biosynthesis (Zhong *et al.* 2006). Therefore, it was not surprising that my results indicated an increase in gene expression (specifically in roots) of the transcription factors acting downstream of SND1 (MYB46, SND3, MYB103 and KNAT7). Given that these transcription factors have been previously shown to be expressed at very low levels in roots (Zhong *et al.* 2006; Zhong *et al.* 2008), my data further confirms that the *4CL1pro-SND1* and *GSTU19pro-SND1* constructs are behaving in a root-

specific manner. These results correlate with the previously characterized hierarchical organization of these transcription factors acting as direct targets of SND1, therefore it seems as though the interactions previously described in aerial tissues, behave in a similar fashion in root tissues (Ko *et al.* 2009; Zhong *et al.* 2008).

However, there is still much that we do not know about the organization, association and interrelation of the entire regulatory cascade involved in the activation and regulation of lignin biosynthetic genes during secondary cell wall formation in stems, let alone in the roots. This could be problematic when trying to determine and interpret what is happening downstream of these master transcriptional switches, such as SND1 and MYB46, and how the lignin biosynthetic pathway is being specifically altered in roots of transgenic plants, an environment within which these TFs do not normally operate. The growing amount of data (and many different interpretations of this data) being generated and subsequently presented in the literature is usually studied within stems and leaf protoplast and is often confusing and sometimes conflicting. Further studies are needed to characterize all the putative TFs involved in regulating secondary cell wall formation, in addition to studies aimed at determining associations between these factors and with biosynthetic genes. These studies should clarify some of the missing links in our current knowledge, at least within aerial tissues. Significantly more work would be required in *Arabidopsis* root systems in order to determine the effects of overexpressing regulatory factors involved in secondary cell wall formation in tissues not normally heavily lignified. This is an important

consideration for future attempts at inducing hyper-lignification in *Arabidopsis* root systems, before attempts can be made at increasing soil carbon stocks in a large-scale crop system through similar approaches and methods.

The genes involved in cellulose, xylan, and lignin biosynthesis need to be turned on in order to make lignified secondary cell walls in *Arabidopsis*. The RT-PCR analysis of three phenylpropanoid pathway enzymes leading to the production of monolignols (4CL1, CCR and COMT) showed no observable difference in gene expression among these lignin biosynthetic genes or among tissue types (root and shoot). Several possibilities could explain this finding, despite the overexpression of *SND1* and its direct targets: (i) they are not involved in the transcriptional control of these particular lignin biosynthetic genes, (ii) they require the involvement of other transcription factor(s) to function, or (iii) they are not directly involved in secondary wall formation (Ko *et al.* 2009). The first explanation could certainly be true where SND3 and MYB103 are concerned, since they were recently shown to induce the GUS reporter gene expression driven by the *CesA8* promoter, from a cellulose synthase gene required for cellulose synthesis during secondary cell wall formation (Zhong *et al.* 2008). This proves that SND1 is involved in regulating certain genes involved in other aspects of the secondary cell wall biosynthetic program, in addition to that of lignin. On the other hand, the MYB46 transcription factor was shown to be a direct target of SND1 and both TFs were previously shown to be capable of turning on a whole set of genes involved in secondary wall synthesis in general (Zhong *et al.* 2007b; Ko *et al.* 2009). Therefore, it is puzzling that overexpression of this gene did not activate key lignin biosynthetic enzymes in

either the root or shoot, where previous studies have shown this to occur. For example, *MYB85* gene expression was previously shown to be upregulated by *MYB46* overexpression, and MYB85 was shown to be able to induce expression of the GUS reporter gene, when driven by the *4CL1* promoter in leaf protoplasts. Studies have shown that overexpression of *MYB85* led to ectopic deposition of lignin in epidermal and cortical cells in stems (Zhong *et al* 2007a; Zhong *et al.* 2006; Zhong *et al.* 2008). Therefore, since MYB46 has been shown to be a direct target of SND1, I am unable to explain (within the current model of this *SND1*-mediated regulatory network) why the overexpression of *SND1* and *MYB46* did not specifically cause the *4CL1* gene to be turned on through induction of the *4CL1* promoter by MYB85. The gene expression level of *MYB85* was not examined in the roots of my transgenic lines, therefore transcriptional analysis of this gene by RT-PCR, could provide further information into determining why the *4CL1* gene was not turned on in response to *SND1* overexpression. Nevertheless, there are several other transcription factors that have been previously reported to regulate secondary wall biosynthesis including, KNAT7, MYB52, MYB54, MYB58 and MYB63. KNAT7, for example, was overexpressed in the roots of my transgenic plants but did not seem to influence the secondary lignin biosynthetic genes tested here. This could be because KNAT7 is not involved in activating lignin biosynthetic genes directly. To test this theory, the characterization of KNAT7 using reverse genetics approaches along with the yeast two-hybrid system for determining protein interactions, may provide some insight into its specific function and interacting partners. In short, TFs in general have diverse roles in regulating gene transcription. For instance, they may

act as part of a complex with other TFs or regulatory proteins, which together might be involved in directly regulating gene expression in a particular biosynthetic pathway. Others, however, might be involved in enhancing or fine-tuning the level of expression of different metabolic pathway genes (Zhong *et al.* 2008).

Interest in lignin biosynthesis and lignin deposition is mainly due to the extensive involvement of lignin in plant biology (Boudet *et al.* 2003; Humphreys & Chapple 2002). Lignin can be defined two ways: 1) from a chemical point of view (chemical composition and structure) or 2) from a functional point of view (what lignin does within the plant) (Hatfield & Fukushima 2005). Regardless of these definitions, it is important to be able to determine the concentration of lignin within a broad assortment of cell wall varieties. One would think that lignin would be relatively easy to measure, given that it is somewhat resistant to both chemical and biological degradation. However, there have been several methods and techniques that have been developed and adapted throughout the years to quantitatively determine total lignin content and composition in different types of plant samples, yet not one of them has been deemed as a standard clear-cut method for all samples (Hatfield & Fukushima 2005). Worth mentioning, however, for the determination of lignin content in plant samples, are non-invasive approaches such as: near infrared spectroscopy (NIRS) and nuclear magnetic resonance spectroscopy (NMR). These methods of lignin content determination offer an advantage over more invasive methods in that they ultimately leave the lignin in the sample chemically unaltered (Hatfield & Fukushima 2005). Alternatively, two procedures (thioglycolate and

acetyl bromide) rely on the solubilization of lignin in an appropriate solvent whereby the lignin in solution can be measured (Hatfield & Fukushima 2005). Lastly, various methods have been proposed using mineral acids to solubilize and hydrolyze carbohydrates leaving the lignin residue to be measured and determined gravimetrically, such as the Klason lignin method (Hatfield & Fukushima 2005). Is seems that the most commonly used method for determining lignin is the Klason lignin or 72% (v/v) H_2SO_4 acid procedure.

Given the relatively low amount of lignin present in roots to begin with, as well as the limited amount of root material available working in the *Arabidopsis* system, it was important to be prudent and judicial with the choice of lignin content determination method. Results from the acetyl bromide analysis of soluble lignin content, showed a marked decrease in total lignin content (~40-50%) in both *GSTU19pro-SND1* lines but only one of the *4CL1pro-SND1* lines. The Klason lignin analysis supported the data obtained from the acetyl bromide-based method by confirming a decrease in insoluble lignin content in the *GSTU19pro-SND1* overexpression line. On the other hand, one *4CL1pro-SND1* line in each of the lignin content analyses showed negligible changes in lignin content. Nevertheless, I reasoned that the decreases seen in lignin content in the majority of lines analyzed (and also lack of change in gene expression of indicative lignin biosynthetic enzymes), could be a result of carbon reallocation to a different area of carbon metabolism (such as production of cellulose and starch). Results from the cellulose and starch content analyses, disproved this theory by showing a similar decrease in

cellulose and hemicellulose content in the *GSTU19pro-SND1* line analyzed. I should note, that results from the cellulose, starch and Klason lignin analyses were absolute values from a single biological replicate making the data somewhat unreliable, however, given that they correlate to certain degree with results seen using the acetyl bromide-based method, I have included them in this thesis.

An overall trend of decreased cell wall composition (lignin, cellulose and hemicellulose) was seen in both *SND1* overexpression constructs analyzed. Previous studies have shown that although *SND1* overexpression induces ectopic secondary wall deposition in cells that are normally not lignified, excess SND1 apparently inhibits normal secondary wall thickening in fibres (Zhong *et al.* 2006). In these studies, *SND1* overexpression was seen to induce secondary cell wall production in many parenchyma cells in leaves and floral organs as well as epidermal cells in stems; however, ectopic secondary wall deposition was seldom seen in the parenchyma cells of other organs. Moreover, *SND1* overexpressors showed that ectopic secondary wall thickening was rarely observed in the epidermal cells of hypocotyls and cortical cells of roots but was not seen in other root cell types (Zhong *et al.* 2006). This finding suggests that different cell types in different organs might exhibit differential competence to induction of secondary wall thickening by SND1. Equally, this differential induction could be a case of substrate availability, meaning that the required precursors for monolignol biosynthesis may be in short supply in the root, given that *Arabidopsis* root tissue does not normally contain high levels of lignin. Either way, any of these reasons

could account for the fact that overexpression of *SND1* in the root-specific (*4CL1pro-SND1*) and inducible (*GSTU19pro-SND1*) constructs did not result in the direct activation of lignin biosynthetic genes as demonstrated using the RT-PCR analysis, which should have resulted in an increase in root lignin content instead of the observed decrease in lignin content.

As for the other cell wall constituents analyzed, Ko *et al.* (2007) reported that cellulose compositions of the cell wall were decreased in the inflorescent stems and roots of plants overexpressing *SND1* driven by the CaMV35S promoter, most likely resulting from defects in xylary fibre formation. However, my results showed an increase in the relative gene expression of *SND3* and *MYB103* as seen in the transcriptional analysis of T3 generation transgenic plants, which should have resulted in an increase in cellulose content given that these downstream targets of SND1 were recently shown to induce the GUS reporter gene expression driven by the *CesA8* promoter (Zhong *et al.* 2008). Instead, I generally observed a decrease in cellulose and hemicellulose content, similar to that observed for lignin.

Overexpression of mRNA can sometimes lead to a drastic reduction in the level of expression of the endogenous genes concerned, i.e. host genes can be silenced as a consequence of the presence of a homologous transgene, thus limiting the potential application of genetic transformation; a phenomenon called co-suppression (Vaucheret *et al.* 2001). One way of understanding this phenomenon is that when RNA transcripts accumulate beyond a critical threshold, they are selectively

degraded by ribonucleases (RNases), a type of nuclease that catalyzes the degradation of RNA. An accumulation of elevated levels of mRNAs might lead to the production of abnormal sense RNA transcripts of the transgene (Vaucheret *et al.* 2001) and accumulation of these anomalous RNA transcripts is proposed to activate the RNA-dependent RNA polymerase, which transcribes the RNA transcripts to produce antisense RNA. The antisense RNA transcripts then bind to the accumulated normal and abnormal RNA transcripts of the transgene as well as the endogenous gene, producing RNA duplexes that are then targeted by double-stranded RNA specific RNases. This often leads to a radical reduction in the level of transgene expression as well as the expression of the endogenous gene and sometimes homologous genes as well. This series of events are collectively referred to as gene silencing and are defined by predominantly taking place at the post-transcriptional level, where RNA does not accumulate even though transcription occurs (Vaucheret *et al.* 2001). The degradation of RNA via gene silencing may be why I observed a decrease in total lignin content in my transgenic plants but it does not explain why I observed an increase in mRNA transcripts in my transgenic lines. One reason for this could be because the transcriptional analysis was performed on three-week-old plants and the lignin analysis was performed on mature plants that were roughly eight-weeks old. It is possible that gene silencing is occurring during secondary cell wall formation in plants that are older then three-weeks. Using RT-PCR to analyze SND1 and its targeted TFs in the root tissue from transgenic lines at different developmental stages could test this hypothesis.

An alternative possibility for this observed difference in increased mRNA versus decreased lignin could be some kind of regulation at the translational level, but the mechanisms for this type of control are poorly understood. It is clear that there are many mechanisms in place to control and maintain normal levels of plant cell wall constituent biosynthesis and deposition. This presents a significant challenge to overcome when designing and engineering genetic constructs for crop improvement. As more knowledge is gained regarding the mechanisms that regulate transcription of secondary cell wall components as a whole (as well as the coordinated expression of the cohort of transcription factors and proteins regulating the lignin biosynthetic pathway), we will undoubtedly be able to gain new insight that will help us to develop more complex and fine-tuned gene expression systems that could complement or counteract any other regulatory mechanisms present that may prevent us from achieving the desired end-product or phenotype.

It is imperative to the process of genetic engineering for agricultural purposes to drive transgene expression in a manner that evades health costs to the plant caused by the constitutive expression of target genes. It was therefore important to survey a variety of different key plant physiological traits that could have a dire impact on the efficacy of crop production. Overexpression of key genes involved in normal plant growth and development, such as secondary cell wall pathways, could be implicated in normal agricultural activities such as seed production and crop yield, thus resulting in major economic consequences if altered inappropriately. The phenotypic analysis of seed-related traits revealed that overexpression of SND1 did

not cause any undesirable pleiotropic effects in seed production and/or viability among my transgenic plants. Given that *GSTU19pro-SND1* lines were induced by herbicide safener at four-weeks post-flowering, it was expected that these lines would not result in a phenotype involving seed-related traits. This data supports my previous analyses showing that both construct promoters were shown to be root-specific, which means that *SND1* overexpression in the roots should not activate gene expression of secondary cell wall biosynthetic genes in seeds.

To determine whether overexpression of *SND1* in roots caused any variation in root architecture, lateral root density (LRD) was analyzed for two lines in each constructs and showed an increase in LRD in three out of the four lines analyzed. It was interesting that the transgenic lines showing an increase in LRD were the same transgenic lines corresponding to a decrease in total lignin content, as seen in the chemical lignin analysis. It has been previously shown, that SND1 is a member of the NAC domain protein family, which comprises approximately 100 genes in the *Arabidopsis* genome and function as plant-specific transcriptional factors. To date, only a small number of NAC domain genes have been characterized and NAC domain proteins have been implicated in a wide variety of processes, including the establishment of the shoot apical meristem, the signaling pathway involved in abiotic stress, defense responses and lateral root formation. Specifically, *AtNAC1* (At1g56010) has been shown to mediate auxin signaling and promote lateral root formation (Xie *et al.* 2000). A multiple sequence alignment of various NAC domain genes, from a previous study, has shown that *AtNAC1* is a distant relative of SND1 (Zhong *et al.* 2006). Plant roots have a distinct organization that is fundamental to

the formation of lateral roots. The outer tissues of dicot plant roots (epidermis, cortex, and endodermis) are organized into separate concentric layers whereas the vascular tissues of the central stele have a more bilateral symmetry (Parizot *et al.* 2008). The outermost layer of the stele, known as the pericycle, is composed of two different cell types: one subset is associated with the xylem, whereas the other is associated with the phloem. The former has the strong capability to initiate cell division but the latter appears to remain inactive (Parizot *et al.* 2008). The formation of lateral roots is a result of a subset of pericycle cells (called the pericycle founder cells) that are positioned at the xylem poles within parent root tissues. Subsequently, the mature pericycle cells form lateral root primordium (LRP) via dedifferentiation, which then undergoes consistent cell divisions to generate a well-organized LRP. Cell expansion causes the LRP to emerge from the parent root, and the lateral root meristem becomes activated resulting in continued growth of the lateral root (Lee *et al.* 2009). The positioning of the pericycle founder cells to the xylem poles may provide a testable hypothesis regarding the decrease in lignin and increase in LRD in transgenic lines overexpressing *SND1* and its downstream target *MYB46*. Previously, ectopic secondary wall thickening in the parenchymatous cells of leaves, floral organs and inflorescence stems was seen in *MYB46* overexpressors (Ko *et al.* 2009). In addition, *SND1* overexpression showed a small increase in the wall thickness of vessels (Zhong 2006). Although I did not specifically look at ectopic secondary wall thickening in my transgenic lines, a possible increase in wall thickening due to *SND1* and *MYB46* overexpression in root xylem vessels may have caused a movement in auxin pools near the xylem poles, causing lateral roots to

form. One way to test this hypothesis would be to: i) to confirm that secondary wall thickness was in fact perturbed and ii) to transform the *SND1* root-specific overexpression constructs (*GSTU19pro-SND1* and *4CL1pro-SND1*) with the promoter-marker gene fusion DR5::GUS activated by auxins to visualize auxin response patterns in the root.

Another possible explanation for this increase in lateral root density found in the *GSTU19pro-SND1* transgenic lines in particular, could be due to the fact that, in addition to high levels of expression in the stele and endo-cortex, *GSTU19* was found to have an even higher level of expression in lateral root cap tissues as seen by the Genevestigator heat map that I generated and the relative probe intensities from the Birnbaum and Benfey dataset (2004). The root cap has been shown to be a complex and dynamic plant organ. Root caps are responsible for sensing and transmitting environmental signals, synthesizing and secreting small molecules and macromolecules, and in some species shedding metabolically active cells (Tsugeki & Federoff 1999). One study reported the identification and use of a root cap-specific promoter to genetically destroy root caps by directing root cap-specific expression of a diphtheria toxin A-chain gene. The roots of these transgenic plants had more highly branched lateral roots than those of wild-type control plants. Root cap ablation (where individual cells are destroyed for experimental purposes) in this study was shown to alter root architecture both by inhibiting root meristematic activity and by stimulating lateral root initiation. These observations implied that

root caps contain essential components of the signaling system that determines root architecture (Tsugeki & Federoff 1999).

If *SND1* overexpression in *GSTU19pro-SND1* lines caused a similar ablation or alteration in lateral root caps this could certainly explain the observed increases in LRD seen among the two transgenic lines analyzed for this particular construct. One way to test this hypothesis would be to visualize longitudinal sections of primary root tips using electon or confocal microscopy in order to determine the differentiation of root cells in my transgenic and empty vector control plants. Another interesting observation from Tsugeki & Federoff (1999) was that despite the abnormal root structure of their transgenic lines, the appearance of the aerial parts of the transgenic plants was normal on both MS agar medium and in soil. The normal aerial phenotype was also observed in my transgenic lines, including those showing increased lateral root formation and a decrease in total lignin content. According to Tsugeki & Federoff (1999), these results could indicate that the formation of more lateral roots might compensate for the effect of the short-root phenotype seen in previous studies involving the *SHORT-ROOT* (*SHR*) gene, which is typified by the absence of gravitropic response in shoots and exhibits a determinate root growth pattern (Benfey *et al.* 1993).

Multiple signaling pathways are responsible for controlling normal plant growth and development. These pathways are able to integrate information from the environment using metabolic and developmental signals. If these normal developmental and signaling pathways, such as the phenylpropanoid pathway, are

disrupted or altered, consequences to overall plant growth and function could result. In order to determine any phenotypes involving flowering time, overall height and shape (leaf and plant), plant growth was examined over a six-week period. My transgenic plants overexpressing *SND1* did not show any observable phenotype among aerial plant tissues, which could mean that: a) my transgenic constructs were sufficiently root-specific that overexpression of *SND1* in roots did not seem to interefere with normal plant growth and development or b) there is no alteration in secondary cell wall composition that could cause an observable phenotype in aerial tissues. Either of these reasons could explain why the pendulous phenotype (as well as other severe phenotypes in flowers and leaves), previously seen in *SND1* overexpressors under the control of the constitutive CaMV35S promoter, was not observed in my transgenic plants (Zhong *et al.* 2006).

Histochemical staining and UV autofluorescence of lignin in root-hypocotyls did not show significant visible phenotypic changes even though considerable variation in lignin content was seen along the 5mm sections of hypocotyl analyzed, which could be due to differences in developmental equivalencies. It is possible that visualizing wall thickness and lignin content at this magnification using this particular type of microscopy, was not sufficient to observe any changes in cell wall thickness or lignin deposition patterns in the roots. A more sensitive method might be needed to distinguish more subtle differences in cell wall thickness among transgenic lines, such as transmission electron microscopy.

5. Conclusions and Future Directions

To my knowledge, this is the first investigation into the manipulation of lignin deposition in *Arabidopsis* roots for the end-use of increasing carbon stocks in agricultural root systems, such as canola or soybean. Using a metabolic engineering approach, SND1, a key transcriptional activator controlling secondary cell wall biosynthesis and deposition in *Arabidopsis*, was identified as a suitable candidate gene to alter the expression of several endogenous genes and transcription factors involved in lignin biosynthesis, through overexpression in root tissues. In my transgenic plant lines overexpressing *SND1* in roots *(*driven by two different root-specific candidate gene promoters, *4CL1* and *SND1*), I found that *SND1* overexpression upregulated previously known downstream targets of SND1, did not result in a modification of lignin biosynthetic pathway genes, generally showed a decrease in total lignin and carbohydrate content, showed an increase in lateral root density and did not exhibited any visible phenotypes regarding seed-related traits, plant growth and development, plant height or lignin deposition patterns in roots.

SND1 did not behave in a predictable manner when overexpressed in an environment that it does not normally operate in. There is still much to discover about the organization, association and interrelation of the entire regulatory cascade of TFs (along with regulatory proteins and cofactors) involved in the activation or supression of lignin biosynthetic genes during secondary wall formation in shoots, let alone in the roots. Further studies are underway, in

Arabidopsis, to characterize the TFs involved in the SND1-mediated regulation of secondary cell wall formation (probably through reverse genetics approaches). The mechanisms in place to control and maintain normal levels of plant cell wall biosynthesis and deposition present a significant challenge to overcome when designing and engineering genetic constructs to ectopically express transcription factors that regulate secondary cell wall metabolic pathways, in plant organs where these factors do not normally regulate this process. TFs in general have very diverse roles in regulating gene transcription and may act as part of a complex with other TFs or regulatory proteins, which together might be involved in directly regulating gene expression in a particular biosynthetic pathway. Others might be involved in enhancing or fine-tuning the level of expression of different metabolic pathway genes. Therefore, studies are needed to determine the specific associations between these factors and with cell wall biosynthetic genes (*in vivo, in vitro* and *in planta*) could also provide more insight into how this particular lignin metabolic pathway is controlled as well as possibly present new candidate genes whose overexpression might induce ectopic lignification in root tissues. All these studies combined should clarify some of the missing links in our current knowledge of secondary cell wall formation, within above ground tissues. Significantly more work is required in *Arabidopsis* root systems to determine how (and even which) secondary cell wall regulatory factors operate in these tissues. These studies may even elucidate new candidate genes controlling lignin deposition specifically in roots.

Suitable promoters and a safener-inducible gene expression system were identified in this project and used to induce root-specific expression in transgenic plants. Experimentally testing the strength and tissue-specificity of all the other putative root-specific promoters that have been previously identified in *Arabidopsis* can be used to assess their ability to drive transgene expression in the context of particular biological questions and objectives. Since *in silico* analysis of regulatory motifs or *cis*-elements in promoter regions indicates that these binding sequences could play an important role in conferring root-specificity, as previously described for these so-called "root-specific" genes, it may be valuable to determine which of these putative motifs are in fact directly linked to root-specific gene expression by more direct experimental approaches. For example, recapitulation studies using intact and mutated versions of the predicted *cis*-element driving a reporter gene (such as luciferase) in transgenic plants could be used to validate the hypothesized function of the *cis*-acting regulatory element *in vivo*. The additional information gathered from these future studies, could provide us with more ways to fully explore the various gene expression resources available for manipulating lignin deposition in roots, thereby enabling us to develop new highly specific gene expression constructs for enhancing lignin deposition in roots. Furthermore, the root-specific gene induction system in dicots using benoxacor and fenclorim as chemical inducers of the *GSTU19* promoter used to drive root-specific transgene expression, showed some promise in conferring spatial and temporal control of transgene expression in the roots of transgenic plants analyzed in this project but this system needs to be optimized with respect to safener concentration, induction time and application

method if it is to be used widely as an acceptable chemical-inducible root-specific gene expression for various root-related biotechnology applications. For example, direct induction of transgene expression in hydroponically grown plants is the most effective way to induce transgene expression in roots by allowing direct access to the safener by root systems, a method that may not be well-suited to large-scale crop systems.

Studies with the overall aim of modifying lignin content and composition in plants have many potential economic and environmental benefits to humans. As a result of this importance, in just over a decade, a number of studies have been conducted to manipulate gene expression in the monolignol pathway within phenylpropanoid metabolism. For instance, cheaper and more easily processed trees for pulp and paper manufacture that could decrease pollution, more readily digestible forage for livestock and improved feedstock for fuel/chemical production (Anterola & Lewis 2002). These research endeavors, along with high throughput transcriptional and metabolic profiling studies, have produced an immense collection of scientific data. These studies are important in gaining significant insight into: 1) the overall dynamics of phenylpropanoid metabolism (i.e. how carbon flux through various pathways is differentially controlled) and 2) how genetic manipulations can alter and disrupt programmed lignin assembly in a predictable manner without affecting overall plant viability (Anterola & Lewis 2002). In fact, metabolic engineering in general is now beginning to take over from single-gene engineering as the best way to manipulate metabolic flux in transgenic plants. The ability to control several

points in a given metabolic pathway at the same time either by overexpressing and/or suppressing several enzymes through the use of transcriptional regulators controlling endogenous genes is a powerful tool in developing complex phenotypes resulting from modifications of entire pathways. Our knowledge of metabolic pathways continues to expand via the use of applied genomics, proteomics and metabolomics, while advances in systems biology help us to model the impact of different modifications. In conclusion, these more recent biotechnological advances are greatly increasing our understanding of the regulatory processes involved in controlling secondary cell wall biosynthesis and deposition.

Bibliography

About Arabidopsis. (2008, August 5). Retrieved August 9, 2009, from The Arabidopsis Information Resource : http://www.arabidopsis.org

Anterola, A. M., & Lewis, N. G. (2002). Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and cascular integrity. *Phytochemistry, 61* (3), 221-294.

Arabidopsis thaliana. (2009, July 28). Retrieved August 9, 2009, from Wikipedia: http://en.wikipedia.org/wiki/Arabidopsis_thaliana

Arce, A. L., Cabello, J. V., & Chan, R. L. (2008). Patents on plant transcription factors. *Recent Patents on Biotechnology, 2* (3), 209-217.

Baghdady, A., Blervacq, A. S., Jouanin, L., Grima-Pettenati, J., Sivadon, P., & Hawkins, S. (2006). Eucalyptus gunnii CCR and CAD2 promoters are active in lignifying cells during primary and secondary xylem formation in Arabidopsis thaliana. *Plant Physiology and Biochemistry, 44* (11-12), 674-683.

Bais, H. P., Weir, T. L., Perry, L. G., Gilroy, S. & Vivanco, J. M. (2006). The role of root exudates in rhizophere interactions with plants and other organisms. *Annual Review of Plant Biology*, *57*, 233-266.

Battle, M., Bender, M. L., Tans, P. P., White, J. W. C., Ellis, J. T., Conway, T. & Francey, R. J. Global Carbon Sinks and Their Variability Inferred from Atmospheric O_2 and $\delta^{13}C$. *Science, 287* (5462), 2467-2470.

Benfey, P. N., Linstead, P. J., Roberts, K., W., S. J., Hauser, M. T., & Aeschbacher, R. A. (1993). Root development in Arabidopsis: four mutants with dramatically altered root morphogenesis. *Development, 119* (1), 57-70.

Besseau, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B., & Legrand, M. (2007). Flavonoid Accumulation in Arabidopsis Repressed in Lignin Synthesis Affects Auxin Transport and Plant Growth. *The Plant Cell, 19*, 148-162.

Bettini, P., Michelotti, S., Bindi, D., Giannini, R., Capuana, M., & Buiatti, M. (2003). Pleiotropic effect of the insertion of the Agrobacterium rhizogenes rolD gene in tomato (Lycopersicon esculentum Mill.). *Theor.Appl.Genet., 107* (5), 831-836.

Bhuiyan, N., Selvaraj, G., Wei, Y., & King, J. (2009). Role of lignification in plant defense. *Plant Signaling & Behavior, 4* (2), 158-159.

Birnbaum, K., & Benfey, P. N. (2004). Network building: transcriptional circuits in the root. *Current Opinion in Plant Biology , 7* (5), 582-588.

Birnbaum, K., Sasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W., et al. (2003). A Gene Expression Map of the Arabidopsis Root. *Science, 302* (5652), 1956 - 1960.

Black, C. C. (1973) Photosunthetic carbon fixation in relation to net CO_2 uptake. *Annual Review of Plant Physiology*. *24*, 253-286.

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin Biosynthesis. *Annual Review of Plant Biology, 54*, 519-546.

Boudet, A., Kajita, S., Grima-Pettenati, J., & Goffner, D. (2003). Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends in Plant Science, 8* (12), 576-581 .

Boudet, A.-M. (2007). Evolution and current status of research in phenolic compounds. *Phytochemistry, 68* (22-24), 2722-2735 .

Boudet, A.-M. (2000). Lignins and lignification: Selected issues. *Plant Physiology and Biochemistry, 38* (1-2), 81-96 .

Brand, L., Horler, M., Nuesch, E., Vassalli, S., Barrell, P., Yang, W., et al. (2006). A Versatile and Reliable Two-Component System for Tissue-Specific Gene Induction in Arabidopsis. *Plant Physiology, 141*, 1194-1204.

Broun, P. (2004). Transcription factors as tools for metabolic engineering in plants . *Current Opinion in Plant Biology, 7* (2), 202-209 .

Burk, D. H., Liu, B., Zhong, R., Morrison, W. H., & Ye, Z.-H. (2001). A Katanin-like Protein Regulates Normal Cell Wall Biosynthesis and Cell Elongation. *The Plant Cell, 13*, 807-828.

Cai, M. C., Wei, J., Li, X., Xu, C., & Wang, S. (2007). A rice promoter containing both novel positive and negative cis-elements for regulation of green tissue-specific gene expression in transgenic plants. *Plant Biotechnology Journal, 5*, 664-674.

Chambers, J. M., Freeny, A and Heiberger, R. M. (1992) Analysis of variance; designed experiments. Chapter 5 of Statistical Models in S eds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole.

Chang, X. F., Chandra, R., Berleth, T., & Beatson, R. P. (2008). Rapid, Microscale, Acetyl Bromide-Based Method for High-Throughput Determination of Lignin Content in Arabidopsis thaliana. *Journal of Agricultural and Food Chemistry, 56* (16), 6825-6834.

Chapple, C., & Carpita, N. (1998). Plant cell walls as targets for biotechnology. *Current Opinion in Plant Biology, 1* (2), 179-185.

Coleman, H. D., Park, J.-Y., Nair, R., Chapple, C. & Mansfield, S. D. (2008) RNAi-mediated suppression of p-coumaroyl-CoA 3′-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *PNAS*, 105 (11), 4501-4506.

De Graaff, M.-A., Six, J., & Van Kessel, C. (2007). Elevated CO2 increases nitrogen rhizodeposition and microbial immobilization of root-derived nitrogen. *New Phytologist, 173* (4), 778-786.

De Veylder, L., Van Montagu, M., & Inze, D. (1997). Herbicide safener-inducible gene expression in Arabidopsis thaliana. *Plant Cell Physiology, 38* (5), 568-577.

Delmer, D. P., & Haigler, C. H. (2002). The Regulation of Metabolic Flux to Cellulose, a Major Sink for Carbon in Plants. *Metabolic Engineering, 4*, 22-28.

DeRidder, B. P., & Goldsbrough, P. B. (2006). Organ-Specific Expression of Glutathione S-Transferases and the Efficacy of Herbicide Safeners in Arabidopsis. *Plant Physiology, 140*, 167-175.

DeRidder, B. P., Dixon, D. P., Beussman, D. J., Edwards, R., & Goldsbrough, P. B. (2002). Induction of Glutathione S-Transferases in Arabidopsis by Herbicide Safeners. *Pltnat Physiology, 130*, 1497-1505.

Dungait, J. A., Stear, N., van Dongen, B. E., Bol, R., & Evershed, R. P. (2008). Off-line pyrolysis and compound-specific stable carbon isotope analysis of lignin moieties: a new method for determining the fate of lignin residues in soil. *Rapid Communications in Mass Spectrometry, 22* (11), 1631-1639.

Ehlting, J., Büttner, D., Wang, Q., Douglas, C. J., Somssich, I. E., & Kombrink, E. (2002). Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms . *The Plant Journal, 19* (1), 9-20.

Elmayan, T., & Tepfer, M. (1995). Evaluation in tobacco of the organ specificity and strength of the rolD promoter, domain A of the 35S promoter and the 35S2 promoter. *Transgenic Res., 4*, 388–396.

Emiliani, G., Fondi, M., Fani, R., & Gribaldo, S. (2009). A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land. *Biology Direct, 4* (7).

Falkowski, P., Scholes, R., Boyle, E., Canadell, J., Canfield, D., Elser, J., et al. (2000). The global carbon cycle: a test of our knowledge of earth as a system. *Science, 290* (5490), 291-296.

Ferrer, J.-L., Austin, M., Stewart Jr., C., & Noel, J. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry, 46* (3), 356-370.

Fukaki, H., & Tasaka, M. (1999). Gravity perception and gravitropic response of inflorescence stems in Arabidopsis thaliana. *Advances in Space Research, 24* (6), 763-770.

Genevestigator (2009). Retrieved January 2007 & September 2009, from Genevestigator: https://www.genevestigator.com

Goujon, T., Sibout, R., Eudes, A., MacKay, J., & Jouanin, L. (2003). Genes involved in the biosynthesis of lignin precursors in Arabidopsis thaliana. *Plant Physiology and Biochemistry, 41*, 677-687.

Grima-Pettenati, J., & Goffner, D. (1999). Lignin genetic engineering revisited . *Plant Science, 145* (2), 51-65 .

Hajdukiewicz, P., Svab, Z., & Maliga, P. (1994). The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. *Plant Molecular Biology, 25* (6), 989-994.

Hatfield, R., & Fukushima, R. S. (2005). Can Lignin Be Accurately Measured? *Crop Science, 45*, 832-839.

Hati, K. M., Swarup, A., Dwivedi, A. K., Misra, A. K., & Bandyopadhyay, K. K. (2007). Changes in soil physical properties and organic carbon status at the topsoil horizon of a vertisol of central India after 28 years of continuous cropping, fertilization and manuring. *Agriculture, Ecosystems & Environment, 119* (1-2), 127-134.

Helariutta, Y., Fukaki, H., Wysocka-Diller, J., Nakajima, K., Jung, J., Sena, G., et al. (2000). The SHORT-ROOT Gene Controls Radial Patterning of the Arabidopsis Root through Radial Signaling. *Cell, 101* (5), 555-567.

Higo, K., Ugawa, Y., Iwamoto, M., & Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999 . *Nucleic Acids Research , 27* (1), 297-300.

Humphreys, J., & Chapple, C. (2002). Rewriting the lignin roadmap . *Current Opinion in Plant Biology, 5* (3), 224-229.

Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W and P Zimmermann (2008) **Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes.** *Advances in Bioinformatics 2008,* 420747.

ImageJ: Image Processing and Analysis in Java. (n.d.). Retrieved August 26, 2009, from ImageJ: Image Processing and Analysis in Java: http://rsb.info.nih.gov/ij/download.html.

Inukai, Y., Sakamoto, T., Ueguchi-Tanaka, M., Shibata, Y., Gomi, K., Umemura, I., et al. (2005). Crown rootless1, which is essential for crown root formation in rice, is a target of an AUXIN RESPONSE FACTOR in auxin signaling. *The Plant Cell, 17*, 1387-1396.

Ishiguro, S., & Nakamura, K. (1992). The nuclear factor SP8BF binds to the 5'-upstream regions of three different genes coding for major proteins of sweet potato tuberous roots . *Plant Mol. Biol., 18*, 97-108.

Jiao, Y., Ma, L., Strickland, E., & Deng, X. W. (2005). Conservation and Divergence of Light-Regulated Genome Expression Patterns during Seedling Development in Rice and Arabidopsis. *Plant Cell, 17*, 3239-3256.

Jobbágy, E. G. & Jackson, R. B. (2001). The distribution of soil nutrients with depth: Global patterns and the imprint of plants. *Biogeochemistry*, *53*, 51-77.

Johnson, J., Franzluebbers, A., Weyers, S., & Reicosky, D. (2007). Agricultural opportunities to mitigate greenhouse gas emissions. *Environmental Pollution, 150*, 107-124.

Jones, M. O., Manning, K., Andrews, J., Wright, C., Taylor, I. B., & Thompson, A. J. (2008). The promoter from SlREO, a highly-expressed, root-specific Solanum lycopersicum gene, directs expression to cortex of mature roots. *Functional Plant Biology, 35* (12), 1224–1233 .

Kagaya, Y., Ohmiya, K., & Hattori, T. (1999). RAV1, a novel DNAbinding protein, binds to bipartite recognition through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res., 27*, 470-478.

Kamiya, N., Nagasaki, H., Morikami, A., Sato, Y., & Matsuoka, M. (2003). Isolation and characterization of a rice WUSCHEL-type homeo box gene that is specifically expressed in the central cells of a quiescent center in the root apical meristem. *The Plant Journal, 35*, 429-441.

Karl, T., & Trenberth, K. (2003). Modern Global Climate Change. *Science, 302* (5651), 1719-1723.

Kim, T., Balish, R. S., Heaton, A. C., McKinney, E. C., Dhankher, O. P., & Meagher, R. B. (2005). Engineering a root-specific, repressor-operator gene complex. *Plant Biotechnology Journal, 3* (6), 571-582.

Klinedinst, S., Pascuzzi, P., Redman, J., Desai, M., & Arias, J. (2000). A xenobiotic-stress-activated transcription factor and its cognate target genes are preferentially expressed in root tip meristems. *Plant Mol. Biol. , 42*, 679-688.

Knox, J. (2008). Revealing the structural and functional diversity of plant cell walls. *Current Opinion in Plant Biology, 11* (3), 308-313.

Ko, J.-H., Kim, W.-C., & Han, K.-H. (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. *The Plant Journal*, published online August 6.

Ko, J.-H., Yang, S. H., Park, A. H., Lerouxel, O., & Han, K.-H. (2007). ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in Arabidopsis thaliana. *The Plant Journal, 50*, 1035-1048.

Kobayashi, T., Nakayama, Y., Itai, R. N., Nakanishi, H., Yoshihara, T., Mori, S., *et al.* (2003). Identification of novel cis-acting elements, IDE1 and IDE2, of the barley IDS2 gene promoter conferring iron-deficiency-inducible, root-specific expression in heterogeneous tobacco plants. *Plant Journal , 36* (6), 780-793.

Koyama, T., Ono, T., Shimizu, M., Jinbo, T., Mizuno, R., Tomita, K., *et al.* (2005). Promoter of Arabidopsis thaliana phosphate transporter gene drives root-specific expression of transgene in rice. *Journal of Bioscience and Bioengineering, 99* (1), 38-42.

Kumar, R., Pandey, S., & Pandey, A. (2006). Plant roots and carbon sequestration. *Current Science, 91* (7), 885-890.

Lal, R. (2008). Carbon sequestration. *Philos.Trans.R.Soc.Lond.B.Biol.Sci., 363* (1492), 815-830.

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science , 304* (5677), 1623-1627.

Lee, H. W., Kim, N. Y., Lee, D., & Kim, J. (2009, August 28). LBD18/ASL20 Regulates Lateral Root Formation in Combination with LBD16/ASL18 Downstream of ARF7 and ARF19 in Arabidopsis. *Plant Physiology*, 109.

Li, K.-F., Pahlevan, K., Kirschvink, J. L., & Yung, Y. L. (2009). Atmospheric pressure as a natural climate regulator for a terrestrial planet with a biosphere. *PNAS, 106* (24), 9576-9579.

Ma, L., Sun, N., Liu, X., Jiao, Y., Zhao, H., & Deng, X. W. (2005). Organ-Specific Expression of Arabidopsis Genome during Development. *Plant Physiology, 138*, 80-91.

Ma, S., & Bohnert, H. J. (2007). Integration of Arabidopsis thaliana stress-related transcript profiles, promoter structures, and cell-specific expression. *Genome Biology, 8* (4), R49.

Maizel, A., & Weigel, D. (2004). Temporally and spatially controlled induction of gene expression in Arabidopsis thaliana. *The Plant Journal, 38* (1), 164-171.

Malhi, Y., Meir, P., & Brown, S. (2002). Forests, carbon and global climate. *Philos Transact A Math Phys Eng Sci., 360* (1797), 1567-1591.

Maruyama-Nakashita, A., Nakamura, Y., Watanabe-Takahashi, A., Inoue, E., Yamaya, T., & Takahashi, H. (2005). Identification of a novel cis-acting element conferring sulfur defi ciency response in Arabidopsis roots. *The Plant Journal, 42*, 305-314.

Mccutchen, B. F., Castle, L. A., Chicoine, T. K., Cho, H.-j., Claus, J. S., Green, J. M., Guida, A. D., Hazel, C. B., Heckert, M. J., Hegstad, J. M., Hutchison, J. M., Liu, D., Lu, A. L., Mehre, W. J., Moy, Y., Olson, P. D., Peeples, K. A., Saunders, D. W., Vogt, M. D., Wilkinson, J. Q. & Wong, J. F. H. (2008) Compositions providing tolerance to multiple herbicides and methods of use thereof. US patent 2008023413.

Millard, P., Sommerkorn, M., & Grelet, G. A. (2007). Environmental change and carbon limitation in trees: a biochemical, ecophysiological and ecosystem appraisal. *New Phytologist , 175* (1), 11-28.

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (2007). NAC Transcription Factors, NST1 and NST3, Are Key Regulators of the Formation of Secondary Walls in Woody Tissues of Arabidopsis. *The Plant Cell, 19* (1), 270-280.

Mitsuda, N., Seki, M., Shinozaki, K., & Ohme-Takagi, M. (2005). The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *The Plant Cell , 17* (11), 2993-3006.

Mizusaki, S., Tanabe, Y., Noguchi, M., & Tamaki, E. (1971). Phytochemical studies on tobacco alkaloids XIV. The occurrence and properties of putrescine N-methyltransferase in tobacco roots . *Plant and Cell Physiology, 12* (4), 633-640.

Mondini, C., & Sequi, P. (2008). Implication of soil C sequestration on sustainable agriculture and environment. *Waste Management, 28* (4), 678-684.

Monolignol. (2008, January 4). Retrieved September 8, 2009, from Wikipedia: http://en.wikipedia.org/wiki/Monolignol

Moore, I., Samalova, M., & Kurup, S. (2006). Transactivated and chemically inducible gene expression in plants. *The Plant Journal, 45* (4), 651-683.

Nessler, C. L. (1994). Metabolic engineering of plant secondary products. *Transgenic Research, 3*, 109-115.

Ni, M., Cui, D., & Gelvin, S. B. (1996). Sequence-specific interactions of wound-inducible nuclear factors with mannopine synthase 2' promoter wound-responsive elements. *Plant Molecular Biology, 30* (1), 77-96.

Nitz, I., Berkefeld, H., Puzio, P. S., & Grundler, F. M. (2001). Pyk10, a seedling and root specific gene and promoter from Arabidopsis thaliana. *Plant Science, 161* (2), 337-346.

Oelkers, E. H., & Cole, D. R. (2008). Carbon Dioxide Sequestration A Solution to a Global Problem. *Elements, 4* (5), 305-310.

Parizot, B., Laplaze, L., Ricaud, L., Boucheron-Dubuisson, E., Bayle, V., Bonke, M., et al. (2008). Diarch Symmetry of the Vascular Bundle in Arabidopsis Root Encompasses the Pericycle and Is Reflected in Distich Lateral Root Initiation. *Plant Physiology, 146*, 140-148.

Petersen, M. (2007). Current status of metabolic phytochemistry. *Phytochemistry , 68* (22-24), 2847-2860.

Pomar, F., Merino, F., & Barceló, A. R. (2002). O-4-Linked coniferyl and sinapyl aldehydes in lignifying cell walls are the main targets of the Wiesner (phloroglucinol-HCl) reaction. *Protoplasma*, *220*, 17-28.

Prestridge, D. (1991). SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements. *CABIOS, 7*, 203-206.

Rasse, D. P., Rumpel, C., & Dignac, M.-F. (2005). Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant and Soil, 269*, 341-356.

Raven, J. A., & Karley, A. J. (2006). Carbon sequestration: photosynthesis and subsequent processes. *Current Biology, 16* (5), R165-7.

Robinson, D. (2007). Implications of a large global root biomass for carbon sink estimates and for soil carbon dynamics. *Proc.Biol.Sci., 274* (1626), 2753-2759.

Rogers, L., Dubos, C., Surman, C., Willment, J., Cullis, I., Mansfield, S., et al. (2005). Comparison of lignin deposition in three ectopic lignification mutants. *New Phytologist, 168* (1), 123-140.

Saballos, A., Ejeta, G., Sanchez, E., Kang, C., & Vermerris, W. (2009). A Genomewide Analysis of the Cinnamyl Alcohol Dehydrogenase Family in Sorghum [Sorghum bicolor (L.) Moench] Identifies SbCAD2 as the Brown midrib6 Gene. *Genetics, 181*, 783-795.

Salinas, J., Oeda, K., & Chua, N. H. (1992). Two G-Box-Related Sequences Confer Different Expression Patterns in Transgenic Tobacco. *The Plant Cell, 4* (12), 1485-1493.

Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., & Lepiniec, L. (2008). Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis . *The Plant Journal, 54* (4), 608-620.

Schimel, D. (1995). Terrestrial ecosystems and the carbon cycle. *Global Change Biology, 1*, 77-91.

Smith, P., & Falloon, P. (2005). Carbon sequestration in European croplands. *SEB Exp Biol Ser*, 47-55.

Smith, P., Martino, D., Cai, Z., Gwary, D., Janzen, H., Kumar, P., et al. (2008). Greenhouse gas mitigation in agriculture. *Philos.Trans.R.Soc.Lond.B.Biol.Sci., 363* (1492), 789-813.

Soltani, B., Ehlting, J., & Douglas, C. J. (2006). Genetic analysis and epigenetic silencing of At4CL1 and At4CL2 expression in transgenic Arabidopsis. *Biotechnology Journal, 1* (10), 1124-1136.

Subedi, K., Ma, B., & Liang, B. (2006). New method to estimate root biomass in soil through root-derived carbon. *Soil Biology and Biochemistry, 38* (8), 2212-2218.

Tang, W., Luo, X., & Samuels, V. (2004). Regulated gene expression with promoters responding to inducers. *Plant Science, 166*, 827-834.

Tao, S., Khanizadeh, S., Zhang, H., & Zhang, S. (2009). Anatomy, ultrastructure and lignin distribution of stone cells in two Pyrus species . *Plant Science, 176* (3), 413-419.

Tremousaygue, D., Manevski, A., Bardet, C., Lescure, N., & Lescure, B. (1999). Plant interstitial telomere motifs participate in the control of gene expression in root meristems. *The Plant Journal, 20* (5), 553-561.

Tsugeki, R., & Federoff, N. V. (1999). Genetic ablation of root cap cells in Arabidopsis. *PNAS, 96* (22), 12941-12946.

Tyo, K. E., Alper, H. S., & Stephanopoulos, G. N. (2007). Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends in Biotechnology , 25* (3), 132-137.

Vanholme, R., Morreel, K., Ralph, J., & Boerjan, W. (2008). Lignin engineering. *Current Opinion in Plant Biology, 11* (3), 278-285 .

Vaucheret, H., Béclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J.-B., et al. (2001). Transgene-induced gene silencing in plants. *The Plant Journal, 16* (6), 651-659.

Vieweg, M. F., Fruhling, M., Quandt, H. J., Heim, U., Baumlein, H., Puhler, A., et al. (2004). The promoter of the Vicia faba L. leghemoglobin gene VfLb29 is specifically activated in the infected cells of root nodules and in the arbuscule-containing cells of mycorrhizal roots from different legume and nonlegume plants. *Mol. Plant–Microbe Interact., 17*, 62-69.

Vijaybhaskar, V., Subbiah, V., Kaur, J., Vijayakumari, P., & Siddiqi, I. (2008). Identifi cation of a root-specifi c glycosyltransferase from Arabidopsis and characterization of its promoter. *Journal of Biosciences, 33* (2), 185-193.

Weber, W., Marty, R. R., Ehrbar, M., Keller, B., Weber, C. C., Zisch, A. H., et al. (2003). Conditional human VEGF-mediated vascularization in chicken embryos using a novel temperature-inducible gene regulation (TIGR) system. *Nucleic Acids Research, 31* (12), e69.

Weng, J.-K., Li, X., Bonawitz, N., & Chapple, C. (2008). Emerging strategies of lignin engineering and degradation for cellulosic biofuel production. *Current Opinion in Biotechnology, 19* (2), 166-172.

West, T., & Marland, G. (2002). A synthesis of carbon sequestration, carbon emissions, and net carbon flux in agriculture: comparing tillage practices in the United States. *Agriculture, Ecosystems & Environment , 91* (1-3), 217-232.

What Is Wood? (2009, April 14). Retrieved September 8, 2009, from Department of Chemistry; Index of CHY431: http://chemistry.umeche.maine.edu/CHY431/Wood/Gosta-Lignin.gif

Xie, Q., Frugis, G., Colgan, D., & Chua, N.-H. (2000). Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes & Development, 14*, 3024-3036 .

Yamaguchi, M., Kubo, M., Fukuda, H., & Demura, T. (2008). VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and shoots. *The Plant Journal, 55* (4), 652-664.

Yamamoto, Y. T., Taylor, C. G., Acedo, G. N., Cheng, C.-L., & Conkling, M. A. (1991). Characterization of cis-acting sequences regulating root-specific gene expression in tobacco. *The Plant Cell, 3* (4), 371-382.

Zhong, R., & Ye, Z.-H. (2007). Regulation of cell wall biosynthesis . *Current Opinion in Plant Biology, 10* (6), 564-572.

Zhong, R., Demura, T., & Ye, Z.-H. (2006). SND1, a NAC Domain Transcription Factor, Is a Key Regulator of Secondary Wall Synthesis in Fibers of Arabidopsis. *The Plant Cell, 18*, 3158-3170.

Zhong, R., Lee, C., Zhou, J., McCarthy, R. L., & Ye, Z.-H. (2008). A Battery of Transcription Factors Involved in the Regulation of Secondary Cell Wall Biosynthesis in Arabidopsis. *The Plant Cell, 20*, 2763-2782.

Zhong, R., Richardson, E. A., & Ye, Z. H. (2007a). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. *Planta, 225* (6), 1603-11.

Zhong, R., Richardson, E. A., & Ye, Z.-H. (2007b). The MYB46 Transcription Factor Is a Direct Target of SND1 and Regulates Secondary Wall Biosynthesis in Arabidopsis. *The Plant Cell, 19*, 2776-2792.

Zhou, J., Lee, C., Zhong, R., & Ye, Z.-H. (2009). MYB58 and MYB63 Are Transcriptional Activators of the Lignin Biosynthetic Pathway during Secondary Cell Wall Formation in Arabidopsis. *The Plant Cell, 21*, 248-266.

Zibilske, L. M., & Bradford, J. M. (2007). Oxygen Effects on Carbon, Polyphenols, and Nitrogen Mineralization Potential in Soil. *Soil Biology & Biochemistry, 71*, 133-139.

Zuo, J., Niu, Q.-W., & Chua, N.-H. (2000). An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *The Plant Journal, 24* (2), 265-273.

Appendices

Appendix A. Primary sequences of gene expression constructs

4CL1pro-SND1 (2381 bp)

```
                                                    EcoR IForward primer
-1280                                          5′-GAATTCTTTTCGGTCTCTAA
-1260 TACCTCCGGTTTTAAAAAAAAACATATCAGTTGAAGGATGAGTTTGGTGAAGGCTATATTGTC
-1197 CATTGATTTTGGAGATATATGTATTATGGTCATGATTATTACGATTTTTATATAAAAGAATAT
-1134 TAAAAATGGTGGGGTTGGTGAAGAAATGAAGATTTATCGTCAAATATTTCAATTTTTACTTGG
-1071 ACTATTGCTTCGGTTATATCGTCAACATGGGCCCACTCTTCCACCAAAGCCCAATCAATATAT
-1008 CTCTCGCTATCTTCACCAACCCACTCTTCTTCTCTTACCAAACCCATTTCCTTTATTTCCAAC
-945  CCTACCCCTTTATTTCTCAAGCTTTACACTTTTAGCCCATAACTTTCTTTTTATCCAAATGGA
-882  TTTGACTGGTCTCCAAAGTTGAATTAAATGGTTGTAGAAATAAAATAAAATTATACGGGTTCA
-819  ATTGTTCAATTGTTCATATACCGTTGACGTTCAATTGTTCATATACGGGTTCCGTGGTCGTTG
-756  GTAATATATATGTCTTTTATGGAACCAAAATAGACCAAATCAACAACAAATGAAGAAATTGTT
-693  AGAGTATGATACACTCATATATACCCAAATATAGCATATATTTATAATATAACTTTTGGCTAT
-630  GTCATTTTACATGATTTTTTTGGCTTATCTATTAAAAGTATCATACAAACTGTTTTTACTTCT
-567  TTTTTTTCTTAGAATATATATGCCCAAAATGGAAAAGAACATATGCCAAGGTTGATTTTATCG
-504  CTTATATGGTAAAAATTGGAAAAACATACAAATCATTACTTTATTTAATTAAATCATGTGAAG
-441  AAACATATTCAATTACGGTAATACGTTATCAAAACATTTTTTTTTACATTAATTGTTACATTT
-378  TTTTTTTTTGCAAATATTCTTAAATAACCATTCTTTTTTTATTTACTATAATTAACATAAAAA
-315  TAAATAAAATATAACATTTCAACAAAGAAATTTGCTTATGAAAAATACAAAATCCAGTTAATT
-252  TTTCAGAAAAATACAAATTTGCTTATAAATATATTACCACTAGTTTATGTGATTTTAAAAGAA
-189  AGAAATGCAGCTTACCAAACGCAACGTGAAAATTTGAGAAACCCATACTCAAAAAAGATTAAA
-126  TGACAAAATCACCCTCAGCAAAATCATGAAACAACAACACTAACATTTTCACCAACCCCACCG
-63   TCTACTCCGGTGAATTGTCTATATGAACTCCTCCGATACAACTCCTGTTTCCTTCAGCCGCGG
                                              Reverse primer     Sac II
   +1
   MetAlaAspAsnLysValAsnLeuSerIleAsnGlyGlnSerLysValProProGlyPheArg
   ATGGCTGATAATAAGGTCAATCTTTCGATTAATGGACAATCAAAAGTGCCTCCAGGTTTCAGA 63
          Forward Primer
   PheHisProThrGluGluGluLeuLeuHisTyrTyrLeuArgLysLysValAsnSerGlnLys
   TTCCATCCCACCGAAGAAGAACTTCTCCATTACTATCTCCGTAAGAAAGTTAACTCTCAAAAG 126

   IleAspLeuAspValIleArgGluValAspLeuAsnLysLeuGluProTrpAspIleGlnGlu
   ATCGATCTTGATGTCATTCGTGAAGTTGATCTAAACAAGCTTGAGCCTTGGGATATTCAAGAG 189

   GluCysArgIleGlySerThrProGlnAsnAspTrpTyrPhePheSerHisLysAspLysLys
   GAATGTAGAATCGGTTCAACGCCACAAAACGACTGGTACTTCTTCAGCCACAAGGACAAGAAG 252

   TyrProThrGlyThrArgThrAsnArgAlaThrValAlaGlyPheTrpLysAlaThrGlyArg
   TATCCAACCGGGACCAGGACGAACCGGGCAACAGTCGCTGGATTCTGGAAAGCTACCGGACGT 315

   AspLysIleIleCysSerCysValArgArgIleGlyLeuArgLysThrLeuValPheTyrLys
   GACAAAATCATCTGCAGTTGTGTCCGGAGAATTGGACTGAGGAAGACACTCGTGTTCTACAAA 378

   GlyArgAlaProHisGlyGlnLysSerAspTrpIleMetHisGluTyrArgLeuAspAspThr
   GGAAGAGCTCCTCACGGTCAGAAATCCGACTGGATCATGCATGAGTATCGCCTCGACGATACT 441

   ProMetSerAsnGlyTyrAlaAspValValThrGluAspProMetSerTyrAsnGluGluGly
   CCAATGTCTAATGGCTATGCTGATGTTGTTACAGAAGATCCAATGAGCTATAACGAAGAAGGT 504
```

```
TrpValValCysArgValPheArgLysLysAsnTyrGlnLysIleAspAspCysProLysIle
TGGGTGGTATGTCGAGTGTTCAGGAAGAAGAACTATCAAAAGATTGACGATTGTCCTAAAATC 567

ThrLeuSerSerLeuProAspAspThrGluGluGluLysGlyProThrPheHisAsnThrGln
ACTCTATCTTCTTTACCTGATGACACGGAGGAAGAGAAGGGGCCCACCTTTCACAACACTCAA 630

AsnValThrGlyLeuAspHisValLeuLeuTyrMetAspArgThrGlySerAsnIleCysMet
AACGTTACCGGTTTAGACCATGTTCTTCTCTACATGGACCGTACCGGTTCTAACATTTGCATG 693

ProGluSerGlnThrThrThrGlnHisGlnAspAspValLeuPheMetGlnLeuProSerLeu
CCCGAGAGCCAAACAACGACTCAACATCAAGATGATGTCTTATTCATGCAACTCCCAAGTCTT 756

GluThrProLysSerGluSerProValAspGlnSerPheLeuThrProSerLysLeuAspPhe
GAGACACCTAAATCCGAGAGCCCGGTCGACCAAAGTTTCCTGACTCCAAGCAAACTCGATTTC 819

SerProValGlnGluLysIleThrGluArgProValCysSerAsnTrpAlaSerLeuAspArg
TCTCCCGTTCAAGAGAAGATAACCGAAAGACCGGTTTGCAGCAACTGGGCTAGTCTTGACCGG 882

LeuValAlaTrpGlnLeuAsnAsnGlyHisHisAsnProCysHisArgLysSerPheAspGlu
CTCGTAGCTTGGCAATTGAACAATGGTCATCATAATCCGTGTCATCGTAAGAGTTTTGATGAA 945

GluGluGluAsnGlyAspThrMetMetGlnArgTrpAspLeuHisTrpAsnAsnAspAspAsn
GAAGAAGAAAATGGTGATACTATGATGCAGCGATGGGATCTTCATTGGAATAATGATGATAAT 1008

ValAspLeuTrpSerSerPheThrGluSerSerSerSerLeuAspProLeuLeuHisLeuSer
GTTGATCTTTGGAGTAGTTTCACTGAGTCTTCTTCGTCTTTAGACCCACTTCTTCATTTATCT 1071
                                        Reverse Primer

Val   HisHisHisHisHisHis
GTATGACATCATCATCATCATCATGGATCC-3'                                  1101
        6xHis tag    BamH I
```

GSTU19pro-SND1 (2558 bp)

EcoR I

-1457 **Forward primer** 5′**GAATTC**GC

-1449 TACGTGTCGTGAGATATCGAACCCAACGCAGATATGAGTATGTTGAGCTAGTTTCTTCTTATG

-1386 AAACAATCATATATGTCTATAATGAATAGATCACATTATCTGCCTGAAAAAAATCCCGTATAT

-1323 TACTCGACGAAATATAAATACCCAATGTAGCTGATTTTGCTTTCTCTGGTGACATATCCAATT

-1260 TGGCTAAATTTGTTAACTAGTCTATTATAGGTTTATAATAGATCTAGCTATGTTAAAGATACT

-1197 AAAGCATCAGTTACATAAATTTTTGGCGCGAGTTTATATCTTTTGGAATTAAAAATAAGAGAA

-1134 TTTAAAAATAAGAAGATCATTTTGTTTGGCCACAGGAGTTCTGAAAGGTCAGGTATGATTTTT

-1071 TTCTTGCTCGCTCTTATGATTTTGTTTTTATTAATGGGTTTTCAAATAAGAAAAACTGTTTTT

-1008 CGAAGCCCGGTTCAGATCCATTGTTTTTTGTAAAATATAGGCCCAATTCACCATAAGTCCATG

-945 ACCAAAACAAAAATAAGATAGAACCAATACTGAACCAGGATCTTCTCTCGCTTTCGTGATCAA

-882 TGTCGCCAAGCTTCTCGAGATCATGTGGTCACGTCAATTGTATAAATACAATTATTGACGTAA

-819 CACAATCTCTACAGTTCCATCGAAATATCTCGAAAATTTCCAGTTAATTCTGGTAACGTGAAC

-756 GTATCTTCCACCTCTTCAACCTACACAGCTTTCTAGAAATTTGGCTCGCTTTTCTAAGTCCTC

-693 TGTATTTTTTTGCACGTTTTTCAACTAAGTTTCAATATGAATCATTTCTTCTATAAATAAATG

-630 ATATTTTCATCAGGTAATGATACATTGTGCCGAAATAAAACGTCAATACTCATTAGTCAAATT

-567 AATTGTTCACATAATTTAAAACTGTGTTAATCCATCCAGTTATTTTCTTACAACAAAATAATC

-504 TTTTCCATCAACTTTTAAAATAATTAAACGCAGTGCTAAGAAATCTAAAATCTTGATTTAGAA

-441 ATCCATTATGGTTTCTGGTCAACTGAAATCCATAATTTCCTTTAACATCCAAAATCCAAATTT

-378 GCTACTATGATAATAGATTTCAGACGATTTTTTTTCTTTTTTCAATCATAGAGTCCACACGAA

-315 TATTTGCAAGTTACTATATAAAACACTATAATGGTCAACAGATAAAAAAAAGGCGAATGAAGA

-252 TATGTTACGTAAAAAGAAAATACTGTAATTATAAATTATTACTTTAAAAAGCTTTAAAATCTG

-189 GCCACATGTTTTTAAAGAGTGGTGTGACGTAACGACTAGAGTCAGCACAATCCATTATTGTAT

-126 CATAAATATTCTCATCTATAAATTACCTAAACCCTTACAGGTAGTGTCCCAACCAAACAAATC

-63 GAGAAAGACGAACACTTACAAAAAAAAATCTCTTTGTGAGCTTTAGCGATCGTAACA**CCGCGG**

SacII

Reverse primer

+1

MetAlaAspAsnLysValAsnLeuSerIleAsnGlyGlnSerLysValProProGlyPheArg

ATGGCTGATAATAAGGTCAATCTTTCGATTAATGGACAATCAAAAGTGCCTCCAGGTTTCAGA 63

Forward Primer

PheHisProThrGluGluGluLeuLeuHisTyrTyrLeuArgLysLysValAsnSerGlnLys

TTCCATCCCACCGAAGAAGAACTTCTCCATTACTATCTCCGTAAGAAAGTTAACTCTCAAAAG 126

IleAspLeuAspValIleArgGluValAspLeuAsnLysLeuGluProTrpAspIleGlnGlu

ATCGATCTTGATGTCATTCGTGAAGTTGATCTAAACAAGCTTGAGCCTTGGGATATTCAAGAG 189

GluCysArgIleGlySerThrProGlnAsnAspTrpTyrPhePheSerHisLysAspLysLys

GAATGTAGAATCGGTTCAACGCCACAAAACGACTGGTACTTCTTCAGCCACAAGGACAAGAAG 252

TyrProThrGlyThrArgThrAsnArgAlaThrValAlaGlyPheTrpLysAlaThrGlyArg

TATCCAACCGGGACCAGGACGAACCGGGCAACAGTCGCTGGATTCTGGAAAGCTACCGGACGT 315

AspLysIleIleCysSerCysValArgArgIleGlyLeuArgLysThrLeuValPheTyrLys

GACAAAATCATCTGCAGTTGTGTCCGGAGAATTGGACTGAGGAAGACACTCGTGTTCTACAAA 378

GlyArgAlaProHisGlyGlnLysSerAspTrpIleMetHisGluTyrArgLeuAspAspThr

GGAAGAGCTCCTCACGGTCAGAAATCCGACTGGATCATGCATGAGTATCGCCTCGACGATACT 441

ProMetSerAsnGlyTyrAlaAspValValThrGluAspProMetSerTyrAsnGluGluGly

CCAATGTCTAATGGCTATGCTGATGTTGTTACAGAAGATCCAATGAGCTATAACGAAGAAGGT 504

TrpValValCysArgValPheArgLysLysAsnTyrGlnLysIleAspAspCysProLysIle

TGGGTGGTATGTCGAGTGTTCAGGAAGAAGAACTATCAAAAGATTGACGATTGTCCTAAAATC 567

ThrLeuSerSerLeuProAspAspThrGluGluGluLysGlyProThrPheHisAsnThrGln
ACTCTATCTTCTTTACCTGATGACACGGAGGAAGAGAAGGGGCCCACCTTTCACAACACTCAA 630

AsnValThrGlyLeuAspHisValLeuLeuTyrMetAspArgThrGlySerAsnIleCysMet
AACGTTACCGGTTTAGACCATGTTCTTCTCTACATGGACCGTACCGGTTCTAACATTTGCATG 693

ProGluSerGlnThrThrThrGlnHisGlnAspAspValLeuPheMetGlnLeuProSerLeu
CCCGAGAGCCAAACAACGACTCAACATCAAGATGATGTCTTATTCATGCAACTCCCAAGTCTT 756

GluThrProLysSerGluSerProValAspGlnSerPheLeuThrProSerLysLeuAspPhe
GAGACACCTAAATCCGAGAGCCCGGTCGACCAAAGTTTCCTGACTCCAAGCAAACTCGATTTC 819

SerProValGlnGluLysIleThrGluArgProValCysSerAsnTrpAlaSerLeuAspArg
TCTCCCGTTCAAGAGAAGATAACCGAAAGACCGGTTTGCAGCAACTGGGCTAGTCTTGACCGG 882

LeuValAlaTrpGlnLeuAsnAsnGlyHisHisAsnProCysHisArgLysSerPheAspGlu
CTCGTAGCTTGGCAATTGAACAATGGTCATCATAATCCGTGTCATCGTAAGAGTTTTGATGAA 945

GluGluGluAsnGlyAspThrMetMetGlnArgTrpAspLeuHisTrpAsnAsnAspAspAsn
GAAGAAGAAAATGGTGATACTATGATGCAGCGATGGGATCTTCATTGGAATAATGATGATAAT 1008

ValAspLeuTrpSerSerPheThrGluSerSerSerSerLeuAspProLeuLeuHisLeuSer
GTTGATCTTTGGAGTAGTTTCACTGAGTCTTCTTCGTCTTTAGACCCACTTCTTCATTTATCT 1071
Reverse Primer

Val **HisHisHisHisHisHis**
GTATGA**CATCATCATCATCATCAT**<u>**GGATCC**</u>**-3′** 1101
6XHis tag ***BamH I***

Appendix B. *Cis*-acting DNA regulatory element analysis of *At4CL1* and *AtGSTU19* promoters

Table 6. *Cis*-acting DNA regulatory element analysis of *At4CL1*, 2000bp upstream of the transcription start site.

Putative root motifs (Vijaybhaskar *et al.* 2008)	**Location**	**Strand**	**Signal sequence**	**Description** (Higo *et al.* 1999; Prestridge 1991)
ARFAT	-424	(+)	TGTCTC	ARF binding site found in the promoters of primary/early auxin response genes of Arabidopsis thaliana
ASF1MOTIFCAMV	-341	(+)	TGACG	ASF-1 binding site involved in transcriptional activation of several genes by auxin and/or salicylic acid
ASF1MOTIFCAMV	-1214	(+)	TGACG	ASF-1 binding site
ASF1MOTIFCAMV	-912	(–)	TGACG	ASF-1 binding site
ASF1MOTIFCAMV	-957	(–)	TGACG	ASF-1 binding site
OSE1ROOTNODULE	-2873	(+)	AAAGAT	A consensus sequence motif of organ-specific elements characteristic of activated promoters found in the infected cells of root nodules
OSE1ROOTNODULE	-1873	(+)	AAAGAT	organ-specific elements
OSE1ROOTNODULE	-494	(–)	AAAGAT	organ-specific elements
OSE2ROOTNODULE	-427	(+)	CTCTT	organ-specific elements
OSE2ROOTNODULE	-973	(+)	CTCTT	organ-specific elements
OSE2ROOTNODULE	-1024	(+)	CTCTT	organ-specific elements
OSE2ROOTNODULE	-1032	(+)	CTCTT	organ-specific elements
RAV1AAT	-960	(+)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-1293	(+)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-1296	(+)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-1713	(+)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-1914	(+)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-1917	(+)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-198	(–)	CAACA	RAV1 transcription factor binding consensus sequence
ROOTMOTIFTAPOX1	-64	(+)	ATATT	Motif found in rolD promoters; organ specificity and strength
ROOTMOTIFTAPOX1	-92	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-307	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-337	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-366	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-687	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-804	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-871	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-918	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1353	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1572	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1644	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1787	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-91	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-115	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-159	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-169	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-469	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-686	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-870	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-917	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-994	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1255	(–)	ATATT	Motif found in rolD promoters

ROOTMOTIFTAPOX1	-1343	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1361	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1454	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1643	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1701	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1784	(–)	ATATT	Motif found in rolD promoters
SP8BFIBSP8BIB	-275	(–)	TACTATT	SPBF binding site
SP8BFIBSP8BIB	-510	(–)	TACTATT	SPBF binding site
SURECOREATSULTR11	-425	(–)	GAGAC	Core of SURE found in the promoter of SULTR1; sulfate uptake and transport
SURECOREATSULTR11	-741	(–)	GAGAC	Core of SURE found in the promoter of SULTR1; sulfate uptake and transport
SURECOREATSULTR11	-1135	(–)	GAGAC	Core of SURE found in the promoter of SULTR1; sulfate uptake and transport
WUSATAg	-416	(+)	TTAATGG	Target sequence of WUS in the intron of AGAMOUS gene in Arabidopsis

Table 7. *Cis*-acting DNA regulatory element analysis of *AtGSTU19*, 2000bp upstream of the transcription start site.

Putative root motifs (Vijaybhaskar *et al.* 2008)	**Location**	**Strand**	**Signal sequence**	**Description** (Higo *et al.* 1999; Prestridge 1991)
ARFAT	-359	(+)	TGTCTC	ARF binding site found in the promoters of primary/early auxin response genes of Arabidopsis thaliana
ASF1MOTIFCAMV	-371	(+)	TGACG	ASF-1 binding site involved in transcriptional activation of several genes by auxin and/or salicylic acid
ASF1MOTIFCAMV	-1267	(+)	TGACG	ASF-1 binding site
ASF1MOTIFCAMV	-1929	(+)	TGACG	ASF-1 binding site
ASF1MOTIFCAMV	-1243	(–)	TGACG	ASF-1 binding site
ASF1MOTIFCAMV	-1504	(–)	TGACG	ASF-1 binding site
OSE1ROOTNODULE	-888	(+)	AAAGAT	A consensus sequence motif of organ-specific elements characteristic of activated promoters found in the infected cells of root nodules
OSE1ROOTNODULE	-34	(–)	AAAGAT	organ-specific elements
OSE1ROOTNODULE	-127	(–)	AAAGAT	organ-specific elements
OSE1ROOTNODULE	-177	(–)	AAAGAT	organ-specific elements
OSE1ROOTNODULE	-934	(–)	AAAGAT	organ-specific elements
OSE1ROOTNODULE	-1587	(–)	AAAGAT	organ-specific elements
OSE2ROOTNODULE	-47	(+)	CTCTT	organ-specific elements
OSE2ROOTNODULE	-158	(+)	CTCTT	organ-specific elements
OSE2ROOTNODULE	-1033	(+)	CTCTT	organ-specific elements
OSE2ROOTNODULE	-1349	(+)	CTCTT	organ-specific elements
OSE2ROOTNODULE	-953	(–)	CTCTT	organ-specific elements
OSE2ROOTNODULE	-1919	(–)	CTCTT	organ-specific elements
RAV1AAT	-1577	(+)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-1814	(+)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-639	(–)	CAACA	RAV1 transcription factor binding consensus sequence
RAV1AAT	-685	(–)	CAACA	RAV1 transcription factor binding consensus sequence
ROOTMOTIFTAPOX1	-336	(+)	ATATT	Motif found in rolD promoters; organ specificity and strength
ROOTMOTIFTAPOX1	-605	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-767	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1464	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1778	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1973	(+)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-267	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-391	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-781	(–)	ATATT	Motif found in rolD promoters

ROOTMOTIFTAPOX1	-1119	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1298	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1434	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1777	(–)	ATATT	Motif found in rolD promoters
ROOTMOTIFTAPOX1	-1972	(–)	ATATT	Motif found in rolD promoters
SORLIP1AT	154	(+)	GCCAC	One of "Sequences Over-Represented in Light-Induced Promoters (SORLIPs) in Arabidopsis; Computationally identified phyA-induced motifs; SORLIP 1 is most over-represented, and most statistically significant
SORLIP1AT	435	(+)	GCCAC	One of "Sequences Over-Represented in Light-Induced Promoters (SORLIPs)
SORLIP1AT	988	(+)	GCCAC	One of "Sequences Over-Represented in Light-Induced Promoters (SORLIPs)
SORLIP1AT	1905	(+)	GCCAC	One of "Sequences Over-Represented in Light-Induced Promoters (SORLIPs)
SP8BFIBSP8BIB	-330	(–)	TACTATT	SPBF binding site
SURECOREATSULTR11	-496	(+)	GAGAC	Core of SURE found in the promoter of SULTR1; sulfate uptake and transport
SURECOREATSULTR11	-360	(–)	GAGAC	Core of SURE found in the promoter of SULTR1; sulfate uptake and transport
WUSATAg	-1053	(+)	TTAATGG	Target sequence of WUS in the intron of AGAMOUS gene in Arabidopsis

Appendix C. Primer sequences

Table 8. List of all primer sequences used for PCR, reverse transcription-PCR and sequencing.

No.	Name	Primer Sequence	Comments
1	*4CL1* Forward	5′-TCCAGAGGTGTAAAGTGACGGTGGC-3′	Native gene expression
2	*4CL1* Reverse	5′-CCGTCATTCCGTATCCCTGACCGAG-3′	Native gene expression
3	*GSTU19* Forward	5′-AGGTGTGGGCGACAAAGGGTG-3′	Native gene expression
4	*GSTU19* Reverse	5′-CCACGCTCTCCCTCTGCAAACAC-3′	Native gene expression
5	*4CL1pro* Forward	5'-GGGCAC**G˅AATTC**TTTTCGGTCTCTAATACCTCC-3′	EcoRI RE site
6	*4CL1pro* Reverse	5′CACGAG**G˅GATCCG˅GTNACCCCGC˅GG**CTGAAGGAAACAGGAGTTGTATC-3′	BamHI, BstEII and SacII RE sites
7	*GSTU19pro* Forward	5′-GGGTCT**G˅AATTC**GCTACGTGTCGTGAGATATCG-3′	EcoRI RE site
8	*GSTU19pro* Reverse	5′-CACGAG**G˅GATCCG˅GTNACCCCGC˅GG**TGTTACGATCGCTAAAGCTCAC-3′	BamHI, BstEII and SacII RE sites
9	*SND1* Forward	5′GAGCTC**CCGC˅GG**ATGGCTGATAATAAGGTCAATCTTTCG-3′	SacII RE site
10	*SND1* Reverse	5′GGGTGT**G˅GATCC****ATGATGATGATGATGATG**TCATACAGATAAATGAAGAAGTGGGTC-3′	BamHI RE site and HIS x6 tag
11	*4CL1pro-SND1* Rev (mid-insert)	5′-GTCACGTCCGGTAGCTTTCC-3′	For sequencing from the middle of the insert
12	*GSTU19pro-SND1* Rev (mid-insert)	5′-TCTCCGGACACAACTGCAGATG-3′	For sequencing from the middle of the insert
13	*MYB46* Forward	5′-CTGGTCGGACCGATAACGAG-3′	300bp fragment
14	*MYB46* Reverse	5′-GGTGGCTGATCATGTTTCCC-3′	300bp fragment
15	*SND3* Forward	5′-ACGCTTGAAGGAGAGAATGG-3′	300bp fragment
16	*SND3* Reverse	5′-CTGATGCATCACCCAATTCG-3′	300bp fragment
17	*MYB103* Forward	5′-AGGTGGGCTCATATAGCTAG-3′	400bp fragment
18	*MYB103* Reverse	5′-CTCTTCCTCCTCTTTGCGTG-3′	400bp fragment
19	*KNAT7* Forward	5′-CAGCACGTGAGGGTTCATGC-3′	300bp fragment
20	*KNAT7* Reverse	5′-CCCAGCCCTTCTCTTCCTCA-3′	300bp fragment
21	*SND1* Forward	5′-GATCATGCATGAGTATCGCC-3′	200bp fragment
22	*SND1* Reverse	5′-CGGGCTCTCGGATTTAGGTG-3′	200bp fragment

23	*4CL1* L1	5′-TCAACCCGGTGAGATTTGTA-3′	From Apurva Bhargava (Ellis Lab)
24	*4CL1* R1	5′-TCGTCATCGATCAATCCAAT-3′	From Apurva Bhargava (Ellis Lab)
25	*CCR1* L1	5′-GTGCAAAGCAGATCTTCAGG-3′	From Apurva Bhargava (Ellis Lab)
26	*CCR1* R1	5′-GCCGCAGCATTAATTACAAA-3′	From Apurva Bhargava (Ellis Lab)
27	*COMT1* L1	5′-GTGCAAAGCAGATCTTCAGG-3′	From Apurva Bhargava (Ellis Lab)
28	*COMT1* R1	5′-CATGGTGATTGTGGAATGGT-3′	From Apurva Bhargava (Ellis Lab)
29	*ACT8F*(QRT)	5′-TCTAAGGAGGAGCAGGTTTGA-3′	From Apurva Bhargava (Ellis Lab)
30	*ACT8R*(QRT)	5′-TTATCCGAGTTTGAAGAGGCTAC-3′	From Apurva Bhargava (Ellis Lab)

Appendix D. Media, Buffers and Reagent Stocks

LB broth (1L)

- Tryptone 10 g
- Yeast Extract 5 g
- NaCl 10 g

*For plates add 15g agar

½ MS media (1L)

- MS salt plus vitamin 2.2 g
- MES hydrate 0.5 g
- Sucrose 10 g (phenotyping)
  20 g (growth)

*For plates add 7g agar
*Adjust pH to 5.7 using 1M KOH

1/10 Johnson solution (20L)

- 20mM Fe-EDTA (use 3mL/20L solution)
- 10mM $CaSO_4$ (use 800mL/20L solution)
- Macro stock (mix: 20mL/L of 1M $MgSO_4$, 40mL/L of 1M KH_2PO_4, 80mL/L of 0.5M K_2SO_4) (use 100mL/20L solution)
- Micro stock (mix: 25mM H_3BO_3, 2mM $MnSO_4$ x H_20, 2mM $ZnSO_4$ x H_20, 0.5mM $CuSO_4$ x $5H_2O$, 0.5mM $NaMoO_4$) (use 3mL/20mL solution)
- 2 spoons of $CaCO_3$ powder
- Add NH_2NO_3 directly to a final concentration of 1mM

Benoxacor 100mM Stock Solution (1000x)

- Benoxacor 125 mg
- Acetone 4.81 mL

Fenclorim 100mM Stock Solution (1000x)

- Fenclorim 125 mg
- Acetone 5.55 mL

Klason lignin procedure solutions

72% H_2SO_4

665 mL conc. H_2SO_4
300 mL DI H_2O
cool, bring to 1L

4% H_2SO_4

37 mL conc. H_2SO_4
950 mL DI H_2O
cool, bring to 1L

Sugar Control (in 50 mL DI H_2O)

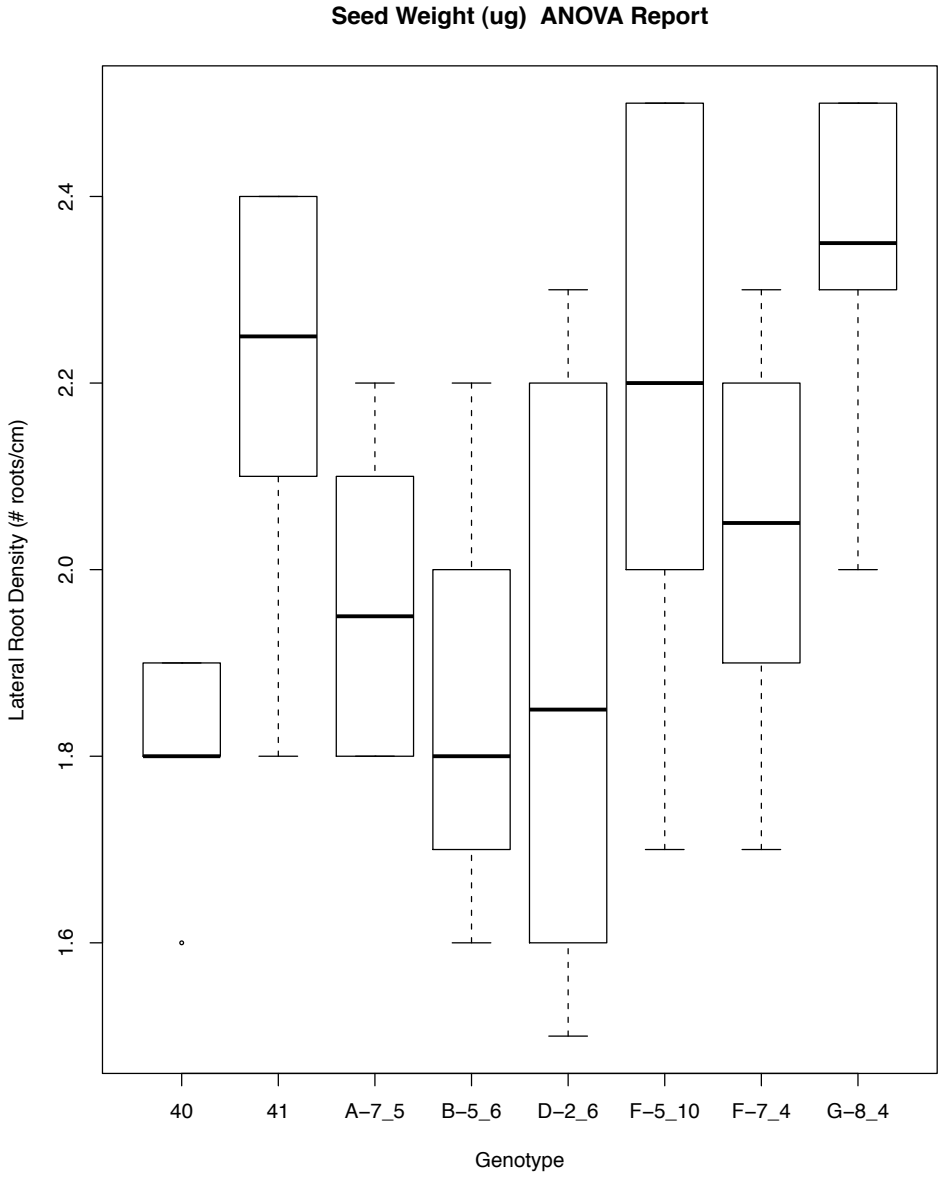
arabinose	10 mg
galactose	10 mg
glucose	200 mg
xylose	60 mg
mannose	60 mg
rhamnose	50 mg

High standard:	sugar stock	30 mL
	DI H_2O	82 mL
	72% H_2SO_4	3 mL
Medium Standard:	sugar stock	10 mL
	DI H_2O	102 mL
	72% H_2SO_4	3 mL
Low Standard:	sugar stock	5 mL
	DI H_2O	107 mL
	72% H_2SO_4	3 mL

Internal Standard

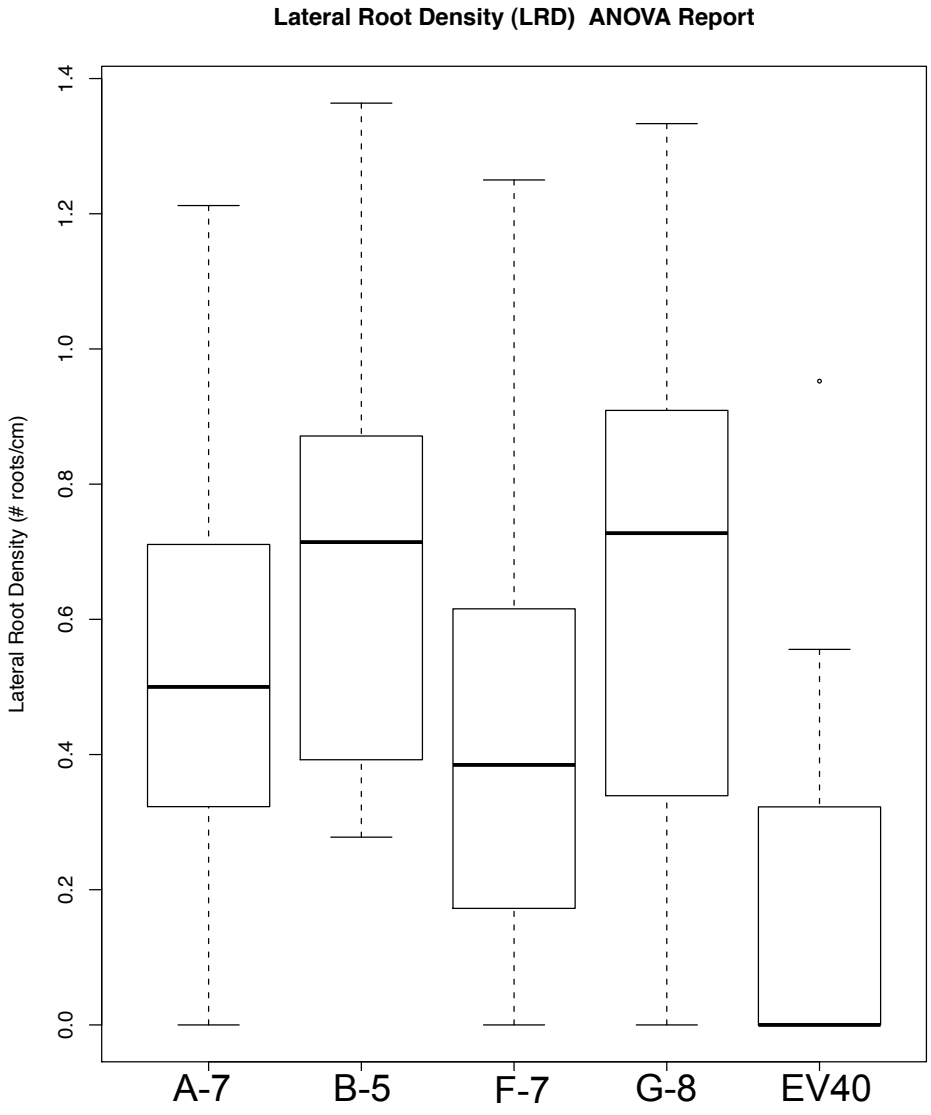
fucose	10 mg/mL

Appendix E. One-way analysis of variance (ANOVA) for average seed weight and lateral root density



	Df	Sum Sq	Mean Sq	F value	Pr(>F) value
Seed weight	7	1.53146	0.21878	4.1053	0.001756
Residuals	40	2.13167	0.05329		

Figure 16. One-way ANOVA statistical analysis to determine differences in average seed weight between genotypes



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
LRD	4	3.7035	0.9259	8.546	6.194e-06
Residuals	95	10.2922	0.1083		

Figure 17. One-way ANOVA statistical analysis to determine differences in average number of lateral roots between genotypes