Investigation into the Cell Wall and Cellulose Biosynthesis in Model Species and in the C4 Model Plant Setaria viridis

Mizuki Tateno

University of Kentucky, mxtateno@gmail.com

Digital Object Identifier: https://doi.org/10.13023/ETD.2016.482
STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student’s advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student’s thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Mizuki Tateno, Student
Dr. Seth DeBolt, Major Professor
Dr. Arthur Hunt, Director of Graduate Studies
INVESTIGATION INTO THE CELL WALL AND CELLULOSE BIOSYNTHESIS IN MODEL SPECIES AND IN THE C4 MODEL PLANT \textit{SETARIA VIRIDIS}

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Agriculture, Food and Environment at the University of Kentucky

By
Mizuki Tateno
Lexington, Kentucky

Director: Dr. Seth DeBolt, Professor of Horticulture
Lexington, Kentucky
2016

Copyright © Mizuki Tateno 2016
ABSTRACT OF DISSERTATION

INVESTIGATION INTO THE CELL WALL AND CELLULOSE BIOSYNTHESIS IN MODEL SPECIES AND IN THE C4 MODEL PLANT SETARIA VIRIDIS

A uniform feature of all plant cells is the presence of a cell wall. The cell wall functions in facilitating directional expansion and is therefore important for cell shape and morphogenesis. All plant cell walls contain cellulose microfibrils embedded in a network of polysaccharides, lignin and protein. Cellulose is evolutionarily conserved and is made by all plants as well as other members of various taxonomic kingdoms. From a human perspective, the field of renewable energy has had an ever increasing interest in using the cell wall for production of renewable platform chemicals and fuels. However, the biosynthesis of these components is complex and the intricacies are still being solved. Herein, a molecular genetics approach was taken to advance our understanding of cell wall biosynthesis. Initial work proposes a C4 model grass Setaria viridis as well as for the crop plant Sorghum bicolor for genetic analysis. The machinery that is responsible for the biosynthesis the cell wall such as the cellulose synthase (CESA) enzymes were identified and the compositional analysis of the cell wall was conducted. In addition, the role of a cellulose synthase like D1 (CSLD1) gene was studied using reverse genetics. Finally, regulatory features called microRNAs that may intersect with lignin biosynthesis were explored through a series of molecular genetic analyses using Arabidopsis.

KEYWORDS: cellulose, cell wall, Arabidopsis, Setaria viridis, molecular genetics, biofuel
INVESTIGATION INTO THE CELL WALL AND CELLULOSE BIOSYNTHESIS IN MODEL SPECIES AND IN THE C4 MODEL PLANT SETARIA VIRIDIS

By
Mizuki Tateno

Seth DeBolt
Director of Dissertation

Arthur G. Hunt
Director of Graduate Studies

11-28-2016
ACKNOWLEDGEMENTS

The work that is presented within these pages would not have been possible without the opportunity given to me first and foremost by my advisor Professor Seth DeBolt. His willingness and allowance for lab members work freely to pursue and find our own niche is something that I will always be grateful of. In addition, this dissertation would also not have been possible if it were not for all of the support and encouragement that I have received from the members of the Debolt lab past and present. I would also like to thank the members of my committee: Dr. Bob Houtz, Dr. Sue Nokes, Dr. Jan Smalle and my outside examiner Dr. Czarena Crofcheck for taking the time and for all of the valuable advice, support and suggestions along the way. Lastly I would like to thank my family for the endless encouragement and support without whom my goals would have been too far to reach.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS........................................................................................................... iii
TABLE OF CONTENTS .................................................................................................................. iv
LIST OF TABLES ............................................................................................................................ viii
LIST OF FIGURES ........................................................................................................................... ix

CHAPTER 1: INTRODUCTION AND BACKGROUND ..................................................................... 11

1.1 Introduction .......................................................................................................................... 11
1.2 Plant Cell Wall .................................................................................................................... 11
1.3 CELLULOSE SYNTTHASES and the CELLULOSE SYNTTHASE LIKE gene family ................. 14
1.4 Cellulose Synthesis Accessory and Interactor Proteins ...................................................... 16
1.5 CESA mutants and Cellulose Biosynthesis inhibitors as tools for better understanding cellulose biosynthesis ................................................................. 17
1.6 Cellulose to Biofuel ............................................................................................................ 21
1.7 Biofuel and the focus on grasses ........................................................................................ 23
1.8 Aims of the dissertation ...................................................................................................... 24

CHAPTER 2: SETARIA VIRIDIS CELL WALL INVESTIGATION AND IDENTIFICATION OF THE CELLULOSE SYNTTHASE ENZYMES ........................................ 25

2.1 Summary ............................................................................................................................. 25
2.2 Introduction ......................................................................................................................... 25
2.3 Results ................................................................................................................................ 27

2.3.1 Setaria cell wall compositional analysis ....................................................................... 27
2.3.2 CESA gene identification and analysis ......................................................................... 28
2.3.3 Setaria CESA expression and cloning ....................................................................... 30
2.3.4 Setaria Transformation ............................................................................................... 30
2.3.5 At transformation with Setaria constructs ................................................................... 32

2.4 Discussion ............................................................................................................................ 33
2.5 Materials and methods ....................................................................................................... 34

2.5.1 Cellulose Quantification ............................................................................................. 34
2.5.2 Acid-insoluble lignin measurement ............................................................................. 34
2.5.3 Phylogenetic analysis ................................................................................................. 35
2.5.4 Setaria cesa cloning ..................................................................................................... 36
2.5.5 Setaria transformation (from the Boyce Thompson Institute, Cornell University) .......................................................... 36
2.5.6 Arabidopsis transformation .......................................................... 37
2.6 Tables and figures ........................................................................ 38

CHAPTER 3: SORGHUM CELL WALL COMPOSITIONAL ANALYSIS AND MUTANT CHARACTERIZATION ......................................................... 48

3.1 Summary ..................................................................................... 48
3.2 Introduction ................................................................................ 48
3.3 Results ....................................................................................... 49
  3.3.1 Identification of the REDforGREEN (RG) mutant .................. 49
  3.3.2 RG mutant displays hyper-accumulation of pigments and reduced plant height ........................................................................................................ 49
  3.3.3 Cell wall composition in the RG mutant reveals modified lignification ...... 50
  3.3.4 Saccharification of total biomass reveals RG mutant influences digestibility .. 51
  3.3.5 Hemicellulose but not cellulose are altered in the RG mutant ............ 52

3.4 Discussion .................................................................................. 53
3.5 Materials and methods ................................................................. 56
  3.5.1 Lignin content determination ................................................... 56
  3.5.2 Neutral sugar analysis .............................................................. 56
  3.5.3 Cellulose estimation ................................................................ 56
  3.5.4 Micro-scale saccharification of biomass .................................... 56
  3.5.5 Semi-purified cellulose digestion ............................................. 57
  3.5.6 X ray diffraction analysis ......................................................... 57

3.6 Figures ....................................................................................... 59

CHAPTER 4: CHARACTERIZATION AND ANALYSIS OF THE HETEROLOGOUS EXPRESSION MUTANT HULK .................................................... 66

4.1 Summary ..................................................................................... 66
4.2 Introduction ................................................................................ 66
4.3 Results ....................................................................................... 67
  4.3.1 Phenotypic analysis ................................................................. 67
  4.3.2 Cellulose analysis ................................................................. 68
4.3.3 Lignin analysis .............................................................. 69
4.3.4 ACC test for ethylene sensitivity ..................................... 69
4.3.5 Insertional position (genome walking) .............................. 70

4.4 Discussion ........................................................................ 71

4.5 Materials and methods ..................................................... 72

4.5.1 Cellulose analysis .......................................................... 72
4.5.2 Lignin analysis .............................................................. 72
4.5.3 ACC testing .................................................................. 72
4.5.4 Identification of insertional position ................................. 72

4.6 Tables and figures ............................................................. 74

CHAPTER 5: CSLD1 INTERACTS WITH THE N-TERMINAL ZINC FINGER OF CESA5 
IN VITRO AND CONTRIBUTES TO MUCILAGE ADHERENCE AND CELL WALL 
SYNTHESIS IN THE ARABIDOPSIS SEED COAT .............................. 78

5.1 Summary .......................................................................... 78
5.2 Introduction ....................................................................... 78
5.3 Results ................................................................................ 80

5.3.1 Development of in vitro screening platform for CESA interacting proteins ..... 80
5.3.2 Yeast two hybrid and BiFC confirms the interaction of CESA5 and CSLD1 ... 81
5.3.3 Mucilage adherence requires CESA5 and CSLD1 during Arabidopsis seed development ................................................................. 81
5.3.4 Histochemical analysis of cesa5, csld1 and wild type epidermal seed coat 
revealed irregular morphogenesis .............................................. 82
5.3.5 Radial wall thickness and height are reduced in cesa5 and csld1 ....... 83
5.3.6 cesa5 and csld1 seeds have less cellulose as compared to wild type .... 83
5.3.7 Rhamnogalacturonan I of seed mucilage increased significantly in csld1 as 
compared to wild type .............................................................. 84
5.3.8 csld1 phenocopied cesa5 in cellulose staining patterns ................ 84
5.3.9 CSLD1 and CESA5 co expresses during seed development .......... 85

5.4 Discussion ......................................................................... 85

5.5 Materials and methods ........................................................ 87

5.5.1 Phage display .................................................................. 87
5.5.2 Zn-Binding capacity of N terminal CesA5 and Mut- CesA5 .......... 87
5.5.3 Y2H confirmation of CESA5 and CSLD1 interaction ........................................... 87
5.5.4 BiFC confirmation of CESA5 and CSLD1 interaction ........................................... 88
5.5.5 Plant material and growth conditions ................................................................. 88
5.5.6 Histochemical analysis ......................................................................................... 89
5.5.7 Scanning Electron Microscopy and Image Analysis ............................................. 89
5.5.8 Confocal Microscopy ......................................................................................... 89
5.5.9 Cellulose Staining ............................................................................................. 90
5.5.10 Mucilage Precipitation .................................................................................... 90
5.5.11 Cellulose Content and monosaccharide analysis in Arabidopsis Seeds .......... 90
5.5.12 Relative transcript abundance analysis ............................................................ 92
5.5.13 Statistical Analysis ......................................................................................... 92

5.6 Tables and figures ............................................................................................... 93

CHAPTER 6: LACCASES AND MICRNORNA INVOLVEMENT IN LIGNIN
POLYMERIZATION ...................................................................................................... 109

6.1 Summary ............................................................................................................. 109
6.2 Introduction ........................................................................................................ 109
6.3 Results ................................................................................................................ 111
  6.3.1 Construct design and reasoning ...................................................................... 111
  6.3.2 Phenotypic characterization .......................................................................... 112
  6.3.3 Alternative areas of potential interest .............................................................. 113
6.4 Discussion ............................................................................................................ 113
6.5 Materials and methods ....................................................................................... 115
  6.5.1 Mutated laccases and microRNA overexressor constructs ................................. 115
  6.5.2 Arabidopsis transformation ............................................................................. 115
  6.5.3 PEG dehydration ............................................................................................ 116
  6.5.4 PEG dehydration seedling assay ..................................................................... 116
  6.5.5 Tetrazolium salt uptake .................................................................................. 116
6.6 Tables and figures ............................................................................................... 117

CHAPTER 7: CONCLUDING REMARKS .................................................................... 121
REFERENCES ............................................................................................................ 123
VITA ........................................................................................................................... 138
LIST OF TABLES

Table 2.1  *Setaria* CESA coding. ........................................................................................................38
Table 2.2 Primers utilized for the amplification and cloning of select *Setaria* CESAs and promoters ..................................................................................................................39
Table 2.3 Calli induction media. ...........................................................................................................40
Table 5.1 List of primers. ..........................................................................................................................93
Table 5.2 *csl1d1-1, csl1d1-2* and *cesa5* exhibit aberrant cell wall polysaccharides compared with wild type. .........................................................................................................................94
Table 6.1 Table of microRNAs with their corresponding laccase(s) under their control. .................................................................................................................................................................117
Table 6.2 PEG germination and dehydration assay media and overlay solution composition. .........................................................................................................................................................118
LIST OF FIGURES

Figure 2.1 Immature and mature Setaria viridis, cellulose and lignin quantification......41
Figure 2.2 Protein alignment of Arabidopsis CESA7 (AT5G17420) and Setaria Italica CESA “S2” (Si028761)........................................................................................................42
Figure 2.3 Phylogenetic tress of CESA proteins (A) and CESA genes (B).....................43
Figure 2.4 Quantitative real time PCR on select Setaria CESAs......................................44
Figure 2.5 Sample images of the various tissue samples utilized in the qRT-PCR. .........45
Figure 2.6 Evidence of loci on calli that were transformed..........................................46
Figure 2.7 Gus staining patterns of Arabidopsis plantlets expressing the S2 and A2 promoter GUS fusion. ................................................................................................................47
Figure 3.1 RG phenotype development (basipetal progression of pigment accumulation in RG leaves)...........................................................................................................60
Figure 3.2 Plant height of RG and wild type for two field seasons.................................61
Figure 3.3 Internode and node assessment........................................................................62
Figure 3.4 Gene expression analysis for key genes in the lignin biosynthetic pathway. .63
Figure 3.5 Lignin quantification from the leaf and stem of RG and wild type sorghum. .64
Figure 3.6 Cellulose and hemicellulose quantification....................................................65
Figure 4.1 Rosette size and diameter. .............................................................................74
Figure 4.2 Mature plant phenotype.................................................................................75
Figure 4.3 Cellulose and lignin quantification of HULK ..................................................76
Figure 4.4 ACC hypocotyl elongation measurements......................................................77
Figure 5.1 Amino acid sequence alignment of the N terminus of CESA5 and mutCESA5. ..............................................................................................................................95
Figure 5.2 SDS Page gel and attachment to microtiter plates.........................................96
Figure 5.3 Phage titer for each biopanning round and gel images of plaques. ...............97
Figure 5.4 Final biopanning hits.......................................................................................98
Figure 5.5 CESA5 interaction with CSLD1 in split-ubiquitin membrane-based yeast two-hybrid system...........................................................................................................99
Figure 5.6 BiFC transient expression illustrating interaction of CSLD1 with CESA5 in Nicotiana benthamiana leaf epidermal cells.........................................................100
Figure 5.7 Altered seed mucilage in csld1. ................................................................101
Figure 5.8 Assessment of cell shape and morphogenesis of seed coat..........................102
Figure 5.9 Altered cell morphogenesis of seed coat assessment though irregular cell angle. .........................................................................................................................104
Figure 5.10 Reduction in radial wall height in csld1.......................................................105
Figure 5.11 Acid Insoluble crystalline cellulose content of seeds.................................106
Figure 5.12 Aberrant mucilage adherence in csld1 and cesa5........................................107
Figure 5.13 Relative quantitative expression of CSLD-1 and CESA5 during wild type seed development. .................................................................108
Figure 6.1 Mature plant height ........................................................................119
Figure 6.2 Tetrazolium salt uptake in seeds.........................................................120
CHAPTER 1: INTRODUCTION AND BACKGROUND

1.1 Introduction

Cellulose is a polysaccharide made up of D-glucose, a simple sugar that is linked by a β1,4 glycosidic bond, which positions the glucose monomers in a 180° angle with its neighboring glucose molecules. It was first described and the molecular formula derived by a French chemist, Payen in 1838 who described a fibrous plant derived solid that was resistant to acid treatment (Payen, 1838; Klemm et al., 2005). Payen derived the molecular formula as $C_6H_{10}O_5$, noting its similarity to the more familiar starch, another sugar.

Humans have a significant connection with cellulose, with a considerable portion of society being shaped by cellulosic material since ancient times (Klemm et al., 2005). There is evidence of cellulosic material from hemp being used for rope and for producing fabric for clothes since 4000BC, as well as papyrus, in Egypt on which numerous documents were written, allowing for the modern age to have an insight on ancient times (Hon, 1994). Cellulose is considered to be the most abundant organic polymer on earth, with the raw source still being used to produce paper and cardboard while industrial uses have also taken off, utilizing cellulose derivatives in food additives to coatings and laminates (Klemm et al., 2005). Although cellulose produced by all land plants, it is well documented that algae, fungi and some tunicates produce cellulose, not to mention the biofilm produced by bacteria including the bacteria used in the making of the popular fermented drink, kombucha.

1.2 Plant Cell Wall

The cell wall is a crucial component of the plant, often used as an example of the main difference between a mammalian and a plant cell (Sampathkumar et al., 2011). The wall is responsible for the plants’ ability to grow upright, as well as being involved in cell differentiation, cell to cell communication and a front line for plant defense (Cosgrove, 2005). The plant cell wall is a complex structure that can largely be divided into the middle lamella, the primary and the secondary cell wall. As the middle lamella is a pectin rich area
in between the primary walls of two cells, we will focus our attention on the primary and secondary cell wall.

The primary cell wall is a hydrated matrix composed of cellulose with soluble polysaccharides in the form of pectin and hemicelluloses. The interaction, rearrangement, and biochemical changes between these components give the cell wall its rigid, yet extensible architectural characteristic. As the name implies, the primary cell wall is the first wall to be deposited in the plant cells as they divide and grow. Anisotropic (directional) elongation is the primary way in which the plant cells elongate and grow in size (Powell and Lenhard, 2012). This directional elongation is driven mainly by turgor pressure (osmotic pressure of the cell) in which the cellulose is deposited in a perpendicular manner to the axis of growth (Geitmann and Ortega, 2009; McFarlane et al., 2014).

The strongest element in the plant cell wall is a network of long cellulose microfibrils which are composed of β 1, 4 linked D-glucose units that are linked end to end in a linear fashion. These chains then use hydrogen bonding and Van der Waals forces to crystalize into the cellulose microfibrils. Unlike the other polysaccharides in the in the cell wall, cellulose is synthesized in the plasma membrane, while the others are made in the golgi and then delivered to be deposited in the cell wall (Scheller and Ulvskov, 2010; Worden et al., 2012; Atmodjo et al., 2013).

The other two components of the primary cell wall- pectin and the hemicellulososes are crucial in their roles to tether and crosslink the cellulose microfibrils. Pectin is generally made of a galacturonic acid backbone, or contains galacturonic acid and has been found to associate and interact with the cellulose microfibrils through hydrogen bonding and is more associated with the dicots while the monocot contain very little (approximately 5% dry weight) (Ishii, 1997; Zykwinska et al., 2005; Vogel, 2008; Wang et al., 2012a).

In addition to the difference in the quantity of pectin, it is in the difference of the composition of the hemicellulose component that differentiate the type I and type II primary cell walls (McCann et al., 1990; Carpita and Gibeaut, 1993). Hemicelluloses are composed of a heterogenous group of pentose (xylose, arabinxylose) and hexose (mannose, glucose, galactose, galacturonic acid) sugar backbones that are generally linked in a β 1,4 linkage similar to cellulose, although the grasses also utilize β (1,3)(1,4) linked glucans
and add stability and strength to the primary cell wall (Smith and Harris, 1999). Unlike cellulose, the hemicelluloses are synthesized in the golgi and then delivered via vesicle trafficking across the plasma membrane to the cell wall (Liepman et al., 2005; Sterling et al., 2006; Jensen et al., 2008). Type I primary cell walls are found in dicots and liliaceous monocots and their most abundant hemicellulose fraction is xyloglucan, which makes up approximately 20-25% of the primary wall but also contain mannan and xylan (Carpita and Gibeaut, 1993; Mohnen et al., 2008; Scheller and Ulvskov, 2010; Sampathkumar et al., 2011). Xyloglucans have the same backbone as cellulose (β 1,4 linked D-glucose units), however in the case of xyloglucans the backbone is decorated by other sugars such as xylose, galactose and fucose. Meanwhile in the type II primary cell wall- which are found in the commelinid monocots, which includes the grasses, contain glucuronoarabinoxylan (GAX) as their main hemicellulose, along with xylan and mixed link glucans, another unique trait of the grasses (Carpita, 1996; Mohnen et al., 2008; Vogel, 2008; Sampathkumar et al., 2011).

The secondary cell wall is deposited after the establishment of the primary cell wall and after the plant cell has stopped expanding and is deposited between the already present primary cell wall and the plasma membrane (Turner et al., 2001). The secondary cell wall is not deposited in all plant cells but rather deposited in cells that require additional reinforcement such as xylem cells, vascular bundles and seed coat cells as well as structures that require water resistance or pathogen resistance (Sarkanen and Ludwig, 1971; Cosgrove, 2005; Mohnen et al., 2008; Haughn and Western, 2012). The secondary cell wall composition is different than that of the flexible primary cell wall with less pectin and soluble polysaccharides and include cellulose, xylans and lignin- a major part of the secondary cell wall (Mohnen et al., 2008; Sampathkumar et al., 2011).

Lignin is a heteropolymer that is made from p- coumaryl, coniferyl and sinapyl hydroxycinnamyl alcohol monomers (monolignols) that are methoxylated to varying degrees (Freudenberg and Neish, 1968; Boerjan et al., 2003). The monolignols are derived from the phenylpropanoid pathway which starts with the modification of the products of the shikimate pathway, mainly phenylalanine that in turn produces various products that are vital for plant defense as well as structural support (Vogt, 2010; Fraser and Chapple, 2011). The phenylpropanoid pathway is elaborate with various pathways branching off of it, such as the anthocyanin biosynthetic pathway (Shi and Xie, 2014). The
three monolignols are produced through the deamination of phenylalanine by PHENYLALANINE AMMONIA LYASE (PAL) which produces cinnamic acid, which is then hydroxylated by CINNAMATE 4-HYDROXYLASE (C4H) to produce p-coumaric acid and the aromatic ring is successively hydroxylated in addition to phenolic o-methylation and alteration of the side chain carboxyl residue to an alcohol by a large number of enzymes such as 4-COUMARATE:CoA LIGASE (4CL), \(p\)-HYDROXYCINNAMOYL-CoA: QUINATE/SHIKIMATE \(p\)-HYDROXYCINNAMOYLTRANSFERASE (HCT), \(p\)-COUMARATE 3-HYDROXYLASE (C3H), CINNAMOYL-CoA REDUCTASE (CCR), FERULATE 5-HYDROXYLASE (F5H), CAFFEIC ACID \(O\)-METHYLTRANSFERASE (CCoAOMT) CAFFEIC ACID \(O\)-METHYLTRANSFERASE (COMT) and CINNAMYL ALCOHOL DEHYDROGENASE (CAD) (Boerjan et al., 2003). The resulting monolignols then undergo a radical coupling reaction that starts with the oxidation of the monolignols that involve the peroxidases and the laccases which then allow for the lignin monomers to be incorporated in the growing chain of the lignin heteropolymer, and will be discussed in more detail in chapter 6 (Freudenberg and Neish, 1968; Sarkanen and Ludwig, 1971; Boerjan et al., 2003). Once incorporated, the \(p\)-coumaryl monolignol are referred to as \(p\)-hydroxylphenyl (H) units, coniferyl monolignols as guaiacyl (G) units and sinapyl monolignols as syringyl (S) phenylpropanoid units that in turn create the three varieties of lignin: H, G and S lignin (Boerjan et al., 2003; Vanholme et al., 2010). Alteration in a number of the enzymes mentioned above have been linked to alteration in the H:G:S lignin ratios in plants (Franke et al., 2002; Coleman et al., 2008; Weng et al., 2010). As a generalization, dicots mainly contain G and S lignin with very little as H lignin, while monocot grasses contain approximately the same ratio of S and G lignin but also contain a much higher amount of H lignin (Baucher et al., 1998; Boerjan et al., 2003).

1.3 CELLULOSE SYNTHASES and the CELLULOSE SYNTHASE LIKE gene family

Cellulose is synthesized at the plasma membrane (PM) by a multi-protein complex referred to as the cellulose synthase complex (CSC). CELLULOSE SYNTHASE A proteins (CESA) are the processive glycosyltransferases responsible for catalyzing the conversion of UDP-glucose to cellulose (Kimura et al., 1999). CESA gene numbers are not consistent between plant species, but in the model plant *Arabidopsis thaliana*, there are 10 CESAs. The CESA s are a part of a larger superfamily of CELLULOSE SYNTHASE-LIKE (CSL) genes with ten other gene families starting with CSLA to CSLJ (Richmond and Somerville,
2000; Cosgrove, 2005; Yin et al., 2009). The CSL family including the CESAs share common features including the D,D,D,Q,X,X,R,W motif that is highly conserved in processive glycosyltransferases and believed to be the active site of the enzymes (Saxena and Brown, 1995; Wolf et al., 2012). As the hemicelluloses have a β 1,4 linked backbones, they are often considered to be made by the CSL family members (Delmer, 1999; Holland et al., 2000; Carpita, 2011). Indeed, it has been found that CSLCs are responsible for the xyloglucan backbone, and the CSLHF/H have been found to synthesize the β (1,3)(1,4) glucans found in monocots although there is no evidence that all of the CSL family members are involved in hemicellulose biosynthesis and there have been reports that show that the CSLDs are involved in cellulose biosynthesis, which will be discussed in more detail in later sections (Hazan et al., 2002; Cocuron et al., 2007; Penning et al., 2009; Park et al., 2011). In the CESAs, there is also the additional Zn finger domain with the motif: CX\(_{5}\)CX\(_{2}\)FXACX\(_{2}\)PXCX\(_{2}\)CXEX\(_{5}\)GX\(_{5}\)CX\(_{2}\)C (X taking the place of any of the amino acids), usually referred to as the CxxC motif that is only present in the CESAs (Richmond, 2000). The active site (a large cytosolic catalytic domain approximately 500 amino acid residue in size) is flanked by two transmembrane domains toward the N terminal in addition to a Zinc finger domain as well as by six transmembrane domains towards the C terminal of the protein (Arioli et al., 1998; Delmer, 1999; Saxena and Brown, 2005; Somerville, 2006). Much of what we know about the structure of the CESAs have been through the work done in the bacteria *Rhodobacter sphaeroides*, in which the crystalized structure of the CSC complex was determined (Morgan et al., 2013).

The CSC assembly takes place in the Golgi (Haigler and Brown Jr, 1986) and is then transported to the PM via the trans-Golgi network and ultimately by cortical microtubule-assisted vesicle trafficking (Taylor et al., 2003; Saxena and Brown, 2005; Paredez et al., 2006; Desprez et al., 2007; Persson et al., 2007; Crowell et al., 2009; Gutierrez et al., 2009; Gonneau et al., 2014). The acentrosomal microtubule array is a dynamic network of α-tubulin and β-tubulin dimers that form protein polymers (Lloyd and Chan, 2004). The microtubule array created molecular rails for PM localized CSCs (Heath, 1974; Paredez et al., 2006; Wightman and Turner, 2008; Li et al., 2015). A series of genetic experiments have shown that three different CESAs are needed to form a functional CSC (Taylor et al., 2003; Desprez et al., 2007; Persson et al., 2007). For the biosynthesis of the primary wall, it has been found in *Arabidopsis* that CESA 1 and 3 are critical with null mutants being gametophytic lethal, while the third position in the primary
CSC can be occupied by CESA 6 or a CESA 6 like (CESA 2, 5, 9 and 10) which are very closely related and participate in a tissue specific manner (Burn, 2002; Hématy and Höfte, 2006; Desprez et al., 2007; Persson et al., 2007). Due to the ability of the CESA 6 like to take the position of CESA6 in the CSC, CESA6 null mutants have more subtle non-lethal phenotypes (Hématy and Höfte, 2006). In the secondary cell wall, the active CESAs include CESA 4,7 and 8. Furthermore, freeze fracture electron microscopy images have revealed that the PM bound CSC is a hexameric rosette-shaped complex (Saxena and Brown, 2005). It is believed that the CESA proteins in each subunit organize into a heterotrimeric complex (Desprez et al., 2007) that possibly involves a stoichiometry of 1:1:1 between these three different CESA subunits (Gonneau et al., 2014; Hill et al., 2014). Rationalizing the number of active CESAs in a CSC has been guided by estimates of the numbers of glucan chains in a microfibril. However, this is still an area of debate and has been revised from a commonly cited 36 glucan chains in a microfibril to an estimate of 18, which has been given a stronger backing by the evidence showing the formation of stable trimers of CESAs seen through small angle neutron and x ray scattering (Fernandes et al., 2011; Vandavasi et al., 2016). In addition, looking at the cross sectional area of the cellulose microfibril, the most current estimation from the x ray diffraction data suggests a cross sectional area of 7 nm$^2$ rather than 11.6 nm$^2$ as was previously believed (Fernandes et al., 2011; Thomas et al., 2013). This would eliminate the possibility of a 36 chain microfibril, as the area is too small to accommodate 36 chains and thus the best fit that is currently being considered is a 18 chain microfibril, implying 18 CESAs in six group of threes to create a rosette (Fernandes et al., 2011; Newman et al., 2013; Thomas et al., 2013). Recently, there has also been evidence of a single CESA from *Populus tremula x tremuloides* (hybrid aspen) has been able to produce a cellulose microfibril in an *in vitro* system (Purushotham et al., 2016).

### 1.4 Cellulose Synthesis Accessory and Interactor Proteins

The deposition of the cell wall is complicated and dynamic requiring controlled cell loosening in which cross linkages are broken and then reformed at an alternative position to accommodate the cell growth (Wolf et al., 2012). Disruption of cellulose biosynthesis or alteration of microfibril alignment in the cell wall causes loss of directional cellular expansion, resulting in cells becoming radially swollen and growth organs becoming dwarfed. The CSC rosettes require numerous accessory proteins for cellulose
biosynthesis in plants, such as KORRIGAN, COBRA and CELLULOSE SYNTHASE INTERACTING protein 1 (CSI1/ POM2), SYP61 and CSI3. These additional proteins are either required at the plasma membrane to work in conjunction with the CSCs or are involved in trafficking of the CSCs or other cell wall components to the plasma membrane and cell wall.

KORRIGAN, an endoglucanase, physically interacts with the CSC and is thought to offer an editing role for the arising cellulose strands as it travels with the CSCs along the microtubule (Roudier et al., 2005; Lei et al., 2014; Vain et al., 2014). kor mutants have been shown to have issues in cell elongation and cytokinesis (Nicol, 1998; Lane et al., 2001; Mansoori et al., 2014; Vain et al., 2014). COBRA, a glycosyl-phosphatidyl inositol-anchored protein, is also required for cellulose biosynthesis and has been found to be highly co-expressed with the primary CESAs (Lane et al., 2001; Persson et al., 2005). While COBRA’s catalytic role remains unclear it has recently been found that it is critical to maintaining cellulose structure (Sorek et al., 2014). In addition, the microtubule-CSC binding protein complex CSI-1/POM2 and CSI3, which is thought of as a molecular cross linker (Gu et al., 2010; Bringmann et al., 2012; Li et al., 2014) was also found to interact with the CSC (Gu et al., 2010; Bringmann et al., 2012). Syp61, another accessory protein was found to be important in the trafficking of the CSCs to the plasma membrane (Drakakaki et al., 2012).

In addition, there have been additional CESA interactors that have been identified such as the Tracheary Element Differentiation Related 6/7 (TED6 and TED7) known to be involved in secondary cell wall formation in the xylem vessels have been shown to interact with the CSC (Endo et al., 2009; McFarlane et al., 2014). Recently, specific members of the CSL family have also been identified as CESA interactors, allowing for new and exciting findings to take place.

1.5 CESA mutants and Cellulose Biosynthesis inhibitors as tools for better understanding cellulose biosynthesis

Much of what we have come to understand about cellulose biosynthesis comes from looking at cell wall mutants and cellulose biosynthesis inhibitors (CBIs). Often times, CESA mutants are screened and identified in a mutagenized population based on their
short stature, or resistance to a known CBI or in the case of the irregular xylem mutants, a collapsed xylem (Turner and Somerville, 1997). In the secondary cell wall, the *irregular xylem* (*irx*) mutants have been indispensable toward understanding cellulose biosynthesis (Turner and Somerville, 1997). In particular, *irx1,3 and 5* mutants which correspond to mutations in *CESA8,7 and 4* were crucial to the finding that all three CESAs are required for secondary cellulose biosynthesis (Taylor et al., 1999; Taylor et al., 2000; Taylor et al., 2003; Somerville, 2006). A considerable breakthrough in examining cellulose biosynthesis was achieved almost a decade ago with the functional complementation of the procuste-cesa6 mutant with a translational fusion between YFP and CESA6, driven by its native promoter (Paredez et al., 2006). This, along with advanced laser scanning (or spinning disc) confocal imaging systems, enabled the quantitative assessment of CESA behavior in living cells.

As chemical inhibitors of cellulose biosynthesis, in addition to their uses in cell wall research, CBIs are useful for weed control in agriculture and are particularly used as pre-emergent herbicides in recreational lawns, golf courses, orchards, vineyards, and railroad tracks with a combined multi-billion dollar value. Herbicides that fall that inhibits cell wall (cellulose) biosynthesis are classified by the Herbicide Resistant Action Committee (HRAC) as group L herbicides.

For a compound to be classified as a CBI, it must meet three criteria: 1) treated seedlings exhibit characteristic CBI symptomology of stunted growth and radial swelling in rapidly expanding tissue, where ectopic lignification is sometimes evident 2) reduced cellulose content in a dose-dependent manner; and 3) rapidly (<2 h) inhibit the incorporation of 14C-glucose into the cellulose fraction of cell walls. The complexity of cellulose biosynthesis makes it difficult to further elucidate the potential inhibitory mechanisms of CBIs. Interestingly, but perhaps not surprising considering the multiple proteins involved in cellulose biosynthesis, different CBIs caused markedly different symptoms. Although this adds to the complexity in our attempts to try and to better understand cellulose biosynthesis, the CBIs are crucial tools for us to tease apart the components and mechanisms for cell wall biosynthesis. To try and use this advanced imaging data to classify CBIs and to further our understanding of the CESAs and of their interactors a categorization system based on how a given CBI disrupts the normal mobility and localization of fluorescently labeled CESA particles was developed (Brabham and
Debolt, 2012). Below we elaborate on the classification system in which the CBIs are divided into three different groups based on their effect on the CESAs and its potential use in understanding newly identified CBIs and complexity of cellulose biosynthesis.

Compounds in Group 1 are based on the phenotype (cellular mode of action (MOA)) of fluorescently labeled CESA-containing CSCs being depleted from the PM focal plane and concomitantly accumulate in cytosolic vesicles. Furthermore, fluorescently labeled CESAs are visually being produced in the Golgi (donut-shaped fluorescence in images), but in one way or another fail to reach and be inserted into the PM. This was demonstrated clearly for the well studied CBI isoxaben (Paredez et al., 2006; Gutierrez et al., 2009). Compounds in this group include isoxaben (Gutierrez et al., 2009), quinoxyphen (Harris et al., 2012b), AE F150944 (Gutierrez et al., 2009), CGA 325’615 (Crowell et al., 2009), thaxtomin A (Bischoff et al., 2010) and two new compounds CESTRIN (Worden et al., 2015) and acetobixan (Xia et al., 2014). The molecular target of some members of Group 1 has been directly associated with CESAs. Ethyl methyl sulfonate (EMS) mutagenized populations of Arabidopsis seed were screened to look for individuals resistant to the CBIs. Researchers have mapped resistance to multiple point mutations in AtCESA1, 3 or, 6 that confer resistance to isoxaben (Scheible et al., 2001; Desprez et al., 2002; Sethaphong et al., 2013) or quinoxyphen (Harris et al., 2012b; Sethaphong et al., 2013). Resistance to other group 1 CBIs have also been mapped to point mutation in AtCESA1 and AtCESA3 (Austin et al., 2011; Shim, 2014). An alternative scenario where the PM can become devoid of CESAs is a result of severe alteration in the trafficking of CESA-containing vesicles between the trans-Golgi network and the PM (Drakakaki et al., 2011; Worden et al., 2015). Delivery of CESA-containing vesicles to the PM is a highly coordinated process and is facilitated by microtubules and cargo transport proteins. Several advances have recently been made in this research area (Crowell et al., 2009; Gutierrez et al., 2009; Gu et al., 2010; Drakakaki et al., 2012; Bashline et al., 2013; Lei et al., 2013). Cortical microtubules, CSI1, CSI3, an adaptin protein μ2 and the SYP61 have all been shown to partially coincide with CESA-containing vesicles indicating their importance in CESA trafficking and therefore could be viable CBI targets as well as for further research.

The second CESA classification group contains DCB (2,6-dichlorobenzonitrile) and indaziflam. The CESA phenotype induced by these CBIs is interesting in that more
CESA particles accumulate at the PM (Herth, 1987; DeBolt et al., 2007b). Accompanying this increase in CESA abundance is an almost complete reduction of CSC velocity (DeBolt et al., 2007b), and while indaziflam (Brabham et al., 2014) too accumulated more PM bound CESA particles it elicits some interesting differences. DCB caused increased accumulation at specific foci at the PM focal plane resulting in brighter and brighter fluorescent ‘dots’ over a time series of 2 h. By contrast, indaziflam induced a more even distribution of particles across the PM. Another variation from DCB was that indaziflam treatment resulted in a reduction rather than cessation of CESA particle movement at the PM. Obtaining resistant mutants using a mutagenesis approach has been unsuccessful for both and only a modestly tolerant (2–4X) DCB mutant DH75 has been reported, which would be interesting to examine further (Heim et al., 1989).

Group 3 CBIs alters the trajectory of CESA particles to and at the PM of which the main CBI is morlin (7-ethoxy-4-methyl chromen-2-one) (DeBolt et al., 2007a). Notably, morlin has the potential to elicit its primary influence on microtubules, which could in turn influence trajectories of CSCs at the PM. While indirect evidence also exists for another CBI named cobtorin (4-[(2-chlorophenyl)-methoxy]-1-ntrobenzene) (Yoneda et al., 2007; Yoneda et al., 2010), it has not been used in combination with YFP::CESA6 or other live cell CSC reporter. Cobtorin alters the methylation ratio and the distribution of pectin in the cell wall and was hypothesized to act by interfering with cellulose pectin associations. Resistance to cobtorin was also conferred by overexpressing a pectin biosynthetic gene (Yoneda et al., 2007; Yoneda et al., 2010). We tentatively place cobtorin as a Group 3 CBI based on existing cellulose and microtubule imaging data (Yoneda et al., 2007) but further work is needed to examine it in Arabidopsis cells expressing YFP:CESA6 and compare its MOA with morlin.

There has been a general trend for CBIs to inhibit dicot root elongation at lower rates as compared with monocots (Corio-Costet et al., 1991; Sabba and Vaughn, 1999). This peculiarity could be due to a number of reasons such as seed size, metabolism, sequestration, herbicide uptake and translocation, or differences in the genetic composition. While seed size and metabolic differences are valid rationales, studies in plant tissue cultures have shown that tolerance to isoxaben in soybean nor wheat callus could not be explained simply by its metabolism or metabolic fate (Corio-Costet et al., 1991). Alternatively, could the composition of the cell wall also influence CBI tolerance,
mainly the difference between type I and type II primary cell wall? When maize tissue and barley cultures (calli) are habituated in DCB and their cell wall analyzed, it was found that it was reduced in cellulose content, but increased in mixed linkage glucans or glucuronooarabinoxylan (GAX) and arabinoxylans, and it was hypothesized that the increase in cell wall phenolics could be a compensation mechanism for the cellulose impoverished cell wall’ (Shedletzky et al., 1992; Mélida et al., 2010). If the cellulose biosynthetic carbon sink is halted, where does the metabolic pool destined for cellulose production go? 14C glucose uptake studies suggest that it can be diverted to pectin and hemicelluloses (García-Angulo et al., 2012). This could be signifying a compensation mechanism in which the excess glucose is being utilized for hemi-cellulose (xyloglucan, heteroxylan) production in grasses and pectin production in dicots. With the notable differences in cell wall composition in the grasses, this diversion to alternative cell wall polysaccharides caused by the CBIs could differentially influence the response. Understanding this divergence will be interesting for the cellulose biosynthesis research community but also the broader weed science community.

Combining genetics with CBIs will continue to assist in elucidating the basic mechanisms of cellulose and cell wall biogenesis, and continued development of new and current CBIs is expected to be driven by their utility in cellulose biosynthesis research but also as weed control agents. The capacity for new inhibitory mechanisms of action in the broad CBI grouping is particularly of interest due to the lack on new herbicidal MOAs developed in the past decades. Additionally, breakthroughs in advanced cellular imaging techniques will also facilitate the use of CBIs as research tools to disrupt cellulose biosynthesis in a targeted way. Beyond cellulose, using chemical genetics to dissect other cell wall processes is anticipated.

1.6 Cellulose to Biofuel

Producing liquid transportation fuel from renewable materials is of considerable interest as the Clean Energy Act of 2007, requires a 36 billion gallons of biofuel to be created in the US by 2022 which roughly calculates to approximately 7.7 fold increase as compared to what was produced in 2007, of the 36 billion gallons, approximately a third must come from cellulosic or biodiesel origin. Conversion of cellulosic material is attractive as cellulose is abundant and can be fermented to make alcohol. The general process
takes biomass (such as corn stover and switchgrass) which goes through a pretreatment process, followed by enzymatic hydrolysis/fermentation and distillation to finally produce the bioethanol. However, at each step there are issues to be resolved.

The thermochemical pretreatment stage in which the biomass is primed to allow for better access of the enzymes to its substrates in the downstream steps is an extremely costly step, both in the monetary aspect as well as the amount of energy required to pretreat the biomass (Sousa et al., 2009; Pauly and Keegstra, 2010). The pretreatment often involve weak acid treatments as well as steam explosion treatments and has been said to be one of the top expenses in the entire process (Wooley et al., 1999).

During enzymatic hydrolysis and fermentation phase, substantial enzyme loading is required for the successful fermentation due to the accessibility of the sugars (Pauly and Keegstra, 2010). In addition, there is also the issue that the yeast used in the fermentation process favors glucose and mannose (hexoses) rather than the abundant pentoses (xylose/arabinose) that comprises the hemicellulose (Pauly and Keegstra, 2010). Currently, new yeast strains are in development that will more readily and happily accept the abundant pentoses found in the biomass (Van Vleet and Jeffries, 2009).

Plant metabolic engineering has been considered to create more ideal biofuel crops. Here, various ideas to try and deal with the recalcitrance of the plant cell wall to ease cellulosic ethanol production have been evaluated. There have been efforts to alter the lignin network through genetic modification to alter the expression of select enzymes in the phenylpropanoid pathway, such as for the enzyme cinnamate 4-hydroxylase (C4H) in which the native promoter was changed that resulted in a 20% decrease in lignin without any negative effects on the plant (Yang et al., 2013). In addition, there have also been research into “redesigning” the lignin composition by introducing ester linkages within the lignin polymer, thereby reducing the biomass’s recalcitrance to pretreatments (Wilkerson et al., 2014). Some have also looked at modifying cellulose itself to make it more digestible (Harris et al., 2012).
1.7 Biofuel and the focus on grasses

Grasses (more precisely, monocot grasses in the Poaceae and Panicoidaea clades) have been studied due to their favorable growth and environmental stress tolerance traits. The Panidoideae clade of grasses use C4 photosynthetic processes and are able to thrive in environments in which the C3 plants would have a high photorespiration rate such as arid climates and as well as being more nitrogen and water use efficient (Sage, 2004; Brutnell et al., 2010). As mentioned above in the plant cell wall section of the chapter, the grass cell wall differs a little in its composition as compared to the dicots, containing glucuronoarabinoxylan (GAX) as the main hemicellulose, along with cellulose and mixed link glucans- another unique aspect of grasses (Mohnen et al., 2008; Vogel, 2008; Sampathkumar et al., 2011). Their hemicellulose composition is also differing from that of the dicots, with GAX surrounding the cellulose microfibrils in the grasses (Carpita, 1996; Vogel, 2008).

As a model species in plant research, Arabidopsis thaliana is often used as the model plant. Other species are utilized by specific research fields-such as Nicotiana benthamiana being used for plant pathogen interaction and Medicago truncatula for the study of nitrogen fixation (Goodin et al., 2008). For the grasses, Brachypodium distachyon has gained favor as the model species, however it does not utilize C4 photosynthesis. Therefore, to fill in this gap, Setaria viridis has been proposed as a new model species for C4 photosynthesis (Brutnell et al., 2010).

Setaria viridis (green foxtail) a cousin to it’s grain variety foxtail millet (Setaria italica) has been of increasing interest as potentially a new model species. Setaria viridis with its small genome size (~515Mb), rapid and prolific growth (under short day conditions, from seed to seed, it only takes half as long as Setaria italica) and in addition to being a C4 grass has made it appealing to be proposed as a model species (Brutnell et al., 2010). In particular, from a biofuel prospective, the cell wall biosynthesis and various components of the cell wall are appealing areas of research.
1.8 Aims of the dissertation

The overall goal of this dissertation is to investigate the machinery that is responsible for the biosynthesis of the cell wall. The cellulose synthase (CESA) enzymes and the compositional analysis of the cell wall of the newly proposed model plant *Setaria viridis* were initially investigated and will be discussed in chapter 2, along with the crop plant *Sorghum (Sorghum bicolor)* in chapter 3. The bioinformatics studies advanced our knowledge of the gene networks involved in biosynthesis of key cell wall polymers and led to a series of mechanistic studies. I carried on a series of molecular genetic analyses using the model plant Arabidopsis. This included looking into the role of a cellulose synthase like family genes (chapter 4) as well as an insertional mutant line that was created (chapter 5). Regulatory processes underlying lignin biosynthesis were also explored by focusing on the laccases and their interactions with their corresponding microRNA in chapter 6. Cellulose biosynthesis was explored using interspecific complementation experiments comparing grasses and dicots as well as identification of new genes involved in biosynthesis.

Copyright statement: Portions of this chapter were published in the Journal of Experimental Botany Tateno et al. (2016).
CHAPTER 2: SETARIA VIRIDIS CELL WALL INVESTIGATION AND IDENTIFICATION OF THE CELLULOSE SYNTHASE ENZYMES

2.1 Summary

Setaria viridis, more commonly known as the green foxtail is the weedy cousin of the grain crop millet (Setaria italica) and has been proposed as a satisfactory model species for the C4 grasses (Brutnell et al., 2010; Li and Brutnell, 2011). The genus Setaria belongs to the subfamily Panicoideae, whose members include major agricultural crops such as Sorghum, maize and switchgrass. The Setaria italica genome was published in 2012 (Bennetzen et al., 2012; Zhang et al., 2012). It was initially our goal to identify Setaria orthologs to known cell wall biosynthesis genes. To do this, we tracked the number and genetic structure of the CESA genes in Setaria viridis. These were examined both as genetic features and their expression pattern studied using RNA abundance. CESA genes were identified bioinformatically and a select few were cloned to be further examined in a heterologous system using Arabidopsis. In addition, attempts at a stable Setaria transformation were conducted by calli transformation.

2.2 Introduction

As the depth of knowledge ever increasingly grows in the model plant Arabidopsis and the utility of this information proves translatable, there remains a need for additional model plants, to decode and translate traits that are absent in Arabidopsis (Rensink, 2004). Monocotyledonous taxa (grasses) display both marked phenotypic and metabolic variability. Moreover, they produce the majority of the world food crops and a growing need for biomass for biofuel production is being driven toward perennial grasses.

Existing examples of model grasses include Brachypodium distachyon and rice (Oryza sativa) have served as model plants for gene discovery and translational biology (Ouyang et al., 2007; Vogel, 2008)). However, there are also important physiological differences among the grasses such as the difference between mechanisms of photosynthesis. Rice and Brachypodium utilize C3 (carbon fixation) photosynthesis whereas the existing and most promising bioenergy feedstocks including maize (Zea mays), sorghum (Sorghum bicolor), switchgrass (Panicum virgatum L.), or Miscanthus
(Miscanthus×giganteus) employ C4 (carbon fixation) photosynthesis (Schmitt and Edwards, 1981; Ghannoum, 2008; Wang et al., 2009).

**Setaria italica**, the domesticated version of *Setaria viridis* has a long history of being an important grain crop that was domesticated more than eight thousand years ago starting in China and India (Zohary et al., 2012). Its ability to grow in arid climates made it a favorable crop as it moved on to be grown in places in the middle east and further into Europe thousands of years after the original domestication event (Zhang et al., 2012; Zohary et al., 2012). In modern times, the millet is less used (although not eliminated) as a grain crop for human consumption but also has now been utilized as a source for bird seed as well as for hay and silage.

*Setaria viridis* has been gaining interest as the a new model species for the C4 grasses with a strong backing from the biofuel sector looking for alternative and renewable energy sources (Bettinger et al., 2010). Although the interest from the biofuel research aspect was that for *Setaria viridis*, originally only the *Setaria italica* genome was sequenced and initially published in 2012, although the scaffold versions were available prior to the publications (Bennetzen et al., 2012; Zhang et al., 2012). Often times, it is widely accepted that the majority of the genetic difference between the domesticated and wild versions of a plant is the mechanism of seed dispersal (Zohary et al., 2012). The trend tends to be that the wild counterpart has a tendency to more readily disperse their seed upon maturation, while the crop variety has been selectively bred for the plants to retain most of their seeds (the grain) until they are ready to be harvested. Bennetzen and colleagues also looked at the genome of *Setaria viridis*, focusing on candidate genes known to be involved in domestication events, but found that there was no significant difference even in these genes that would have been considered to be most likely to be differing. Therefore, in the bioinformatic analysis performed, we utilized the available *Setaria italica* genome while all wet bench experiments were conducted on *Setaria viridis*.

In addition to being a diploid monocot, *Setaria viridis* has multiple characteristics that are favorable for a model species such as a relatively small genome size (~515Mb) with synten to the Panicoideae grasses (Li and Brutnell, 2011) a rapid lifecycle with high seed production and ability to self-fertilize (Doust et al., 2009). Phenotypically, the annual grass has glaborous leaf blades that can be up to 40cm long and the plant can grow to
1.5m or taller in height, with dense panicles 20cm or longer full of spikelets (de Wet et al., 1979)

As the *Setaria viridis* cell wall has not been examined in great detail and as the various enzymes involved have not been characterized, our initial goals were to better understand the compositional makeup of the *Setaria viridis* cell wall and to bioinformatically characterize and identify the CESAs present in *Setaria viridis* as well as the other family members of the cellulose synthase like families. In addition, we aimed to further characterize specific *Setaria viridis* CESA genes out of the bioinformatics results that may be of particular interest.

2.3 Results

2.3.1 *Setaria* cell wall compositional analysis

Due to the interest of *Setaria* being mostly in the field of cellulosic ethanol and biofuel production, we thought it to be important to first understand the compositional make-up of the *Setaria* cell wall. It is recognized that the relative abundance of cellulose, lignin, and hemicellulose varies depending on tissue type, age, and environmental/biological condition of the plant tissue (Meinert and Delmer, 1977; Rogers and Campbell, 2004). Cellulose, a natural highly abundant macromolecule made up of glucose is a principle substrate for bioethanol conversion (Hidayat et al., 2012). By contrast, the presence of lignin inhibits enzymatic processes (Li et al., 2010; Kim et al., 2011) and its down-regulation through genetic approaches has been shown to improve the saccharification efficiency (Chen and Dixon, 2007a; Saathoff et al., 2011; Wang et al., 2012b).

Cellulose and lignin content of acid insoluble residue (AIR biomass) was quantified in *Setaria*. The cellulose content in the mature (Figure 2.1B) biomass was compared with that of immature tissue (Figure 2.1A). Results revealed an expected and consistent proportional decrease in cellulose during development (Figure 2.1C). This proportional decrease in cellulose content may be a target for future transcriptional rewiring efforts that aim to increase cellulose without significant cost to the fitness of the plant.
The majority of lignin was measured as insoluble lignin. Lignin content increased significantly between the first and second sampling in both aerial and root tissue (Figure 2.1C). Increased lignification during development was consistent with diversification of the primary cell wall into lignified secondary cell walls in maturing stems, leaves, and roots (Redfearn and Nelson, 2003; Parrish and Fike, 2005; Stork et al., 2010). Collectively, the data highlights a contiguous developmental program during the deposition patterns for cellulose and lignin in Setaria.

2.3.2 CESA gene identification and analysis

After quantifying the major cell wall composition in both immature and mature Setaria samples, the first step toward understanding the cellulose biosynthesis in Setaria was to identify how many CESA orthologs were in the Setaria genome and which known CESA genes they were most similar to. In order to ask whether similarities in cellulose biosynthesis were reflected by similar organization of the CESA gene family, we performed a phylogenetic analysis of annotated genes. To aid in providing confidence in the bioinformatic analysis, two separate phylogenetic trees were established using distinct protein prediction compared with genomic sequence. Known CESA genes from Arabidopsis and maize were utilized for the assembly of the two trees. After the initial BLAST searches, the putative CESA genes of Setaria, Sorghum and Panicum were further analyzed to determine if the CESA unique CxxC motif (Richmond, 2000) were present. The putative CESAs were lettered alphabetically in the order that we identified them (Table 2.1). Figure 2.2 represents an example of the homology seen between the CESA genes in Arabidopsis and Setaria. The protein alignment is between CESA7 in Arabidopsis and the closest Setaria homolog SvCESAS2 (as found through the phylogenetic analysis). As mentioned earlier, the CxxC motif conserved in just the CESAs are found toward the N terminal of the protein, while the QxxRW motif found in all CSL family members are found toward the C terminal of the protein sequence. Except for a very few 10-20 amino acid stretches, most of the two CESA are highly conserved. Homology-based associations were made between CESA identified in the genomic data derived from Arabidopsis (AT), Setaria italica (SI), Z. mays (ZM), Sorghum bicolor (SB), and P. vulgaris (PV). The MUSCLE multiple alignment tool (Edgar, 2004) was utilized followed by the PhyML program (Guindon et al., 2010) to create the phylogenetic tree based on the translational predictions (protein sequences) of the putative CESAs (Figure
2.3A). GARLI was utilized in parallel to produce a phylogenetic tree based on genomic sequence information (Figure 2.3B).

The phylogenetic tree inferred by protein sequences clustered in five classes with the Arabidopsis CESA1, 3, 7, 6-like, and 4. It was noteworthy that the Arabidopsis CESA6-like clade, which also comprised proteins similar to Arabidopsis CESA2, 5, and 9 (Persson et al., 2007) was the most highly represented cluster. By contrast, the Panicoideae homologs with Arabidopsis secondary cell wall CESA4 were least common, which only Z. mays and Setaria containing putative members of the CESA4-associated clade. With the exception of the Arabidopsis CESA4 clade-comprising of only two putative CESA members, the remaining clades contain at least one putative CESA member from each of the five species investigated. Using genomic data to assemble and predict phylogenetic relationships, the putative CESAs divide readily into four distinct clades, clustering around Arabidopsis CESA1 (primary cell wall; (Arioli et al., 1998)), CESA3 (primary cell wall; (Scheible et al., 2001)), CESA4/7/8 (vascular secondary cell wall; (Taylor et al., 2003)), and CESA6-like (CESA2, 5, and 9, included in primary and seed coat secondary cell wall; (Stork et al., 2010; Sullivan et al., 2011; Mendu et al., 2011b)). It was found that at least one putative CESA member from each of the five species was represented within each sub-clade. However, in the heterotrimeric model for cellulose synthase complex (CSC) formation in the primary and secondary cell wall (Delmer and Amor, 1995; Tanaka et al., 2003), the lack of at least three distinct CESAs within the CESA4, 7, and 8 sub-clade suggests either divergence among heterotrimeric associations or an example of large incompletions within the genome sequencing for all of the Panicoideae. In comparing the two phylogenetic trees, there were no marked differences in the CESAs that comprised a particular sub-clade. For example, in both phylogenetic trees the Arabidopsis CESA6-like CESA2, 5, 6, and 9 cluster closely with each other and were broadly represented. The discrepancy in the number of clades between the two trees can be attributed to the fact that within the protein sequence phylogenetic tree, the Arabidopsis CESA4 clade and CESA7 clade are separated while they are combined into one clade within the genetic sequence phylogenetic tree. From the two phylogenetic trees compiled, it was found that Setaria has ten putative CESAs. Ultimately, targeted reverse genetics are urgently needed in the model grasses to elucidate questions related to the tissue specificity, functionality, redundancy, and stoichiometry of the CESA genes.
2.3.3 *Setaria* CESA expression and cloning

From the information gathered from the phylogenetic analysis, we focused our attention on two *Setaria* CESAs, *SvCESA0D* and *SvCESA0I* from the phylogenetic tree. The reasoning behind choosing the two CESAs were due to their position on the phylogenetic tree, in which *SvCESA0D* (from now on referred to as A2 in the manuscript) was most closely homologous to the *Arabidopsis CESA3* - a primary cell wall CESA, while *SvCESA0I* (from now on referred to as S2) was closest to *Arabidopsis CESA7* - a secondary cell wall CESA. Initially, quantitative real time PCR was performed to look at the expression of S2 and A2 in various tissue types in *Setaria* (Figure 2.4 and 2.5). From the qRT, S2 in particular had significantly higher expression levels in the internode regions, which require reinforcement and are areas with secondary cell wall and lignification as well as in the roots, which are also areas requiring secondary cell wall formation, this led us to have a greater interest in S2 as one of CESAs.

As our next step, full length *Setaria* CESAs were amplified from *Setaria* genomic DNA. The full genomic sequence was used rather than the coding sequence in case the exon intron junctions may be critical in the splicing and correct expression of the CESA proteins (Table 2.2). The amplified CESAs were ligated into the pMDC32 vector, which contains the double 35S promoter for constitutive overexpression. In addition to the CESAs, their corresponding promoters were also investigated. Approximately 2000 base pairs upstream of the start codon was taken to be considered the promoter and was also amplified to be used for promoter- β-glucuronidase (GUS) fusion constructs to look at expression patterns with in the plant.

2.3.4 *Setaria* Transformation

One of the favorable traits of a model species is its ability to be easily genetically modified/ transformed, which based on our work and others was not supported experimentally. Below, efforts to transform *Setaria* and regenerate stable transformants are described, albeit unsuccessfully. Simple transformation techniques such as floral dip transformation method developed by Bent and Clough for *Arabidopsis* or the leaf disk transformations of *Nicotiana* were not found to be effective for *Setaria* at the time that we were attempting our transformations, although there have been recent publications in the
last year for *Setaria* floral dipping protocols (Horsch et al., 1989; Clough and Bent, 1998; Martins et al., 2015; Saha and Blumwald, 2016). We therefore focused on cell culturing and calli transformation. The initial task was to grow the seed under sterile conditions without contamination. To achieve this, various concentrations of bleach and sodium dodecyl sulfate SDS were tested with and without seed hull removal. Results supported that the most crucial factor was seed hull removal prior to the sterilization process (data not shown). Next, we established the germination rate of seeds and vernalization treatments that would result in the highest germination rate (data not presented). A vernalization overnight at -80°C, a 24 hour -80°C treatment as well as a 4°C treatment were compared. The highest germination resulted from the 4°C vernalization, however even with the treatment the rate of germination was only about 34%. The next step was to determine a suitable calli induction media for *Setaria*. Various vitamin, auxin and sucrose levels were considered amongst other variables. Ultimately, amongst the different medias tested, the GVS+ dicamba media seemed to be the most promising in terms of calli induction and formation (Table 2.3).

However, while looking into the germination rates and media composition ourselves, we were also shared the protocol/outline that was developed by the Van Eck group at the Boyce Thompson institute of Plant Research at Cornell University, which has now been published (Van Eck and Swartwood, 2015). As it seemed that the protocol had been streamlined and was giving positive results to the group, we decided to base our attempts at producing a stable transformant line with the shared protocol.

In brief, the institute’s media composition had a base of Murashige and Skoog salt with an addition of a supplemental vitamin solution and maltose on which the dehulled, sterilized seeds were grown in the dark to induce calli formation and subcultured after a month onto fresh media. A week after the subculturing, the calli were transformed by incubating the calli in an agrobacterium solution containing a construct of choice. After the co-incubation, the calli were dried and allowed to recover for three days in the dark after which they were transferred to calli inducing media with antibiotic selection and placed back in the dark. After approximately two weeks, the calli were transferred to plant regeneration media and put under 16-hour photoperiod conditions and transferred to fresh media every three weeks. If there were to be positive transformants (0.8cm shoots) they were transferred to root promoting media and grown in magenta boxes.
The approximate time line from calli induction to transformed plantlets was estimated to be about four to five months, followed by another two months to seed collection of the next generation. We continuously tried for 3 years following the protocol from the institute as well as trying variations to see if we would be successful in obtaining at least one stable transformant line. Many of our transformation attempts were with the *Setaria* promoter GUS fusion constructs and although there were loci on the calli that had been transformed -as indicated by the GUS staining visible on the calli (Figure 2.6), unfortunately we were not able to successfully regenerate a mature plant from the transformation attempts.

2.3.5 At transformation with *Setaria* constructs

In tandem with the *Setaria* transformation efforts, we also made stable Arabidopsis transformant lines with the *Setaria* gene constructs to study their effects in a heterologous system. We created both overexpressor constructs of the *Setaria* gene with the 2x35S promoter as well as reporter lines, using the *Setaria* promoter to drive the GUS gene. The overexpressor lines were used to determine if there would be an obvious phenotype that may suggest an increase in cellulose production or if there would be a knockdown (therefore a stunted phenotype) associated with the expression of the *Setaria CESA*. However, the overexpression did not yield any notable phenotypes (neither an increase or decrease in stature). Due to the lack of obvious phenotypes, the lines were not further examined.

The promoter GUS fusion lines on the other hand did show interesting GUS staining patterns that indicated a similar localization as expected from the phylogenetic tree and the clustering with the *Arabidopsis* CESAs on the phylogenetic tree (Figure 2.7). The A2 promoter GUS line (which has the closest homology to *Arabidopsis CESA3*) showed GUS staining in the younger developing leaves, developing inflorescence and toward the meristematic region of the stem. On the other hand, S2 (closest homology to *Arabidopsis CESA7*, a secondary cell wall CESA) displayed a completely saturated leaf as well as on the exterior of the siliques as well as on the sepals below the flower petals all pointing towards a secondary cell wall CESA expression, neatly corroborating the qRT data and the results of the phylogenetic trees.
2.4 Discussion

In examining *Setaria viridis*, we focused in particular on the cellulose synthase genes first with their identification followed by a molecular genetic approach for further investigation. Through the phylogenetic analysis, we were able to see that the *Setaria CESA* genes were similar to well characterized *CESAs* in other model species such as *Arabidopsis thaliana* (Figure 2.2), as was demonstrated by the qRT and the GUS expression (Figures 2.4 and 2.7). The *Setaria* genome contained a comparable number of *CESAs* with similar clustering based on their genomic and proteomic sequence that suggest cellulose biosynthesis for the primary and secondary cell walls (Figure 2.3). As the sample protein alignment demonstrated, the *Arabidopsis* and *Setaria* CESAs had similarities in their protein structures such as the eight predicted transmembrane helices and the large cytosolic catalytic region (Figure 2.2).

From the information gained from the phylogenetic analysis, we tried our hand at creating stable transformant lines in *Setaria* in the hopes that we would be able to further study the cellulose biosynthesis mechanism in vivo. *Setaria* agrobacterium mediated transformation had been in the early stages with various suggestions and potential protocols in circulation. After a continuous effort over three years, and testing various protocols and numerous alterations, we were still unable to produce a transgeneic line although we could see that there were loci on the calli that had been transformed (Figure 2.7).

In an effort to further investigate the CESAs of *Setaria* we turned to heterologous expression in Arabidopsis. From the promoter GUS fusion constructs, it was visible that the *Setaria CESA*s showed a similar GUS expression that was expected from the phylogenetic analysis and the clustering that we saw on the two phylogenetic trees. This indicated to us that the *CESAs* are indeed a well conserved group of genes and that their roles in cellulose biosynthesis are well conserved across plant families.
2.5 Materials and methods

2.5.1 Cellulose Quantification

Cellulose content of the samples were measured colorimetrically as described by Updegraff (1969). Samples were homogenized using a grinder (Arthur H. Thomas Co Scientific, Philadelphia, PA, USA) equipped with a 1-mm sieve. Twenty-five milligrams plant material were de-starched as by the specification in National Renewable Energy Laboratory (NREL), LAP-004 (1996) guidelines by incubating the samples in 1 ml 70% ethanol overnight at 65°C, washed twice with 1 ml 70% ethanol for 1 hour and once with 1 ml acetone for 5 min. The samples were dried overnight at 30°C with shakings. Cellulose content was determined by weighing out 5 mg of dry biomass extract in quadruplicate and boiled in acetic-nitric reagent (acetic acid: nitric acid: water 8:1:2) for 30 min to remove lignin and hemicellulosic carbohydrates (Updegraff, 1969). Remaining material was then washed twice with 8 ml MQ-water and 4 ml acetone and dried under vacuum. The cellulose samples were then hydrolyzed in 67% sulfuric acid for 1 h. The glucose content of the samples was determined by the anthrone method (Updegraff, 1969). Ten microliters of the sulfuric acid hydrolyzed samples were mixed with 490 µl water and 1 ml 0.3% anthrone in concentrated sulfuric acid on ice. The samples were boiled for 5 min then placed immediately back on ice. The absorbance of the samples was measured using a Bio-Mate Thermo Scientific spectrophotometer (Thermo Fischer, Waltham, MA, USA) set at OD 620 nm and the cellulose content was calculated by multiplying the measured glucose concentration of each sample by the total volume of the assay and then by the hydration correction factor of 0.9 to correct for the water molecule added during hydrolysis of the cellulose polymer.

2.5.2 Acid-insoluble lignin measurement

Porcelain filter crucibles containing a glass microfiber filter (Whatman 934-AH) were placed in a furnace at 575°C for 4 h. The crucibles were removed from the furnace and placed into a desiccator to cool for a minimum of 1 h before being weighed. The autoclaved samples were vacuum filtered through the crucibles. The crucibles and their contents were dried in an oven at 105°C overnight before transferring into a desiccator to cool. The weight of the crucibles were recorded and placed in a furnace at 575°C for 4 h
before reducing the temperature to 105°C overnight. The crucibles were removed and placed into a desiccator to cool and weighed.

2.5.3 Phylogenetic analysis

The sequences of the *A. thaliana* cellulose synthase genes were obtained from the TAIR database1 and were utilized as the initial reference in order to obtain the putative *CESA* genes of the grasses. The cDNA sequences of the ten cellulose synthase genes of *A. thaliana* were employed to conduct the BLASTX (Altschul et al., 1997) search within the *Setaria italica*, *Sorghum bicolor*, *Panicum virgatum*, and *Z. mays* genomes. BLASTX searches were conducted using the phytozome database provided by JGI2. Due to the very few number of putative *CESA* genes for *Panicum virgatum* that resulted from the BLASTX search, an additional BLASTP search was conducted using the 12 CESAs of the *Z. mays* to determine if there were any additional putative CESAs. After the identification of the putative *CESA* genes, the protein sequences of each of the putative genes were screened to determine if they contained the CxxC motif (Richmond, 2000) considered to be unique to the CESAs and distinguishing them from the cellulose synthase like genes. A total of 36 putative cellulose synthase genes were identified.

The putative CESA that contained the CxxC motif of the four monocot species as well as the known CESAs of *A. thaliana* were then uploaded onto the PhyML program (Guindon et al., 2010). The putative CESA protein sequences that contained the CxxC motif of the four monocot species above as well as the CESA protein sequences of *A. thaliana* were analyzed using PhyML (v.3.0; (Guindon et al., 2010)). The advanced setting within PhyML conducts a sequence alignment using MUSCLE (v.3.7; (Edgar, 2004)), then automatically utilizes the alignment to create the phylogenetic tree. A total of 500 bootstrap replicates were conducted under the WAG (Whelan And Goldman) protein substitution model (Whelan and Goldman, 2001).

In addition to the amino acid-based tree, a tree was constructed using the DNA sequences of the 36 putative *CESA* genes as well as the 10 *Arabidopsis CESA* genes. The sequences were obtained from the Phytozome database and the TAIR database, respectively. The sequences were submitted to the Guidance web server (Penn et al., 2010) and aligned with MAFFT (Katoh, 2005). Guidance is a program that both aligns
sequences and assesses confidence in each position in the alignment, in order to objectively identify and remove regions of ambiguous alignment prior to phylogenetic analysis. Any column in the alignment that had a confidence score of less than 95% was removed by Guidance. GARLI (v. 2.0; (Zwickl, 2006)) was used to analyze the resulting alignment under the GTR+I+G model (Rodríguez et al., 1990), in which a best tree search (five replicates) and a 200-replicate bootstrap analysis were conducted, both using the default settings. The resulting maximum likelihood tree was then visualized via FIGTREE3.

When grouping CESA genes or gene products within the broad CESA phylogenetic tree, it was necessary to visualize groups (sub-clades) of putative CESAs within the tree, and therefore they were referred to based on the closest homology to the established Arabidopsis CESA nomenclature of Richmond (2000).

2.5.4 *Setaria cesa* cloning

All cloning was done utilizing standard cloning methodologies. Select *Setaria* CESAs were amplified from genomic DNA and cloned into the pMDC 32 vector that uses a double 35S overexpressor promoter to drive the expression of the gene of interest with kanamycin as the bacterial selection marker and hygromycin as the plant transformant selection marker (Curtis and Grossniklaus, 2003). The full length *Setaria* CESAs were amplified with the Phusion polymerase (Thermo-Scientific) with primers listed in table 2.2 from genomic DNA and subsequently digested with Ascl and PacI restriction enzymes (Thermo-Scientific) and ligated using T4 DNA ligase (Thermo-Scientific) into pMDC32 which had been similarly digested. The ligated product was then transformed into E. coli strain DH5α. The promoters were similarly amplified out of genomic DNA and restriction sites Pmel and KpnI were utilized to digest both the PCR product and the vector, which were then ligated using T4 DNA ligase and the product subsequently transformed in to E. coli strain DH5α.

2.5.5 *Setaria* transformation (from the Boyce Thompson Institute, Cornell University)

The attempts at *Setaria* transformation followed the protocols provided by the Van Eck group at Cornell University (Van Eck and Swartwood, 2015). Briefly, the media composition had a base of Murashige and Skoog salt with an addition of a supplemental
vitamin solution and maltose on which the dehulled, sterilized seeds were grown in the
dark to induce calli formation. After three to four weeks, the calli were subcultured and
divided into smaller portions onto fresh calli inducing media. A week after the subculturing,
the calli were transformed by incubating the calli in an agrobacterium solution containing
a construct of choice. This entailed the calli to be incubated with the agrobacterium in a
liquid calli inducing media with added acetosyringone to increase virulence of the
agrobacterium. The co incubation should only last about 5 minutes after which the
agrobacterium solution was decanted and the calli were dried on sterile filter papers,
wrapped and allowed to recover for three days in the dark. After the recovery period, the
calli were transferred to calli inducing media with antibiotic selection and place back in the
dark. After approximately two weeks, the calli were transferred to plant regeneration
media and put under 16 hour photoperiod conditions and transferred to fresh media every
three weeks. If there were to be positive transformants (0.8 cm shoots) they were
transferred to root promoting media and grown in magenta boxes.

2.5.6 Arabidopsis transformation

The floral dip method developed by Bent and Clough was used for all Arabidopsis
transformations with the Setaria promoter GUS constructs as well as the pMDC 32 vectors
containing the Setaria CESA genes.

Copyright statement: Portions of this chapter were published in Frontiers in Plant Science:
Petti et al. (2013b)
2.6 Tables and figures

**Table 2.1 Setaria CESA coding.** Coding utilized on the phylogenetic trees and corresponding accession number from Phytozome.

<table>
<thead>
<tr>
<th>Phylogenetic tree coding</th>
<th>Phytozome gene Accession number</th>
<th>Phytozome protein Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI-CESA0A</td>
<td>Si034009m.g</td>
<td>Si034009m.g</td>
</tr>
<tr>
<td>SI-CESA0B</td>
<td>Si028764m.g</td>
<td>Si028764m.g</td>
</tr>
<tr>
<td>SI-CESA0C</td>
<td>Si028762</td>
<td>Si028762</td>
</tr>
<tr>
<td>SI-CESA0D</td>
<td>Si034016</td>
<td>Si034016</td>
</tr>
<tr>
<td>SI-CESA0E</td>
<td>Si028770m.g</td>
<td>Si028770m.g</td>
</tr>
<tr>
<td>SI-CESA0F</td>
<td>Si000133m.g</td>
<td>Si000133m.g</td>
</tr>
<tr>
<td>SI-CESA0G</td>
<td>Si005742m</td>
<td>Si005742m</td>
</tr>
<tr>
<td>SI-CESA0H</td>
<td>Si021050m</td>
<td>Si021050m</td>
</tr>
<tr>
<td>SI-CESA0I</td>
<td>Si028761m.g</td>
<td>Si028761m.g</td>
</tr>
<tr>
<td>SI-CESA0J</td>
<td>Si034020m.g</td>
<td>Si034020m.g</td>
</tr>
</tbody>
</table>
Table 2.2 Primers utilized for the amplification and cloning of select *Setaria* CESAs and promoters.

<table>
<thead>
<tr>
<th>primer name</th>
<th>restriction sites</th>
<th>primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2p-F-Pme1</td>
<td>pme1 (GTTCATAC)</td>
<td>GGC GTT TAA ACC ACA TTA ATC AAA GAA AA</td>
</tr>
<tr>
<td>S2p-R-Kpn1</td>
<td>kpn1 (GGTACCC)</td>
<td>GCG GGT ACC TTT GAA AAG ACA CGG GC</td>
</tr>
<tr>
<td>S2cds-F-Asc1</td>
<td>asc1 (GGCGCGGCC)</td>
<td>ATC GGC GCG CCA TGG CGG GAA AGC CAA CT</td>
</tr>
<tr>
<td>S2cds-R-Pac1</td>
<td>pac1 (TTAATTAA)</td>
<td>CAG TTA ATT AAT CAG CAA TTG ATG CCA</td>
</tr>
<tr>
<td>A2p-F-Pme1</td>
<td>pme1 (GTTCATAC)</td>
<td>GGC GTT TAA ACA GAC CGC CAG CCG CCA GG</td>
</tr>
<tr>
<td>A2p-R-Kpn1</td>
<td>kpn1 (GGTACCC)</td>
<td>GCG GGT ACC GGC GCC CGC GCC GAT CAC T</td>
</tr>
<tr>
<td>A2cds-F-Asc1</td>
<td>asc1 (GGCGCGGCC)</td>
<td>ATC GGC GCG CCA TGG ACG GGG GCG GCC GA</td>
</tr>
<tr>
<td>A2cds-R-Pac1</td>
<td>pac1 (TTAATTAA)</td>
<td>CAG TTA ATT AAC TAG CAG TTG ATG CCA CAT T</td>
</tr>
</tbody>
</table>
Table 2.3 Calli induction media.

<table>
<thead>
<tr>
<th>GVS media</th>
<th>per Liter</th>
</tr>
</thead>
<tbody>
<tr>
<td>GVS</td>
<td>4.4g</td>
</tr>
<tr>
<td>phytagel</td>
<td>5g</td>
</tr>
<tr>
<td>sucrose</td>
<td>30g</td>
</tr>
<tr>
<td>dicamba</td>
<td>4mg</td>
</tr>
<tr>
<td>copper sulfate</td>
<td>5uM</td>
</tr>
</tbody>
</table>

pH: 5.7-6.0
Figure 2.1 Immature and mature Setaria viridis, cellulose and lignin quantification. Images depict example of plant growth stage: immature, vegetative growth (A) and mature, reproductive stage (B). Bar graph depicts cellulose and lignin (insoluble and soluble) in aerial tissue as well as root samples (C). Significant differences are seen in the areal cellulose content between the two stages, with the younger material having a greater percent of cellulose. The trend is reversed when looking at insoluble lignin with the older material accumulating a greater amount of lignin as compared to the younger tissue. The cellulose and lignin quantifications are the result of three biological replicates and the error bars represent the standard error from the mean. (Scale bar = 10cm) Adapted from Petti et al. (2013)
Figure 2.2 Protein alignment of Arabidopsis CESAl7 (AT5G17420) and Setaria italica CESA “S2” (Si028761). Red amino residues represent high consensus (identical), blue amino residues represent low conservation (weak conservation). Highly conserved residues are in capital letters on the consensus line, while low conservation is depicted as a lower case letter. No conserved residues between the proteins are represented by a dash on the consensus line (Corpet, 1988).
Figure 2.3 Phylogenetic tress of CESA proteins (A) and CESA genes (B).
Phylogenetic relationships were determined for *A. thaliana* (AT), *Setaria italica* (SI), *Z. mays* (ZM), *Sorghum bicolor* (SB) and *P. Vulgare* (PV). Reprinted from Petti et al. (2013)
Figure 2.4 Quantitative real time PCR on select Setaria CESAs. A total of 16 different tissue samples were tested to look at the expression levels of the two CESAs of interest. Bars are the mean of three biological replicates. Error bars indicate standard deviation from the mean.
Figure 2.5 Sample images of the various tissue samples utilized in the qRT-PCR. Scale bar = 10cm in the image of the mature plant.
Figure 2.6 Evidence of loci on calli that were transformed. Transformed loci indicated by the blue staining by GUS. Scale bar= 500μm
Figure 2.7 Gus staining patterns of Arabidopsis plantlets expressing the S2 and A2 promoter GUS fusion. Error bars in A and E=1cm, B and C= 2mm, D and F = 5mm.
CHAPTER 3: SORGHUM CELL WALL COMPOSITIONAL ANALYSIS AND MUTANT CHARACTERIZATION

3.1 Summary

A phylogentic tree to identify the CESAs from *Setaria viridis* was expanded to include an assessment of CESA genes from *Sorghum bicolor*, a crop and important member of the Panicoidae clade. Identification of genes was coupled with a screen of a large mutant population of Sorghum for variation in cell wall composition. A resulting mutant found through the screening was further analyzed.

3.2 Introduction

Despite the many beneficial traits of C4-grasses (Carmo-Silva et al., 2009; Wang et al., 2009), the conversion of cellulosic carbohydrate to liquid transportation biofuel is limited by the recalcitrant influence upon enzymatic hydrolysis of one the most abundant component of the plant cell wall, lignin- a product of the phenylpropanoid pathway (Chandrakant and Bisaria, 1998; Oh et al., 2000; Akin, 2007; Muratov, 2007; Wang et al., 2010; Chu et al., 2011; Van Dyk and Pletschke, 2012). Many advances have occurred in elucidating the genes and the genetic modules involved in this complex pathway and it has been shown that regulatory mechanisms are often mediated through metabolic intermediates (Mavandad et al., 1990; Carpin et al., 2001; Zhong and Ye, 2009; Vanholme et al., 2010) suggesting a dynamic balancing act within the phenylpropanoid pathway.

Here, we examined phylogenetic trees for examples of Sorghum genes involved in cell wall biosynthesis and found that there are 9 putative CESA genes in Sorghum that contain the CxxC and QxxRW motif thus satisfying the criteria to be considered a cellulose synthase gene. In turn, we performed a screen of mutants for modified cell wall properties. We describe the identification of a low lignin mutant arising from diversion of the phenylpropanoid pathway and place this into a molecular genetic context for cell walls.

Although lignin is a vital structural support system for the plant kingdom, allowing for upright growth and water transport through the xylem; as mentioned above, it does not take kindly to enzymatic hydrolysis and has been shown that effectiveness (or lack
thereof) of acid pretreatment and enzymatic digestion is directly correlated to lignin content and thereby being the main factor that limits saccharification yields (Chen and Dixon, 2007b; Van Acker et al., 2013). Fungal and steam explosion pretreatments have also been investigated, however lignin is still recalcitrant to saccharification (Sawada et al., 1995). As lignin saccharification has been a major challenge in the field of cellulosic ethanol and biofuel production, we came across a sorghum mutant that we named REDforGREEN (RG), that seemed like a promising mutant to allow us to further investigate the complex nature of the phenylpropanoid pathway and gain a better insight into the lignin biosynthetic pathway.

3.3 Results

3.3.1 Identification of the REDforGREEN (RG) mutant

The dominant REDforGREEN (RG) mutant was generated through chemical mutagenesis EMS and was identified by its’ strikingly enhanced red pigmentation in plant tissues. We found that the red pigmentation was a result of increased phenolics such as anthocyanins which suggested that there may had been an alteration caused by the EMS mutagenesis somewhere in the phenylpropanoid pathway. Upon closer investigation, the RG mutant displays an antithetic abundance and reduction of lignin in a tissue specific manner. As such, we focused on the physical and chemical properties of the cell wall in RG to determine the suitability of RG for bioconversion as a potential biofuel.

3.3.2 RG mutant displays hyper-accumulation of pigments and reduced plant height

The RG mutant was phenotypically identified by a marked accumulation of red/purple pigments in the leaf blades (Figure. 3.1). This trait was characterized as leaf-specific and it evolved in a basipetal manner, from the tip of the leaf to the base and further to the leaf sheath (Figure 3.1). Between 3 and 5 days post-germination RG displayed no distinguishable difference from the sorghum wild type seedling. Onset of pigmentation was evident from the cotyledon and each true leaf lamina thereafter (Figure 3.1). Upon the leaf completing expansion, light-driven pigment accumulation initiated at the leaf apex and progressed to the base, termed basipetal accumulation (Figure 1A–N). Subsequent basipetal progression of red coloration occurred in all lamina during the plant life cycle.
Visual examination indicated that coloration was predominantly interveinal (Figure 3.1B,C). As a general trend, basipetal red/purple coloration of lamina began after cellular expansion had plateaued. The red/purple coloration therefore progressed in a hierarchical manner (with the earliest emerged leaves being fully red, and newly emerged and expanding leaves being green). This phenotype led to a continuum of green terminal leaves and red fully expanded leaves. Red/purple leaf coloration extended partially into the leaf sheath (shown in Figure 3.1L−N). The process of complete red/purple coloration of the lamina took progressively longer (2 days for leaves 1−3 to 21 days for leaves subtending the inflorescence).

To explore the mature RG plant phenotype outside the greenhouse, a field performance trial was established over two growing seasons. RG displayed a reduction in maximum height, compared with wild type. Results showed an average decrement of 37% (P < 0.0001, Mann–Whitney) during the first year and 40% (P < 0.0001, Mann–Whitney) for the second year (Figure 3.2). Reduction in height could arise in association with three phenotypes: 1) reduced internodal length, 2) fewer nodes initiated during the life cycle or 3) both fewer internodes of a shorter length. A large internodal length decrease was calculated for the RG mutant as compared to the wild type (Figure 3.3 A and E, P < 0.0001). Notably, the number of nodes was significantly greater in RG versus wild type (Figure 3.3 D, E; P < 0.0001) but this could not compensate for the shorter internode length for RG, resulting in shorter plants. Additionally, the frequency distributions of internode lengths between the two genotypes indicated that the RG mutant is skewed toward shorter internodes relative to a more normal distribution of overall longer lengths in WT (Figure 3.3 B and C). Leaf length and width was also estimated but was not significantly different (Data not shown). Further, it was found that the average seed weight (1000 seeds weight, with replication) was not different between RG and the wild type which was consistent with no visible changes in the RG inflorescence.

3.3.3 Cell wall composition in the RG mutant reveals modified lignification

As mentioned above, the RG mutant displays visible red coloration of leaves that develops basipetally (Figure 3.1). The red/purple accumulation in RG leaves was consistent with an alteration in the phenylpropanoid pathway. Gene expression and chemical compositional analyses were performed which indicated an increase in transcript
abundance for key genes involved in the lignin biosynthetic branch of the phenylpropanoid pathway in RG red leaves compared with wild type (Figure 3.4 A). The converse was true for gene expression in the stems for all but cinnamyl alcohol dehydrogenase (Figure 3.4 B). Consistent with these data, transcript abundance extended from those genes whose products are responsible for the earliest committed metabolic conversions in the phenylpropanoid pathway (data not shown). To establish whether lignin biosynthesis was increased, we visually observed lignin by counterstaining with phloroglucinol stain (Maule’s reagent, (Bate et al., 1994)). Here, we examined transverse cross sections of RG and wild type stems and leaves. Results illustrated a more pronounced lignin staining in the wild type stem section relative to RG (Figure 3.5 A,C) and vice versa in the leaf (Figure 3.5 B,D). Therefore, histochemical data suggest that stems of the RG mutant have reduced lignin biosynthesis whereas the leaves display increased levels. Transcriptional characterization of 4-week old RG and WT stems demonstrated a down-regulation for key genes involved in the lignin pathway for RG compared to the wild type, corroborating the histochemical evidence for lignin reduction (Figure 3.4 B). One exception was CAD, whose expression was up-regulated, contradictory to reduced lignification of this tissue.

The content of (soluble and insoluble) lignin in leaves and stems of field-grown RG and wild type plants was determined. Both forms of lignin were increased significantly in the leaf tissue of RG compared with wild type (Figure 3.5 E,F). By contrast, acid insoluble lignin content decreased significantly in the stem of RG compared with wild type. Acid soluble lignin, which accounts for a small proportion (2-3%) of the total lignin, was unchanged in the RG and wild type stems. Results showed that the acid insoluble lignin content of the RG leaf was similar to that of the wild type stem (Figure 3.5 E, P > 0.05). Taken together, these results demonstrate that lignin accumulated in an antithetic pattern in RG.

3.3.4 Saccharification of total biomass reveals RG mutant influences digestibility

Based on the modified lignin content of the RG mutant, we sought to determine whether the lignocellulosic biomass displayed a different response to saccharification compared with wild type. It was anticipated that increased lignification in the leaves would influence saccharification efficiency. Indeed, it was found that WT leaves were more efficiently converted to fermentable sugars than RG leaves (data not shown). These data
are consistent with the increased lignin content of RG leaves reducing digestibility. In the stems of the RG mutant, where lower insoluble lignin was quantified (Figure 3.5 E), we observed that the RG mutant displayed higher saccharification efficiency than WT. This too supports lignin content having an influence on digestibility. Secondly, these data are consistent with the RG mutant phenotype being a suitable marker for digestibility traits in a tissue specific manner.

As the digestibility of the various tissues (stem vs leaves) were varied, we also looked at the biomass crystallinity which examines the crystalline signature relative to other amorphous biomass components using x ray diffraction. The analysis showed that there was no significant difference in relative crystallinity of the samples. To determine if cellulose structure was influencing saccarification efficiency, enzymatic conversion of semi purified cellulose was conducted. However, both the derived Vmax and Km were very similar between that of the RG and wildtype (data not shown).

The data collectively showed that improved cellulose digestibility in the RG mutant stem, relative to WT, correlated best with aberrant lignification as opposed to cellulose abundance or structural changes.

3.3.5 Hemicellulose but not cellulose are altered in the RG mutant

Lignin has been suggested to chemically associate with hemicellulose in the cell wall (Scheller and Ulvskov, 2010), particularly xylan side groups have an important role in the bonding of lignin to hemicellulose e.g. ester linkages between lignin and methylglucuronic acid residues and ether bonds from lignin to arabinosyl groups have been reported. Moreover, phenolic lignin components, such as ferulic acid and p-coumaric acid, are covalently bound to hemicelluloses. Therefore, we hypothesized that further rearrangements in the plant cell wall may exist in the RG plants. To further explore, we analyzed cellulose and neutral sugar composition in the leaves and stems of both genotypes. No significant difference in the percentage of cellulose was determined between the RG and wild type in either tissue type (Figure 3.6 A, P > 0.05). Further, when we examined the distribution and localization of crystalline cellulose by calcofluor white staining in stems of the RG and wild type, we observed similar fluorescence intensity in transverse sections (Figure 3.6 B). Neutral sugars, which contribute to the hemicellulosic
fraction of the cell wall displayed variation between the RG mutant and wild type. Fucose and rhamnose, which in general account for small portions of the total sugars (Hatfield et al., 1999), were unchanged in the stems. In leaves, rhamnose was significantly greater in RG (Figure 3.6 C). Also in leaves, arabinose, galactose, and glucose were significantly more abundant in RG than wild type (Figure 3.6 D, P < 0.05). By contrast, leaf xylose decreased in RG from 26% in wild type to 19%. The stem composition also displayed differences. Here, galactose was decreased significantly and glucose was increased in RG (Figure 3.6 D, P < 0.05). Therefore, neutral cell wall polysaccharides were aberrant in RG accompanying antithetic changes in lignification, consistent with association between the fractions.

3.4 Discussion

In studying and analyzing the Sorghum mutant, we were able to have a more thorough and in depth look at the cell wall composition and biosynthetic mechanisms involved than was possible to with Setaria. In investigating the RG mutant in detail, while we originally were trying to gain a broader view of the cellulose biosynthesis machinery, instead we were able to gain a better understanding of more toward the lignin biosynthetic pathways. In particular, the phenylpropanoid pathway and its regulatory effects were able to be studied in more depth with the use of the RG mutant giving us a greater insight into the complex nature and the not so straight forward regulations occurring in the cell wall.

The Sorghum RG phenotype identified represents a prominent marker for redistribution of lignin (and broadly phenylpropanoid metabolism) and improved stem saccharification efficiency for conversion of lignocellulosic biomass to fermentable sugars (Figure 3.5). Despite evidence that lignin content and composition were altered, as was the hemicellulose fraction, no data supported a change in cellulose biosynthesis in the RG mutant when the proportion of cellulose as expressed per gram of cell wall was quantified. XRD as well as pseudo-kinetics of semi-purified cellulose saccharification failed to support a structural alteration of the cellulose microfibrils. Therefore, it seems most plausible that the alterations in saccharification arose as a consequence of the abundance and composition of lignin in the cell wall.
Indeed, there had been a lignin modification in the mutant that cause an increase in lignin in \textit{RG} leaves and the opposite, a lesser amount of lignification in the stems, when compared with wild type, this improved saccharification efficiency in the \textit{RG} stem (Figure 3.5). These results correlated with an irregular distribution of insoluble lignin in the above mentioned tissues and support lignin’s inhibitory role on enzymatic processes and lignin content having the most pronounced effect on saccharification efficiency (Sutcliffe and Saddler, 1986; Draude et al., 2001; Chen and Dixon, 2007b; Papa et al., 2012).

In addition, we were also able to extrapolate our findings to a biofuel production application. Specifically, \textit{RG} leaf samples displayed approximately 30\% lower efficiency and \textit{RG} stem was approximately 2.5-fold higher saccharification efficiency than wild type. Changes in S:G lignin ratio can influence degradation properties, in terms of hydrolytic enzymes creating fermentable sugars from lignocellulosic feedstocks (Davison et al., 2006). Thus, the metabolic re-writing evident in \textit{RG} may represent a strategy to modulate lignin structure. Lignin and other phenolic plant metabolites have typically been viewed as a waste stream in a biorefineries, which instead focus on polysaccharides as feedstock inputs (Simmons et al., 2010). In the case of the \textit{RG} mutant, the leaf biomass would need to be removed and separated from stem material either during or post-harvest in order to optimize biorefining costs.

The removal and separation of the stem material may not necessarily be a favorable attribute; however, \textit{RG} leaf tissues have an alternatively encouraging trait which is the increased micronutrients within them. Coupled with increased phenolics such as anthocyanins, the \textit{RG} leaf tissue may provide a feedstock for the development of nutritional supplements for livestock (Prior and Cao, 2000). In looking at the current literature, high value by-products of lignocellulosic crops developed during bioprocessing has been under-explored. Often times, the lignosulfonates that are a byproduct of the paper industry can be used as the base to produce higher value products using the lignosulfate as the substrate to produce products such as vanillin, pesticides, industrial cleaners and dyes (Doherty et al., 2011). Indeed, in looking at the GC/MS data from the \textit{RG} analysis- it demonstrated that the increased lignin and phenolic components of the \textit{RG} leaf tissue released a differential spectrum of high value pyrolysates including vanillin, a structural compound related to vanillin, namely the 4-hydroxy-3- methoxyacetophenone (Apocynin) and 4-hydroxybenzaldehyde; these compounds have been investigated for
their pharmaceutical potentials and effect on human health (Ha et al., 2000; Castro et al., 2012).

Regulation of the PPP is complex, and numerous levels of feedback inhibition and transcriptional regulation exist. For example, intermediates of the pathway negatively regulate gene expression and thus co-regulate the accumulation of precursor or the next metabolite (Mavandad et al., 1990; Carpin et al., 2001) and as discussed in (Gray et al., 2012). In the stems of RG, down-regulation of the lignin specific PPP genes was generally observed (Figure 3.4). Taken together, the RG mutant leaf biomass has potential to be implemented as a dual use crop.

Lignin alteration in RG was also qualitative. A change in the S:G ratio was measured in RG compared with wild type. Since deposition of lignin types (H, G and S) is a spatial-temporal process (Donaldson, 2001; Boerjan et al., 2003), the quantitative/qualitative changes could indicate that the branches of the lignin pathways are differentially regulated leading to a differential deposition of say S versus G. It is tempting to speculate that bypassing the regulatory networks responsible for physiological and developmental control may be associated with a dominant negative mutation (Veitia, 2007). Similarly, a mutation in the promoter region of transcription factors has been associated with the activation of the PPP. For instance, multiple repeats within a MYB10 transcription factor in apples (Malus domestica L.) have been coupled with an auto-regulation and a dramatic all-red phenotype (Espley et al., 2009). However, evidences of single gene auto-regulation are also available in literature for instance the case for the sorghum bmr2 mutant and a consistent down-regulation of lignin in both tissue types (Saballos et al., 2012).

The RG mutant displayed an overall reduction in the total plant height in the mature plant to around 60-70% of wild type (Figure 3.2). Several plausible explanations exist to explain this observation. Firstly, the reduction in size was consistent with the correlation between lignin content and biomass highlighted in various studies (Pedersen et al., 2005; Novaes et al., 2010). We hypothesize that the reduced lignin results in reduced water transport capacity and suppressed height potential.
3.5 Materials and methods

3.5.1 Lignin content determination

Acid-soluble lignin, acid-insoluble lignin and ash were measured using the laboratory analytical protocols NREL, LAP-004 (1996) and lignin distribution was visualized by phloroglucinol staining. For lignin determination, the first internodal region of field grown plants was used as well as leaf material, and 4 biological replicates (each containing 4 technical replicates) of 300 ± 1.0 mg were employed. The localization of lignin was determined histochemically on stems of axillary secondary shoots of the main mutant/WT, stem as well as on 4–6 weeks old plantlets, using a saturated solution of phloroglucinol in 20% HCl. All chemicals were reagent grade (Sigma-Aldrich, St. Louis, MO), unless noted otherwise.

3.5.2 Neutral sugar analysis

For neutral sugar analysis, a protocol described in (Harris et al., 2009) was followed. Ribose was used as an internal sugar standard and authentic standards, used for all sugars, identified ribose and fucose, rhamnose, arabinose, galactose, glucose, mannose, xylose, galacturonic- and glucuronic-acid. The HPLC profile used was described previously (Mendu et al., 2011a). Measurements comprised four biological replicates as well as four technical replicates.

3.5.3 Cellulose estimation

Cellulose content was estimated as acid soluble glucose (Updegraff, 1969). A crude cell wall extract was prepared according to (Reiter et al., 1993). Cellulose estimation was completed on 5 mg of field grown plant material and used four biological and technical replicates.

3.5.4 Micro-scale saccharification of biomass

Saccharification efficiency was measured for leaf and stem material from the RG mutant and wild type. Briefly, total biomass was collected from field grown samples.
Samples were dried at 50°C for seven days prior to homogenization using a grinder (Arthur H Thomas Co Scientific, Philadelphia, PA, USA) equipped with a 1.0 mm sieve. A cocktail of Cellulase from Trichoderma reesei (Sigma) and Cellobiase from Aspergillus niger (Sigma, USA), equivalent to 2 Filter Paper Activity Units (FPU), was used to digest 100 mg of raw biomass (ethanol and acetone washed) in a 50 mM citrate buffer (pH 4.8) in a final volume of 2 ml at 50°C. A time course deconstruction assay was performed over a period of 48 hours (at 0, 4, 12, 24 and 48 hours) whereby 3 100 µl aliquots (without replacement) were collected for glucose analysis. Quantification used the Yellow Springs Instruments (YSI)-glucose analyzer standardized for glucose determination using YSI buffer and membranes purchased from YSI (Yellow Springs, OH, USA). Glucose release was converted into saccharification efficiency and expressed as a percentage of cellulose within the biomass convertible to glucose.

3.5.5 Semi-purified cellulose digestion

Cellulase digestion and pseudo apparent kinetic analysis of semi-purified cellulose from both the RG mutant and wild type tissue was performed as described previously (Harris et al., 2012a). Semi-purified cellulose (2, 5, 7.5, 10, 12.5, 15, 17.5 and 20 mg) obtained from leaf and stem material from RG and WT was digested in the presence of equal amounts of Cellulase and Cellobiase, equivalent to 2 FPU, in a 50 mM citrate buffer (pH 4.8) in a final volume of 2 ml at 50°C for 2 hours. The enzyme mixture was heat inactivated (100°C, 3 min) prior to glucose measurement using the YSI-glucose analyzer (Yellow Springs, OH, USA) (as described above). These data were generated in triplicate for average ± standard error and the glucose values converted to mol min⁻¹ unit protein⁻¹ and used to determine the apparent kinetics values using the program GRAPHPAD PRISM-4 (Graphpad, La Jolla, CA, USA). The inability to exactly calculate the number of catalytic ends in the complex mixture of cell wall biomass allowed only for the calculation of a relative estimation, expressed as relative kinetics (relative Vmax and Km).

3.5.6 X ray diffraction analysis

X-ray diffraction was performed as described previously (Harris and DeBolt, 2008). Diffractograms were collected between 5° and 35° (for samples with little baseline drift), with 0.02° resolution and a 2 s exposure time interval for each step.
Copyright statement: Portions of this chapter were published in the Journal of Agricultural Food Chemistry: Petti et al. (2014). As well as in Biotechnology for biofuels: Petti et al. (2013a).
3.6 Figures
Figure 3.1 *RG* phenotype development (basipetal progression of pigment accumulation in *RG* leaves) (A, D) full-length leaves of *RG* partially and fully pigmented; (G) wild type leaf with no pigment (scale bar = 2cm); (B, E, H) detailed image of the abaxial leaf surface with evidence of transition from red to green (B), completed accumulation and no green tissue (E), and all green tissue in WT (H); (C, F, I) detailed view of the adaxial leaf surface; (L, M, N) leaf sheath displaying pigment accumulation for *RG* but not for the wild type (scale bar = 1cm); (A) 2-week-old plantlet with evidence of pigment accumulation (scale bar = 1cm); (B) comparison and progression of pigment accumulation on cotyledon, first and second leaf of *RG*, and WT (scale bar = 1cm). Reprinted with permission from Petti et al. (2014). Copyright 2014 American Chemical Society.
Figure 3.2 Plant height of RG and wild type for two field seasons. For each growing season, plant height was determined based on maximal flag termini (n = 25 replicates). Error bars represent the standard error from the mean. Significance (P < 0.05) is indicated by a star (*). Adapted from Petti et al. (2013a)
Figure 3.3 Internode and node assessment. The length of the internode was compared between RG and wild type (n = 25) (A). Histograms display the distribution of internode lengths for each genotype (B, RG; C, WT). The average number of nodes per stems (D). Error bars in all figures are standard error from the mean. Significance (P < 0.05) is indicated by a star (*). Visually, the node and internodes are depicted in (E). Scale bar = 3.5cm. Reprinted from Petti et al. (2013a)
Figure 3.4 Gene expression analysis for key genes in the lignin biosynthetic pathway. (A) Transcriptional analysis for leaf tissues; (B) Transcriptional analysis for stem material at 4-week old for RG and wild type. HCT, Hydroxycinnamoyl transferase; CAD, Cinnamyl alcohol dehydrogenase; COMT, Caffeic acid O-methyltransferase; CCoAOMT, Caffeoyl-CoA O-methyltransferase; CCR, Cinnamoyl-CoA reductase, C3H1, 4-coumaric acid 3′-hydroxylase 1. Significance (P < 0.05) is indicated by a star (*). Reprinted from Petti et al. (2013a).
Figure 3.5 Lignin quantification from the leaf and stem of RG and wild type sorghum. Maule's staining for lignin in cross sections of sorghum stem (A, C) and leaf tissue (B, D). (E) Total insoluble lignin, (F) total soluble lignin; Each bar comprises the mean of four biological and four technical replicates. Error bars indicate the standard error from the mean. Significance (P < 0.05) is indicated by a star (*). Scale bar = 1mm. Reprinted from Petti et al. (2013a).
Figure 3.6 Cellulose and hemicellulose quantification. (A) Cellulose content from leaf and stems of RG and wild type sorghum (n = 4 biological and technical replicates, error bars are standard error from the mean). B) Histochemical analyses of cellulose deposition by Calcofluor White (Sigma Aldrich, St Louis, MO) staining of transverse sections from noted genetic background, scale bar = 1mm. (C) and (D) Neutral sugars composition for RG and wild type leaf and stem biomass. (n = 4 biological and technical replicates, error bars are standard error from the mean). Significance (P < 0.05) is indicated by a star sign (*). Reprinted from Petti et al. (2013a).
CHAPTER 4: CHARACTERIZATION AND ANALYSIS OF THE HETEROLOGOUS EXPRESSION MUTANT HULK

4.1 Summary

This chapter describes experiments to heterologously express a Setaria ortholog into Arabidopsis. The selected Setaria gene SvCESAS2 was selected based on expression and phylogenetic similarity to secondary cell wall CESAs characterized in Arabidopsis. Transformed individuals were identified and studied. SvCESAS2 was driven by the constitutive cauliflower mosaic virus 35S promoter. The resulting selected line was named HULK due to extended period in vegetative growth as compared to wild type and larger than normal plant size. Even once it had started to bolt, the main stem was thicker than wild type. It was found that the phenotypes were not due to expression of the CESA allele based upon results indicating that independent lines with the same construct did not show the same phenotypes. Evidence is presented that phenotypes are due to the insertion of the Setaria CESA as a T-DNA. Currently, data indicate that HULK is an ethylene insensitive mutant and attempts were made to locate the position of the insertional event.

4.2 Introduction

The Arabidopsis mutant, named HULK came out of the Seteria CESA homologous expression study that had been carried out in Arabidopsis. Often times, heterologous expression systems are used to produce economically valuable compounds in yeast and fungi (Chiang et al., 2013) as well as to improve crop yields as shown by the expression of the Arabidopsis phytochrome B in potatoes (Thiele et al., 1999). Heterologous expression has also been used in the field of biofuel production to create transgenic plants that produce heterologously expressed cellulases for biomass that is more readily digestible for biofuel production platforms (Dai et al., 2005; Hood et al., 2007; Lambertz et al., 2014). In addition, heterologous studies also allow for the study of specific genes of interest in a non-native system which is especially helpful for plants in which genetic modification is difficult, such as Setaria. Native genes as well as heterologously expressed transgenes, with the intent of being overexpressed have been studied and documented to be found that they can be silenced transcriptionally or post transcriptionally (Napoli et al.,
Our end goal was that we would be better able to study the *Setaria CESAs* in a heterologous system without consequence.

There were a few questions that we hoped could be answered by the heterologous overexpression studies, such as would a single CESA from *Setaria* be able to function in *Arabidopsis*? It is well documented that cellulose biosynthesis is by a rosette structure that requires three different CESAs to complete. Would a lone *Setaria* CESA be able to incorporate itself into the CSC and rosette structure? Even without being able to be incorporated into the native cellulose biosynthesis rosette, would simply the overexpression of the *Setaria* CESA allow for an increase in cellulose production?

As a result of the transformations, there was one line that had quite the striking phenotype. The robustness and significantly larger phenotype of the line led to us referring to the transgenic line as “HULK”. Initially we had hoped that the overexpression of the *Setaria CESA* had caused the alteration in phenotype. However, upon creation of more independent lines, we quickly realized that we were looking at an insertional mutant effect caused by a disruption from our construct to a native gene and from looking at the other independent lines, there appeared to be no negative effect on the plants by an overexpression of a heterologous *CESA*, however there were also no striking differences.

Due to the interesting phenotypes that *HULK* displayed, we set off to document and characterize the various striking features of the mutant and also try to find exactly which gene within the genome had been disrupted to cause such a striking set of phenotypes.

### 4.3 Results

#### 4.3.1 Phenotypic analysis

The *HULK* mutant continuously showed a distinct difference in their development and phenotype as compared to their wildtype counterparts. The first indication of a deviation from the norm is the alteration in light sensitivity of the mutant seeds as seen by the delayed germination of the mutants in dark grown conditions. In order to induce the growth of etiolated hypocotyls, *HULK* and wildtype seeds were plated on media and
wrapped in aluminum foil and placed in the growth chamber. Three days later, the plates were unwrapped to find that while the wildtype seeds had produced etiolated hypocotyls, the majority of the mutant seeds had not yet even germinated. Placing the seeds under the light for three to four hours after they were plated broke the dormancy and then were able to germinate under dark conditions.

The next clear difference in development was the in the growth and number of rosette leaves of the mutant. As they matured, the rosettes of the Hulk and continued to prolifically produce a significantly greater number of rosette leaves as compared to their wildtype counterpart (Figure 4.1). We hypothesize that this also led to the other characteristic of the HULK, delayed bolting as more resources were diverted into vegetative growth- to produce a larger rosette with larger and more number of leaves, and as such a significantly greater above ground biomass, that this delayed the transition into the reproductive stage for the plant (bolting and flowering). When the mutant did start to bolt (approximately 2 weeks later than wildtype) and continue on, we observed that the main stem was much thicker in addition to being significantly taller than its wild type counterpart (Figure 4.2).

In looking at the seeds of the mutant, it was found that the HULK seeds were significantly smaller than wildtype (3.47 ± 0.15 mg, 4.97 ± 0.05 mg) To see whether this could be a seed coat defect (potentially a seed coat mucilage defect), the seeds were imbibed and shaken in the ruthenium red stain. After imbibing and shaking the seeds in the stain, Hulk displays no abnormality in mucilage production or mucilage adhesion to the seed surface.

4.3.2 Cellulose analysis

It was our hypothesis that HULK phenotypes were a result of an overexpression of a CESA gene, which in turn led to bigger tissues. Alternatively, if an insertional mutation arose from the CESA T-DNA, it was expected that there would be no change in cellulose content. In order to investigate, we isolated part of the HULK and wild type plants including, stem internodes and leaf tissue and then measured cellulose content by quantifying the acid soluble glucose (Figure 4.3 A). Results showed that there was no significant difference (P > 0.05, Student T Test) in the cellulose quantities between the mutant and
wildtype. Thereby indicating that the phenotype was not due to the overexpression of the CESA construct nor was it producing any additional cellulose.

4.3.3 Lignin analysis

As the cellulose analysis showed no significant difference in quantity between the wild type and mutant, we also decided to investigate the lignin content (Figure 4.3 B). As described previously, the HULK’s main stem was much thicker than that of the wild type (data not shown)- could it be that there was a greater degree of lignification in the mutant to support the taller plant, thus causing a thicker stem?

As the base of the stem (closest to the rosette) is naturally the most lignified area, we decided to sample a five centimeter section of the main stem from the base of both the wild type and mutant plant to quantify the lignin content. In doing so, there was an appreciable difference in the amount of biomass from the wildtype as compared to the mutant (4.3±0.67 mg vs 15±1.56 mg) indicating that indeed the stem of the HULK were much thicker than the control. However, in quantifying the amount of lignin, we found that there was no significant difference in the amount of lignin between the wild type and the mutant.

4.3.4 ACC test for ethylene sensitivity

From the cell wall analysis, there appeared to be no alteration in the main components of the cell wall, we considered if the HULK may have an alteration in its ability to sense and respond to ethylene, a major plant hormone. Ethylene is most often associated as the plant hormone responsible fruit ripening but also plays vital roles in seed germination as well as stress responses (Lin et al., 2009). The triple response phenotype exhibited by etiolated seedlings is one of the most common screens utilized to look for ethylene mutants (Bleeker et al., 1988; Guzman, 1990; Lin et al., 2009). The triple response of wild type seedlings in the presence of ethylene includes the inhibition of hypocotyl/ root growth, the radial swelling of the hypocotyl and root as well as an exaggerated apical hook upon germination(Guzman, 1990; Gagne et al., 2004; Lin et al., 2009).
In order to investigate if the mutant could be an ethylene mutant, the seeds were grown on media containing with varying ACC (1-amino- cyclopropane-1-carboxylic acid) concentrations. ACC is an ethylene precursor commonly used to test for the triple response. After five days on the media with the corresponding ACC concentration, the length of the dark grown etiolated hypocotyls were measured (Figure 4.4 A/B). Even from the lowest ACC concentration (5μM) tested there was a drastic difference in the hypocotyl length between the wild type and the HULK. At the highest concentration (50μM), the difference was even more pronounced in how severely the wild type hypocotyls were stunted as compared to the mutant. In the control plates with no ACC, there were no significant difference in the hypocotyl length between the mutant and wild type (data not shown). This ability of the HULK seedlings to grow in the presence of ACC indicates an insensitivity of the HULK mutants to respond to the compound.

4.3.5 Insertional position (genome walking)

The insensitivity of HULK to the ethylene precursor demonstrated that the mutant is not a cell wall mutant but rather a mutant in the ethylene signaling pathway. Although it was not the outcome that we had hoped for, we decided to attempt to determine exactly what gene has been disrupted in the HULK; therefore, a genome walking experiment was designed to try and pinpoint the exact location of the disruption.

As the inserted fragment of DNA that had caused the mutation is known, the known sequence was utilized to try to sequence the flanking region within the mutant to find the interrupted gene using a technique called Restriction Site Extension PCR RSE-PCR (Ji and Braam, 2010). Briefly, genomic DNA was extracted from the HULK which was then digested individually with the restriction enzymes SpeI, NcoI, NheI and MluI- enzymes which were verified to not cut within the inserted fragment (SvCESAS2 gene). The digested DNA was then used as the template for two back to back PCRs in which a primers targeted the ends of the randomly cut genomic DNA (utilizing the restriction sites) while the other primer was designed to anneal to the known inserted sequence of DNA (SvCESAS2). This in theory should allow for the genomic sequence that was flanked by the restriction site and inserted DNA to be sequenced and the gene interrupted identified.
The experiment was repeated multiple times starting with the extraction of new Hulk DNA and new enzyme digestions each time. Of the four different enzymes utilized, Ncol and MluI produced specific bands that could be used for sequencing. The PCR products were submitted for sequencing, but were not able to be sequenced, therefore the PCR products were cloned into the pGEM T Easy vector (Promega) and sent for sequencing. Although the sequencing itself was successful, the results were inconclusive as to where the insertion had occurred.

4.4 Discussion

What was initially considered to be the effect of the overexpression of the Setaria CESA turned out to be an insertional disruption of endogenous gene. Although the end result and findings from the HULK were not cell wall related, the Hulk revealed a novel insertional mutant in the ACC ethylene pathway.

Addressing our initial question, because the HULK was not a result of the heterologous overexpression and because the other independent lines transformed with the same construct had no alteration in phenotype, it also showed that the introduction of a overexpressed cellulose synthase gene had no effect in a heterologous system. These data were consistent with prior experiments where the overexpression of a native cellulose synthase driven by a native promoter fails to induce changes in cellulose biosynthesis. Similarly, many groups have tried to create plants that produce a higher amount of cellulose with little success. Many times it has been though introducing and overexpressing a CESA to see whether it is able to incorporate itself into the well documented rosette of the cellulose synthase complex. However, there has been very little evidence to show that this is possible. In some cases, there have been reports that the overexpression of a native CESA has caused an adverse effect, which resulted in the suppression of the endogenous CESA, thereby causing a reduction in cellulose, quite opposite to the original intentions. Overall, the HULK is significantly larger than wild type. Ethylene signaling appears to be involved.
4.5 Materials and methods

4.5.1 Cellulose analysis

The cellulose content was measured by quantifying the acid soluble glucose as described in Updegraff et al (1969) by utilizing 5mg samples of stem tissue from mature plants.

4.5.2 Lignin analysis

The lignin content was quantified using the acetyl bromide method and by using the bottom 5cm of the wild type and HULK main stem. The extinction coefficient of 23.35 g\(^{-1}\) L cm\(^{-1}\) was used and the absorbance read at 280nm (Xue et al., 2008; Ralph et al., 2015)

4.5.3 ACC testing

A range of concentrations of ACC (from 5μM to 50μM) were incorporated into solid media on which the seeds were allowed to germinate as described by Gagne et al (2004). The seeds were germinated on their designated plates for five days in the dark, after which their hypocotyl lengths were documented.

4.5.4 Identification of insertional position

The known sequence was utilized to sequence the flanking region within the mutant to find the interrupted gene utilizing the methods utilizing the RSE-PCR as described by Ji et al. Briefly, genomic DNA was extracted from the HULK and digested with Ncol, SpeI, Mlul and Nhel independently. The digested DNA was then used as the template for two PCRs. The first PCR utilized a primer that annealed to the digested ends of the DNA, while also adding an adaptor sequence to be used for the second PCR and the other primer was designed to anneal to the S2 gene that had been inserted. The second PCR further amplified the products from the first PCR in which one primer annealed to the adaptor sequence that had been added in the first PCR and the other primer was designed to be a nested primer within the S2 sequence. This in theory should
have allowed for the genomic sequence that was flanked by the restriction site and inserted DNA to be sequenced and the gene interrupted identified.

The products from the second PCR were sent for sequencing as a PCR product as well as in the pGEM-T Easy vector and both the primers utilized in the PCRs described above as well as the M13 forward and reverse primers were utilized for sequencing.
4.6 Tables and figures

![Bar graph showing the significant difference in rosette diameter between the wild type and mutant. Difference in the size of the rosettes between wild type (left) and the HULK (right). The bars are the mean of three biological replicates and error bars indicate standard deviation from the mean. The HULK continued with a longer vegetative growth stage to produce a larger rosette with a greater number of leaves as compared to the wild type before bolting approximately two weeks after the wild type. Scale bar= 5cm.](image)

**Figure 4.1 Rosette size and diameter.** Bar graph showing the significant difference in rosette diameter between the wild type and mutant. Difference in the size of the rosettes between wild type (left) and the *HULK* (right). The bars are the mean of three biological replicates and error bars indicate standard deviation from the mean. The *HULK* continued with a longer vegetative growth stage to produce a larger rosette with a greater number of leaves as compared to the wild type before bolting approximately two weeks after the wild type. Scale bar= 5cm.
**Figure 4.2 Mature plant phenotype.** Image of mature mutant on left, with wild type comparison on right, scale bar= 10cm. Bar graph depicting significant difference in plant height. The bars represent the mean of three biological replicates and error bars indicate standard deviation from the mean.
Figure 4.3 Cellulose and lignin quantification of HULK. Neither cellulose quantification (A) and lignin quantification (B) in the mutant indicate a significant difference in either cellulose or lignin quantities as compared to the wild type. The bars represent the mean of three biological replicates and the error bars indicate standard deviation from the mean.
Figure 4.4 ACC hypocotyl elongation measurements. Significant difference in dark grown hypocotyl length between the wild type and mutant on two concentrations of ACC 5μM (A), 50μM (B). Wild type hypocotyls were significantly stunted as compared to the HULK, indicating an insensitivity of the mutant to detect ethylene. Bars are the mean of three biological replicates. Error bars indicate standard deviation from the mean.
CHAPTER 5: CSLD1 INTERACTS WITH THE N-TERMINAL ZINC FINGER OF CESA5 IN VITRO AND CONTRIBUTES TO MUCILAGE ADHERENCE AND CELL WALL SYNTHESIS IN THE ARABIDOPSIS SEED COAT

5.1 Summary

The N-terminus of the CESAs contains a family-conserved, cysteine-rich domain referred to as a RING-type zinc finger. These RING fingers have been implicated in mediating a wide variety of protein-protein interactions such as oxidation driven dimerization between CESA subunits. It has been well documented that there are various interactors that work in conjunction with the cellulose synthase such as the CSI1 and CSI3 (Bringmann et al., 2012; Lei et al., 2013). To add to the list, we have found other interactors such as members of the cellulose synthase like family D (CSLD1) and SNRK through a technique known as phage display in which we used the CESA5 RING finger module as bait for interacting proteins via phage displayed cDNA libraries derived from seeds of wild type Arabidopsis. Interaction between CESA5 and CSLD1 was confirmed using Split-Ubiquitin Yeast-Two-Hybrid and bimolecular fluorescence complementation. By reverse genetics, we discovered that seeds from mutated csld1 alleles phenocopied cesa5 seeds and displayed aberrant cellulose ray formation within hydrated mucilage, dysfunctional mucilage adherence and cell wall rearrangements including reduced cellulose content and increased pectin. In vitro biochemical, histochemical and genetic analyses support a role for CSLD1 in cell wall formation in mucilage secretory cells in Arabidopsis seed coats. In the following studies, we have taken a closer look at the role of CSLD as an interactor and what role it may play for cellulose biosynthesis and for an overall better understanding of the cell wall biosynthesis in a specific plant tissue, in this case the seed coat.

5.2 Introduction

Seed coat mucilage is produced in all myxospermous seeds and is thought to play a role in seed germination, storage, dispersal and protection (Western, 2012) and much of the mechanistic studies on mucilage have been performed in Arabidopsis thaliana. Specialized epidermal cells are the sites of mucilage production. Due to the pectin rich nature of the mucilage, when the seeds are hydrated the mucilage rapidly increases in volume and break the outer tangential cell wall to form a halo. The mucilage halo remains
attached to the seed but forms an outer non-adherent layer and an inner adherent layer (Western et al., 2000). Chemically, these inner adherent and outer non-adherent mucilage layers are composed of different amounts of polysaccharides (Dean et al., 2007; Macquet et al., 2007; Rautengarten et al., 2008; Arsovski et al., 2009a; Harpaz-Saad et al., 2011; Mendu et al., 2011a; Sullivan et al., 2011; Saez-Aguayo et al., 2013; Voiniciuc et al., 2013; Yu et al., 2014; Ralet et al., 2016), which in turn are important for the physical properties of mucilage and how it maintains association with the seed during hydration. The inner adherent mucilage layer can be visualized histochemically as a series of rays arising perpendicular to the cell plane on the seed surface (Harpaz-Saad et al., 2011; Mendu et al., 2011a; Sullivan et al., 2011). Rays terminate at the perimeter of the adherent layer. Comprised of pectin, cellulose and mannan (Haughn and Western, 2012; Griffiths et al., 2014; North et al., 2014; Yu et al., 2014; Ralet et al., 2016) the rays are a chemical matrix. Recent studies have suggested that this matrix is reliant on association between xylan chains and pectin (rhamnogalacturonan-I or RG-I) chains, which in turn mediate an adsorption of mucilage to the core cellulosic structures that extend from the host seed (Ralet et al., 2016). Without cellulose biosynthesis, the adherent layer becomes disassociated from the seed upon mild agitation and disperses in aqueous solution (Harpaz-Saad et al., 2011; Mendu et al., 2011a; Sullivan et al., 2011; Griffiths et al., 2015). In addition to complex physiology of this cell type, the seed coat mucilage cells represent a model cell type to examine cell wall biosynthesis, in particular cellulose biosynthetic processes, using molecular genetics.

In *Arabidopsis*, biosynthesis of cellulose in seed coat mucilage cells is part of a two-phase secondary cell wall deposition process (Haughn and Western, 2012). A pectinaceous mucilage pocket is developed followed by synthesis of a reinforced columella (Arsovski et al., 2010; North et al., 2014). Here, cellulose is synthesized in a controlled manner whereby cellulose coils around the columella and unwinds during mucilage hydration to form a linear ray (Griffiths et al., 2015). To date, it has been found that cellulose biosynthesis for the linear rays in seed coat epidermal cells involves particular CESA proteins, including CESA5 (Harpaz-Saad et al., 2011; Mendu et al., 2011a; Sullivan et al., 2011) and CESA3 (Griffiths et al., 2015). It was the goal of the current study to discover additional genetic elements contributing to mucilage adherence and cellulose biosynthesis.
5.3 Results

5.3.1 Development of in vitro screening platform for CESA interacting proteins

To identify interacting proteins with CESA, we focused on our capacity to make phage display libraries from cDNA derived for target plant tissues, in particular a library enriched in seed specific transcripts (Chen et al., 2010; Kushwaha et al., 2012). Previously, we and others had identified that CESA5 was required for cellulose biosynthesis in the Arabidopsis mucilage secretory cells (Mendu et al., 2011a; Sullivan et al., 2011). Therefore, we hypothesized that CESA5 may be a useful bait to screen for interacting proteins. A 262 amino acid fragment of the N-terminus of CESA5 from start codon MNTGGR to KSSKINP 262 was cloned into the recombinant protein expression vector pET29b (Figure 5.1). This fragment contained the full RING-type zinc finger domain and was designated CESA5N for simplicity. Timmers et al. (2009) established that double substitution of the first and last Cysteine (C), corresponding to C39 and C84 with Alanine, in the Zn-finger domain additively reduced the protein interaction more than a single substitution (Timmers et al., 2009). Therefore, using a multi-site directed mutagenesis approach, we substituted C39A and C84A (Figure 5.1, Table 5.1) and designated this fragment mutCESA5N. Before being able to screen our phage display library using these fragments, we needed to establish whether or not these RING-type zinc finger fragments were capable of binding to the microtiter plate wells. Results showed that both mutated and non-mutated CESA5N modules bound to the microtiter plate and their binding capacity depends on the concentration of protein used. Highest binding capacity was observed when the concentration of protein was 10 μg·mL⁻¹ and lowest at 0.01 μg·mL⁻¹ (Figure 5.2). In order to perform the screen, we probed the phage displayed seed proteins against CESA5N and mut-CESA5N. Phage titer was found to increase with successive biopanning rounds with the highest PFU/mL (Plaque forming Units) at 4th round when N-terminus CESA5N was used as bait (Figure 5.3A). By contrast, the control BSA and mutCESA5N revealed a steady decline in phage titer (Figure 5.3A). Thus, BSA and mutCESA5N were incapable of retaining phage demonstrating that the phages were not binding indiscriminately to the N-terminus CESA5 proteins, the blocking agent used, or the well itself (Figure 5.3A).
Over four rounds of biopanning (Figure 5.3B) there was a continuous percentage increase of phage containing inserts from 44 to 77% in CESA5N. By contrast, with BSA and mutCESA5N this percentage declined from 16 to 0% and 27 to 16%, respectively. By the end of four rounds, when the DNA from randomly selected phage containing insert were isolated, amplified, sequenced and compared, about 60% insert were found to be in frame. Four independent hits of CELLULOSE SYNTHASE LIKE - D1 (CSLD1) were found to associate with CESA5N. The interacting region of CSLD1 that exhibited association with the bait CESA5N was a 42-44 amino acid motif (Figure 5.4), which was well conserved among CSLD and CESA proteins.

5.3.2 Yeast two hybrid and BiFC confirms the interaction of CESA5 and CSLD1

To further examine the interaction between CESA5 and CSLD1 we performed in vitro validation using the split-ubiquitin membrane-based yeast two-hybrid (Y2H) system. Results of the Y2H showed that CESA5N interacted with CSLD1 when grown in SD/-Leu-Trp-His media (Figure 5.5), consistent with a biochemical association between the two translated products. A lack of interaction was recorded in either of the controls (Figure 5.5) consistent with the phage display results. In addition, BiFC transient expression was conducted in tobacco (Nicotiana benthamiana) in which the results also confirmed the phage display and yeast two hybrid (Figure 5.6).

5.3.3 Mucilage adherence requires CESA5 and CSLD1 during Arabidopsis seed development

Upon hydration, the mucilage layer in Arabidopsis seed rehydrates and rapidly expands while remaining attached to the seed coat. This creates what is referred to as the mucilage halo (Western et al., 2000). Loss or disruption of the mucilage halo is also amenable to screening by staining and visualizing the process using light microscopy. In order to determine if CSLD1 was needed for adhesion of the mucilage to the seed, the seeds of the wild type (negative control), cesa5 (positive control), and two alleles of T-DNA disrupted csld1 (csld1-1, csld1-2) were imbibed in a solution of Ruthenium Red and screened for visual alteration in the mucilage halo (Figure 5.7). Under the light microscope, all seeds were shown to be capable of producing an intact halo of mucilage surrounding the seed. However, upon agitation (via the use of an orbital shaker) the mucilage layer
surrounding the \textit{csld1-1}, \textit{csld1-2} and \textit{cesa5} almost completely dissociated from the parent seed leaving a naked seed. The wild type mucilage halo remained intact and associated with the seed. Therefore, we are able to identify that CSLD1 associated with the N terminal RING finger domain of CESA5 and genetic dysfunction in \textit{csld1} phenocopied the \textit{cesa5} mucilage adherence phenotype.

In order to quantify the mucilage that was produced and visualized, regardless of association or disassociation to the seed coat, we used the ammonium oxalate method (Arsovski et al., 2009b). Mucilage content of \textit{cesa5} was found to be highest $1.24 \pm 0.087$ mg$\cdot$100 mg seed$^{-1}$ which was almost double that observed in wild type $0.64 \pm 0.0704$ mg$\cdot$100 mg seed$^{-1}$. (P<0.001 Tukey means separation). \textit{csld1-2} ($1.09 \pm 0.17$ mg$\cdot$100 mg seed$^{-1}$) and \textit{csld1-1} ($1.052 \pm 0.077$ mg$\cdot$100 mg seed$^{-1}$) were both found to produce significantly greater amounts of mucilage than the wild type (P<0.001 Tukey means separation) but less than the \textit{cesa5}, this difference was not statistically significant (P>0.05) (Figure 5.7).

5.3.4 Histochemical analysis of \textit{cesa5}, \textit{csld1} and wild type epidermal seed coat revealed irregular morphogenesis

Features at the \textit{cesa5}, \textit{csld1} and wild type epidermal seed coat surface can be visualized and quantified using careful measurements of scanning electron micrographs. Seed cell surface area was calculated using 20 different seeds from each genetic background and 10 cells from each seeds (n = 200). The seed coat cell area was not significantly different between \textit{cesa5} and \textit{csld1}. Between wild type and \textit{cesa5} or \textit{csld1} a significant difference (P < 0.01 Tukey means separation) in cell surface area was measured (Figure 5.8). The cell surface area of wild type seeds was highest 847$\mu$m$^2$ followed by \textit{csld1-1} (791.6 $\mu$m$^2$), \textit{csld1-2} (758.2 $\mu$m$^2$) and \textit{cesa5} (752 $\mu$m$^2$).

It was obvious from visual examination that the regular hexagon shape of the epidermal seed coat cells in wild type plants was distorted in both \textit{cesa5} and \textit{csld1} alleles. One way to document this was to examine the distribution of irregular angles produced (5 cells each of 20 seeds). There was significant difference between the cell angles of all the studied mutants and the wild type. Consistent with their hexagonal shape, wild type seed coat
cells have angles roughly around 120°. Angle distribution varied to a greater degree in cesa5, csld1-2 and csld1-1 (Figure 5.9).

5.3.5 *Radial wall thickness and height are reduced in cesa5 and csld1*

The radial cell wall of seed coat mucilage cells is thickened during cellular development as an important secondary differentiation. The height of this radial wall has been linked to deficit in cellulose biosynthesis (Mendu et al., 2011a; Griffiths et al., 2015). To estimate the height of the radial cell wall we sought to use laser scanning microscopy to create optical cross sections and measure the dimensions of the wild type versus csld1 and cesa5 alleles. In order to visualize the radial wall cell walls were stained with fluorescent dyes calcofluor white or pontamine S4B dyes and allowed to hydrate for 15 min prior to imaging the radial wall. Consistent with prior studies (Mendu et al., 2011a), the cesa5 seed displayed severe reduction in radial wall height. Radial wall height was also significantly reduced in the csld1 alleles compared with wild type (P<0.01 Tukey mean separation) (Figure 5.10). The exhibited reduction in radial wall height in csld1 was less severe than cesa5 (P>0.01 Tukey mean separation).

5.3.6 *cesa5 and csld1 seeds have less cellulose as compared to wild type*

Adherent mucilage contains approximately 12-19% cellulose (Harpaz-Saad et al., 2011; Mendu et al., 2011a). Total cellulose content of the seeds was quantified according to the method of Updegraff (1969). Adherent mucilage in csld1 and cesa5 contained significantly less cellulose as compared to wild type. Wild type seeds have highest cellulose (167.75 ± 2.73mg•g⁻¹ wall dry weight) followed by csld1-1, csld1-2, (142.05 ± 4.6, 139.63 ± 3.6) and cesa5 (110.34 ± 2.1mg•g⁻¹ wall dry weight). Difference between the amount of cellulose in wild type adherent mucilage and that of cesa5 and csld1 was found to be highly significant (P<0.01). Variation in cellulose content of adherent mucilage in cesa5 and csld1 was not significant (P>0.05) (Figure 5.11).
5.3.7 Rhamnogalacturonan I of seed mucilage increased significantly in csld1 as compared to wild type

Seed mucilage is mainly composed of polygalacturonic acid (PGA) and rhamnogalacturonan I (RG I) (Brett and Waldron, 1990; Carpita and Gibeaut, 1993; Cosgrove, 1997). Beside these other polysaccharides present in mucilage are homogalacturonan, arabinans, galactans, xyloglucan, glucomannans, and cellulose (Arsovski et al., 2010; Western, 2012; Yu et al., 2014). To get a better insight into the mucilage composition of wild type versus csld1 seeds we performed monosaccharide analysis of soluble seed mucilage as well as whole seeds by HPLC as described by Mendu et al. (2011). We hypothesized possible feedback changes in wall composition in response to cellulose deficit. Soluble mucilage analysis indicated a significant (P<0.05) increase in the rhamnogalacturonan I backbone sugars i.e. Rha and GalA in the cesa5 and csld1 as compared with the wild type (Table 5.2). On the contrary, a significant (P<0.05) decline in Rha level was observed in whole cesa5 and csld1 seeds compared with the wild type. Gal and Glc level in the mutants were found to be unchanged in both soluble mucilage as well as whole seed compared to wild type. The csld1 exhibited no change in Ara and Xyl content in soluble mucilage compared with wild type but there was a significant increase in Ara levels and decrease in Xyl levels (Table 5.2).

5.3.8 csld1 phenocopied cesa5 in cellulose staining patterns

The reduction in cellulose content in seeds of csld1 was similar cesa5. Thus, we sought to explore csld1, cesa5 and wild type mucilage ray structures in detail using laser scanning confocal microscopy as the ray structures are composed of cellulose (Griffiths et al., 2015). Pontamine S4B shows primary binding affinity to cellulose and, to a lesser degree, xyloglucans while Calcofluor White is less specific, showing binding affinity to cellulose, arabinans, and pectic galactans (Herth and Schnepf, 1980; Anderson et al., 2010). In wild-type seeds, both Calcofluor and Pontamine stained the columella, the remnants of the primary cell wall attached to the columella tips upon mucilage extrusion, and a set of organized rays extending from the tip of the columella at the seed surface and across the adherent layer of seed mucilage (Figure 5.12). Although both stains label the columella and remnants of the primary cell wall in the mutants, the rays appear substantially reduced and malformed in csld1 compared with the wild type. This reduction
in staining intensity was observed most strikingly in *cesa5* but still remained prominent in the *csld1* seeds compared with wild type. These data supported that *CSLD1* and *CESA5* are independently important for mucilage ray formation (Figure 5.12).

**5.3.9 CSLD1 and CESA5 co expresses during seed development**

To get an insight into the relative expression pattern of *CSLD1* and *CESA5* genes during seed development, their relative transcript abundance was studied at different intervals during seed /siliques development. For relative expression analysis, plant samples from wild type plants were taken at 0, 1, 3, 5, 7 days post-anthesis (DPA) and were further processed for RNA isolation. Quantitative RT-PCR showed that there was a significant fold increase in relative abundance of *CESA5* from 1 to 7 DPA. This increase was 6 fold as compared to a 3-fold increase in *CSLD1* transcript abundance over the same period (Figure 5.13). Although their expression level was different, data were consistent with both transcripts being upregulated at a common temporal stage of development in the seed consistent with common functionality.

**5.4 Discussion**

Synthesis and adherence of seed coat mucilage requires the involvement of a complex protein network and a two-step specialized differentiation (Haughn and Western, 2012). According to Fracoz et al. (2015) more than 60 enzymes are required in these processes, and many of these remain unknown (Francoz et al., 2015). Recent finding elegantly illustrate that the adsorption of the mucilage matrix to a cellulose scaffold involves specific Xyl-ramifications on RG-I (Ralet et al., 2016). But biosynthesis of the cellulose scaffold used to extrude mucilage is still being elucidated. To date, data supports the involvement of CESA3 (Griffiths et al., 2015) and CESA5 (Harpaz-Saad et al., 2011; Mendu et al., 2011a; Sullivan et al., 2011) in scaffold cellulose production needed for mucilage adherence. Cellulose formation is broadly considered to take place via the combined action of a cellulose synthase complex (CSC). The CSC uses a 1:1:1 stoichiometry of individual CESA protein subunits (Gonneau et al., 2014). Interestingly, the CSLD family of proteins has been found to share much of the core domain homology with the CESA proteins (Richmond and Somerville, 2000). Exceptions to the above-mentioned CSC was identified in the case of CSLD1, CSLD3 and CSLD4 (Wang et al.,
CSLD3 is capable of (1→4)-β-glucan synthase activity in apical plasma membranes of root hairs consistent with a role in cell wall synthesis in particular cell types (Park et al., 2011). Similarly, CSLD1 and CSLD4 were found to be required for cellulose biosynthesis in pollen tubes (Wang et al., 2011). Based on the reduction in scaffold cellulose rays in the cslD1 mutant mucilage (Fig 8), our results support a role for CSLD1 in either directly or indirectly facilitating cellulose-pectin associations during mucilage attachment.

An indirect mechanism is that a different cell wall polysaccharide, other than cellulose, is synthesized by CSLD1 and this is needed for mucilage adherence. Members of the CSLD family proteins have been found to be required for hetero-mannan (HM) (β-1,4-linked Man units) synthesis. Here, CSLD2, CSLD3 and CSLD5 were suggested to be capable of forming a complex to produce HM (Verhertbruggen et al., 2011; Voiniciuc et al., 2015). This hypothesis is not clear in the case of seed coats. Data observed for both quantitative assessments of cellulose and histochemical assessment of crystalline cellulose by direct red 23 staining both support a requirement for CSLD1 in forming cellulose rays that extend from the seed coat. It is plausible that the feature of polarized secretion in secondary cell wall cellulose biosynthesis observed in seed coats differentiation (Western et al., 2000) is comparable to the polarized secretion mechanisms at play in pollen tubes and root hairs (Wang et al., 2001; Bernal et al., 2008; Park et al., 2011; Wang et al., 2011) that utilize CSLD proteins to synthesize cellulose.

When transcript abundance was carefully measured during siliques/seed development we showed that both CESAl5 and CSLD1 were expressed in the same temporal manner, consistent with pathway association. However, in vitro biochemical association between CSLD1 and CESAl5 was not necessarily indicative of functional association. The CSLD family has five genes and one pseudogene (Bernal et al., 2007) and are closely related to CESA (Richmond and Somerville, 2000). Thus, interaction between CSLD1 and CESAl5 zinc finger domain may be explained by their homology and potential for interaction in vitro such as illustrated by homodimerization between CESA zinc finger domains (Kurek et al., 2002). We therefore do not want to overemphasize the in vitro biochemical data as being relevant in vivo. It remains compelling that the phage display approach was capable of revealing positive interacting proteins that, when examined using molecular genetics, fell into the same pathway. In this case a new
molecular player involved in mucilage adherence was identified. Hence, the involvement of CSLD1 in cell wall biosynthesis during mucilage secretion in Arabidopsis seed coat cells was unexpected and provides a useful new player in the pathway.

5.5 Materials and methods

5.5.1 Phage display

The phage display biopanning, cloning and purification of the CESA5 bait was performed as described in Griffiths et al. (2015). As per Timmers et al. (2009), double substitution in the Zn-finger domain affects the protein interaction more than the one substitution. Therefore, quick change Multisite directed mutagenesis Kit from Stratagene (200514) was used to substitute the Cysteine (C) at position 39 and 84 with Alanine (A) (Fig. S1) by using primer CESA5C39 and CESA5C84, respectively (Table 5.1). (described in Griffiths et al., 2015). Biopanning and titration were carried out by methods described by Chen et al. (2010), Kushwaha et al. (2012) and Griffiths et al. (2015).

5.5.2 Zn-Binding capacity of N terminal CesA5 and Mut- CesA5

Different amounts (40, 80, 100, 200μg) of purified recombinant proteins were incubated with 10μM Zn Acetate for 30min. at room temperature in TSD buffer (50mM Tris-Cl [pH 7.5], 200mM NaCl, 1mM DTT). After 30 min. samples were passed through Sephadex™ G-25 (Amersham Biosciences, Uppsala Sweden) columns to remove unbound divalent cations. For the determination of bound Zn, protein samples were digested with 4 units of Proteinase K (Fischer Scientific, Pittsburg, PA U.S.A) at 60°C overnight. Free Zn ions were quantified colorimetrically according to Hunt et al. (1985). One hundred μM PAR [4-(2-pyridylazo) resorcinol] (Sigma-Aldrich, St. Louis, MO, U.S.A.) was added to 100μl of samples and absorbance was measured at 490nm.

5.5.3 Y2H confirmation of CESA5 and CSLD1 interaction

In vitro interaction between N-terminal CESA5 and CSLD1 was performed using the split-ubiquitin-Yeast-two-hybrid assay based on the Dualmembrane kit
The identified binding region of CSLD1 was cloned into the pGAD-BT2-N-His33 at SpeI and Ncol restriction sites. N-terminal-CESA5 mutated/unmutated were cloned in to pPR-N-RE (NubG) at the restriction sites BamHI and Nhel and were sequenced (Table 5.1). Sequence-confirmed plasmids were co-transformed in Yeast strain NMY51 and plated onto Trp−/Leu− synthetic minimal medium and allowed to grow at 30°C. Transformed colonies were plated (Trp−/Leu−/His) plates to test for CSLD1: CESA5 interaction. Two negative controls were also used (1) empty vectors pGAD-BT2-N-His33: pPR-N-RE and (2) mutCESA5: CSLD1.

5.5.4 BiFC confirmation of CESA5 and CSLD1 interaction

Bimolecular fluorescence complementation was conducted to confirm the results of the Y2H and phage display. The N terminal region of CESA5 and the interacting region of CSLD1 were cloned into the gateway compatible pSITE-BiFC vectors (Martin et al., 2009). Transient expression in Nicotiana benthamiana leaves were carried out by methods described in Schütze et al., 2009. Imaging was performed on an Olympus FV1000 laser scanning confocal microscope using a 40X N.A 0.95 objective. All image processing was performed by using Olympus Fluoview software (Olympus) and ImageJ (W. Rasband, National Institutes of Health, Bethesda, MD) software.

5.5.5 Plant material and growth conditions

All Arabidopsis thaliana mutants used in this study were of the Columbia-0 ecotype. We obtained T-DNA insertion or point mutants from the Arabidopsis Biological Resource Center as follows: CESA5 (cesa5 [SALK_118491]), CSLD1 (csld1-1 [SALK_141530] and csld1-2 [SALK_097300]). T-DNA insertion lines were genotyped and PCR confirmed homozygous alleles were utilized. Seeds were surface sterilized using 30% (v/v) household bleach supplemented with 5% (w/v) sodium dodecyl sulfate (SDS) solution for 20 min. After three sequential washes with sterile water the seeds were moist chilled for 2 days at 4°C in 0.15% (w/v) agar. For germination surface sterilized seeds were exposed to 16hr light (200 mmol• m⁻²•s⁻¹) and 8hr dark cycles at 22°C on plates containing 0.5× Murashige and Skoog mineral salts (Sigma) and 1% w/v agar.

88
5.5.6 Histochemical analysis

Seeds incubated in 0.01% (w/v) ruthenium red (Sigma) in water for 30min at 25°C were used to test for aberrant mucilage production (Beeckman et al., 2000). Alternatively, seeds were shaken on an orbital shaker for 2 hours, then stained with 0.01 % (w/v) ruthenium red in water for 30min at 25°C and imaged with an Olympus MVX-10 Stereo Fluorescence Microscope.

5.5.7 Scanning Electron Microscopy and Image Analysis

*Arabidopsis* seeds were attached to standard electron microscope stubs, and sputter coated with gold-palladium alloy using HummerVI sputtering system (Anatech). Specimens were visualized using Hitachi model S-800 scanning electron microscope (Hitachi), images were captured using Evex Nano Analysis digital imaging system. Image analysis was performed using the ImageJ (W. Rasband, National Institute of Health) software. Measurement for each cell’s planar area used area measurement output after tracing the polygon via the freehand selection tool (ImageJ). Data were organized in the spreadsheet program Microsoft Excel and transferred to GraphPad Prism (Prism-4, GraphPad) for generating histograms and performing statistical comparisons.

5.5.8 Confocal Microscopy

Whole seed were stained to Calcofluor White for seed histochemical analysis. Seed mounts were created between a slide and a cover slip separated by perimeter of vacuum grease that acted to keep seed in contact with Calcofluor White and avoid crushing seed structures. Once mounted, specimens were imaged in darkness, after being exposed to Calcofluor White for 5 min. Imaging was performed on an Olympus FV1000 laser scanning confocal microscope using a 60X N.A 1.4 water-immersion objective. The microscope is equipped with lasers for excitation wavelengths ranging from 405–633 nm; Calcofluor White was visualized at 405 nm. All image processing was performed by using Olympus Fluoview software (Olympus) and ImageJ (W. Rasband, National Institutes of Health, Bethesda, MD) software.
5.5.9 Cellulose Staining

*Pontamine S4B now available as Direct Red 23 [Sigma-Aldrich; 212490-50G]* was used for cellulose staining. Seeds were shaken then stained with *Direct Red 23* as per Mendu et al. (2011).

5.5.10 Mucilage Precipitation

Mucilage content was determined using the ammonium-oxalate method (Arsovski et al., 2009b). Briefly, approximately 50 mg (dry weight) of *Arabidopsis* seed were incubated in 0.2 % (w/v) of aqueous ammonium-oxalate solution at 37°C for 8 hours. To facilitate extraction of mucilage, test tubes were vortexted for 1 min every hour. The seeds were removed by centrifugation after incubation and the resultant supernatant was transferred to a fresh tube. Five volumes of ethanol were added to the supernatant followed by 30 min incubation on ice to precipitate the polysaccharides. Samples were centrifuged in a bench-top centrifuge for 30 min at 21,000 g. Precipitated mucilage was washed with 70% (v/v) ethanol, air dried, and weighed. Average weights were calculated based on three mucilage replicates for mutant and wild type seeds.

5.5.11 Cellulose Content and monosaccharide analysis in Arabidopsis Seeds

Cellulose content was measured colorimetrically (Updegraff, 1969) and neutral sugar composition was determined by high performance liquid chromatography (HPLC) (Blakeney et al., 1983). Briefly, cell wall material from mature seed samples were prepared by sequential washing (five times) with 70% (v/v) ethanol for 45 min at 70°C followed by five sequential acetone washes at room temperature for 2 min each. To remove the starch, the vacuum dried samples were re-dissolved in 1 mL of \( \alpha \)-amylase solution (Sigma Aldrich no. A8220) according to manufacturer's instruction and incubated for 15 min at 37°C. The \( \alpha \)-amylase treated samples were washed and vacuum dried. Exactly 4 mg of the air-dry cell wall material was measured into glass tubes and standards of different neutral sugars (Arabinose, Fucose, Galactose, Glucose, Mannose, Rhamnose and Xylose) were added to individual glass tubes (5.6 μL of 10 mM stock) and dried at 60°C. Thirty-five μL of 72% w/v \( \text{H}_2\text{SO}_4 \) was added to each sample and standard and incubated on ice for 2 hours,
mixing every 30 minutes with the vortex. Samples and standards were next diluted to 4% w/v H₂SO₄ by adding 965 μL water and autoclave at 121°C for 1 hour.

For soluble mucilage samples, approximately 20 mg of seeds (exact weight recorded) were hydrated with 1.2 mL of water and 10 μL of 5 mg•mL⁻¹ D-erythritol was added as an internal standard and the samples were shaken on an orbital shaker for 2 hours. One mL of mucilage was removed and dried at 60°C under a stream of nitrogen. Sulfuric acid (72% w/v, 17.5 μL) was added to each sample, and they were incubated on ice for 2 hours and shaken every 30 min. Water (480 μL) was then added to give a final concentration of 4% (w/v) sulfuric acid, and the samples were autoclaved for 60 min at 121°C and neutralized with 9M NH₄OH before being filtered through 0.45-µm pore nylon syringe filters. Neutral sugar standards (fucose, arabinose, rhamnose, galactose, glucose, mannose, and xylose) and acid sugar standards (galacturonic acid) were also processed in the same way at the same time.

Neutral and acidic cell wall sugars were identified and quantified by Pulsed Electrochemical Detection (PED) using a Dionex ED50. Sugar separation was achieved using a BioLC GS50 HPLC and Carbo PAC-PAI pellicular anion exchange column with column guard (Dionex, Sunnyvale, CA, USA). Column and detector temperature was maintained at 30°C using a Dionex LC25 chromatography oven. Samples (25μL) were introduced to the column from a Dionex AS50 autosampler, and, for neutral sugars, were eluted isocratically in 22mM sodium hydroxide (NaOH) over a 30-min period after which the column was washed in 1M sodium acetate (NaOAc), 200mM NaOH for 5 min. The column was recharged by passing 200mM NaOH through it for 10 min, then re-equilibrating the column in 22mM NaOH for 10 min prior to injection of the next sample. This protocol resulted in optimal (near baseline) resolution of galactose, glucose, mannose and xylose (Downie and Bewley 1996). Acidic sugar separation and quantification used the same HPLC but followed the eluent profile of (Dean, Zheng et al., 2007). For soluble mucilage, monosaccharides were quantified with the D-erythritol internal standard after correction of response factors with monosaccharide standards of different concentrations to allow area to be converted to molar amounts. Sugars from mucilage were normalized to the mass of seed used in the extraction. For cell wall monosaccharides the neutral sugars were normalized to the amount of cell wall material used for the sample preparation.
5.5.12 Relative transcript abundance analysis

Relative quantification of CESA5 and CSLD1 gene transcription products was studied in developing seeds. Flowers/siliques were collected on 0, 1, 3, 5 and 7 days post anthesis (DPA). Total RNA from 100 mg of plant tissue was isolated using QIAGEN RNeasy plant Mini Kit (Cat No. 74904). Applied Biosystems™ (AB) High Capacity cDNA Reverse transcription Kit (Cat No. 4368814) was used for cDNA synthesis following manufacturer’s instructions. Quantitative RT-PCR (qRT-PCR) analyses were performed using the SYBR Green master mix (Applied Biosystems™ SYBR Green® (Cat No. 4312704) with the StepOnePlus™ Real-time PCR System (Applied Biosystems). The PCR program was as follows: 95°C for 10s, 58°C for 35s, repeated for 40 cycles. The relative expression level was normalized to that of the ELONGATION FACTOR 1-A (EF1α) and UBIQUITIN (UBQ11) as mentioned in Gutierrez et al. (2008) as internal controls, with the \(2^{-\Delta CT}\) method representing the relative quantification of gene expression.

5.5.13 Statistical Analysis

All experiments in the study were independently performed in triplicate (unless indicated). Each result shown in the figures was the mean of at least three replicated treatments. The significant differences between treatments were statistically evaluated by standard deviation and one-way analysis of variance (ANOVA). The data between differently treated groups were compared statistically by ANOVA, followed by a Tukey test if the ANOVA result was significant at \(P<0.01\).
5.6 Tables and figures

Table 5.1 List of primers. List of primers used for the cloning of CESA5 N terminus regions, subsequent mutagenesis and yeast two hybrid experiments in addition to primers utilized for the quantitative RT-PCR as well as the genotyping primers utilized to verify homozygosity in the two csld lines. *left border primer for T-DNA insertion from the Salk Institute Genomic Analysis Library.

<table>
<thead>
<tr>
<th>primer name</th>
<th>primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CESA5 NTR F</td>
<td>CCCAAGCTTGGGATGAATACTGGTGGTCGGCTCATCG</td>
</tr>
<tr>
<td>CESA5 NTR R</td>
<td>CCGCTCGAGCGCGGATTTATTTGCTCGACTTTATCG</td>
</tr>
<tr>
<td>CESA5 C-39 F site directed mutation</td>
<td>GAAGTTAATGGAACACACACGTCAATCTGGGAGATGAG</td>
</tr>
<tr>
<td>CESA5 C-39 R site directed mutation</td>
<td>CTGATCTCCCCAGATTTGTGTGTGTCCACTTTAGTTCA</td>
</tr>
<tr>
<td>CESA5 C-84 F site directed mutation</td>
<td>GGGAACAAATCTTTGCTCTCAGGCCAAACTCGTTACAGCGC</td>
</tr>
<tr>
<td>CESA5 C-84 R site directed mutation</td>
<td>GCGCTTTGTAACGAGTTTTCGCTGAGAAGATTTGTTTCC</td>
</tr>
<tr>
<td>CESA5 pPRN-RE F</td>
<td>GCGGAGTCCATGAATCGTGGTGGCGG</td>
</tr>
<tr>
<td>CESA5 pPRN-RE R</td>
<td>CGCGGCTAGCCTAGGATTTATTTGCTCGAC</td>
</tr>
<tr>
<td>CSLD1 pGADBT2N F</td>
<td>GACAGTTCATCACCACACATCACCTATCGTGCCAGCCCTCG</td>
</tr>
<tr>
<td>CSLD1 pGADBT2N R</td>
<td>CGCCATGGCTACTGTTTCATTTCCACCATTCC</td>
</tr>
<tr>
<td>Lb1b.3*</td>
<td>ATTTGGCGATTTCGGAAC</td>
</tr>
<tr>
<td>SALK_097300 LP</td>
<td>GCTTTCCCATTACGAAACG</td>
</tr>
<tr>
<td>SALK_097300 RP</td>
<td>CTCTCCAGCTATACCGCCTTTACC</td>
</tr>
<tr>
<td>SALK_141530 LP</td>
<td>CCATAACACCATTGCATCCC</td>
</tr>
<tr>
<td>SALK_141530 RP</td>
<td>TTCTGGATCCCCCTCTATGTTG</td>
</tr>
<tr>
<td>CESA5F qRT</td>
<td>GGACCTCTATACCTCCTCGG</td>
</tr>
<tr>
<td>CESA5R qRT</td>
<td>CATCTCGAGAATGCCGTTT</td>
</tr>
<tr>
<td>CSLD 1F qRT</td>
<td>CGCGTTTATCTCAATCTG</td>
</tr>
<tr>
<td>CSLD 1R qRT</td>
<td>TGCATATCGCATGGGATGGTCG</td>
</tr>
</tbody>
</table>
Table 5.2 *csld1-1, csld1-2* and *cesa5* exhibit aberrant cell wall polysaccharides compared with wild type. High performance liquid chromatography (HPLC) was employed to measure neutral and acidic sugars derived from whole seed and soluble mucilage. Data reported as the average of 3 replicates for each genetic background, ± S.E (standard error from the mean), nd- not detected.

<table>
<thead>
<tr>
<th>Sugar</th>
<th>Whole Seed</th>
<th>Soluble Mucilage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>wt</strong></td>
<td><strong>csld1-1</strong></td>
</tr>
<tr>
<td>Rha</td>
<td>36.22±1.05</td>
<td>30.31±3.74</td>
</tr>
<tr>
<td>Ara</td>
<td>50.77±0.68</td>
<td>66.40±1.64</td>
</tr>
<tr>
<td>Gal</td>
<td>52.37±1.93</td>
<td>55.56±0.98</td>
</tr>
<tr>
<td>Xyl</td>
<td>26.51±1.53</td>
<td>21.68±0.73</td>
</tr>
<tr>
<td>Glu</td>
<td>11.23±1.05</td>
<td>14.33±1.06</td>
</tr>
<tr>
<td>GalA</td>
<td>nd</td>
<td>nd</td>
</tr>
</tbody>
</table>
Figure 5.1 Amino acid sequence alignment of the N terminus of CESA5 and mutCESA5. Alignment shows the substitution of Cys (C) at position 39 and 84 with Ala (A). Highlighted part corresponds to Zn finger domain. Substitutions with Alanine are shown in red. Adapted from Griffiths et al. (2015).
Figure 5.2 SDS Page gel and attachment to microtiter plates. A 12% SDS-Page gel showing recombinant protein purification for use as biopanning bait (A) CESA5N (1), marker (2) and mutCESA5N (3). ELISA assays showing the bait proteins' attachment to the microtiter plate wells at varying concentration (B). Error bars represent the standard error from the mean. Asterisks indicates significant difference from the BSA based on Tukeys multiple comparison test (P<0.01). Adapted from Griffiths et al. (2015).
Figure 5.3 Phage titer for each biopanning round and gel images of plaques. Phage titer (plaque forming units (PFU/ml⁻¹)) at each of the four rounds of biopanning for the phage display library (A). Asterisks represent significant difference between the titers for the BSA relative to mutCESA5N or CESA5N in each biopanning round. Gel images show the percentage of randomly chosen plaques that contained insert greater than 200bp from the phage display library for each round of biopanning (B). Adapted from Griffiths et al. (2015).
Figure 5.4 Final biopanning hits. After four rounds of biopanning four independent hits 1-4 (shown in black box) that aligned with CSLD1 were discovered. These hits correspond to the sequence of the interacting region of CSLD1 (shown in green box). Scale corresponds to the length of the protein.
Figure 5.5 CESA5 interaction with CSLD1 in split-ubiquitin membrane-based yeast two-hybrid system. NMY51 yeast cells expressing pPR-N-RE CESA5N/ CSLD1 pGADBT2-NHis33 (A) pPR-N-RE MutCESA5N/ CSLD1 pGADBT2-NHis33 (B) and empty vector (pPR-N-RE/ pGADBT2-NHis33) (C) were selected on synthetic growth medium without Leu-Trp-His.
Figure 5.6 BiFC transient expression illustrating interaction of CSLD1 with CESA5 in *Nicotiana benthamiana* leaf epidermal cells. Positive control (A), CYFPc CSLD1+ NYFPn CESA5 (B), CYFPn CSLD1+NYFPc CESA5 (C) and negative control (D). Scale bar: 20µm.
Figure 5.7 Altered seed mucilage in csld1. Mature seed were imbibed in a solution of ruthenium red stain without (A) and with agitation (1.5 hour) (B) and imaged using light microscopy; wild type, csld1-1, csld1-2 and cesa5. Scale bar = 100μm. Extracted mucilage weights were measured for wild type, csld1-1, csld1-2 and cesa5 (C). Error bars represent standard error from the mean (n = 3). Asterisk indicates a significant difference from the wild type based on results of a Tukey’s multiple comparison test (P < 0.01).
Figure 5.8 Assessment of cell shape and morphogenesis of seed coat. Cell area (A) and columella area (B) of wild type, csld1-1, csdl-2, and cesa5 were measured using scanning electron micrographs (n = 200 cells in 10 different seeds). Error bars represents the standard error (SE) of three technical replicates from three independent batches of seeds as biological replicates. Asterisk indicates a significant difference from the wild type based on results of a Tukey’s multiple comparison test (P < 0.01). Representative micrographs from mature seeds cell area (C) and columella area (D) were bordered with white color. Scale bar = 30 μm.
Figure 5.9 Altered cell morphogenesis of seed coat assessment though irregular cell angle. The regular hexagonal cell shape of wild type seeds promotes a relatively uniform 120° angle. To determine the shift in angle distribution we visually examined seeds using SEM micrographs. Interior angles within individual cells were measured (100 cells in 20 different seeds) for wild type (A), csld1-1 (B), csdl-2 (C), and cesa5 (D). An obvious shift arose in the distribution of angles from the csld1 alleles and cesa5 compared with wild-type. Asterisk indicates a significant difference from the wild type based on results of a Tukey multiple comparison test (P < 0.01). Standard error of three technical replicates from three independent batches of seeds were used as biological replicates. At right, (E,F,G,H) representative micrographs from mature seeds. Cell area bordered with white and arrows indicate an example of the cell angle measured. Bars = 30 μm.
Figure 5.10 Reduction in radial wall height in csl1. Laser scanning microscopy was used to measure the radial cell wall height in wild type, csl1-1, csl1-2 and cesa5. Error bars indicate standard error from the mean, asterisks indicate significant difference compared to the wild type based on Tukey mean separation (P<0.01).
Figure 5.11 Acid Insoluble crystalline cellulose content of seeds. Acid Insoluble crystalline cellulose content of seeds was determined calorimetrically. Error bars represent the standard error of three technical replicates from three independent batches of seeds as biological replicates. Asterisks indicate a significant difference from the wild type based on the results from a Tukey multiple comparison test (P<0.01).
Figure 5.12 Aberrant mucilage adherence in csld1 and cesa5. Mature seeds were imbibed in aqueous solution and stained with either Calcofluor White (A and B) or direct red23 (C and D). Wild-type, csld1-2, csld1-2 and cesa5 from left to right, scale bars = 100 μm. Whole seed were imaged using laser scanning confocal microscopy with 20 X objective (A and C) as well as 60X objective (B and D) to highlight the scaffold rays. DAPI filter was used for visualizing calcofluor white and DsRed filter for direct red23. Scale bars (B) = 10, (D) = 50μm
Figure 5.13 Relative quantitative expression of *CSLD-1* and *CESA5* during wild type seed development. Individual flowers were tagged at flowering, and developing siliques were collected at 0, 1, 3, 5, and 7 days post-anthesis (DPA). Error bars represent the standard error of three technical replicates from three independent batches of three biological replicates.
CHAPTER 6: LACCASES AND MICRONA INVOLVEMENT IN LIGNIN POLYMERIZATION

6.1 Summary

One of the central foci of the dissertation topic was to identify means to improve cell wall digestibility in grasses for biofuel. For lignin alteration we took a long term approach and focused on upstream regulation via microRNA control. MicroRNA regulation of lignin could have numerous targets and we looked at the downstream laccases as a potential target. Three different microRNAs were identified as displaying appropriate expression profiles. Overexpression and knock out variants were developed. This chapter described the molecular genetics approach to characterizing these microRNA lines.

6.2 Introduction

Lignin is essential to the plant and is a key player in water transport, upright growth and essential for the structural properties of plants (Chabannes et al., 2001; Jones et al., 2001). As mentioned in the chapter 1, lignin is primarily formed from the polymerization of three precursors: coniferyl, coumaryl and sinapyl alcohols, the products of the phenylpropanoid pathway, that are then oxidized to produce S, H and G lignin. One of the two enzymes responsible for the last step in creating the polymerized chain of monolignols are the laccases (LAC). Once the monolignols are activated in the apoplast, the highly reactive radical species couple with other radical species in a free radical coupling reaction that gives rise to the lignin polymer (Freudenberg and Neish, 1968; Sarkanen and Ludwig, 1971; Boerjan et al., 2003) This last step of lignin polymerization is also conducted by the peroxidases through a different chemical mechanism.

The laccases are copper containing oxido reductases that use oxygen to oxidize their substrate by a radical catalyzed mechanism (Claus, 2004). The laccases have a wide range of substrates that they can work with and are found in both prokaryotes as well as eukaryotes (Mayer and Staples, 2002; Claus, 2004). Laccases and their connection to the cell wall have been demonstrated in protoplasts as well as leaf wounding experiments conducted by De Marco and Roubelakis-Angelakis. In the protoplast experiment, they demonstrated that there was a steady increase in laccase activity after protoplast isolation,
while in their leaf wounding experiment, they reported that while laccase activity was detected immediately after wounding, peroxidase activity took four days to reach its peak (De Marco and Roubelakis-Angelakis, 1997). As a conclusion, they suggested that perhaps the laccases were involved in the initial healing of the cell wall. Leaf wounding often leads to lignification of the wound site, therefore their findings could perhaps be extrapolated to lignification as well.

In *Arabidopsis*, there are 17 laccases (*LAC*) that have been found through promoter GUS fusion constructs to be active in various tissue types and at various stages of development (Turlapati et al., 2011). Various laccase mutants have been studied to better understand the importance of the laccases, such as in the case of lac 2,8 and 15 where it was found that lac2 mutants had reduced root growth under PEG dehydration conditions, while the lac8 mutant flowered early and lac15 mutants had an altered seed color and a reduction in lignin quantity (Cai et al., 2006). Although it has been observed that single laccase mutations cause some altered phenotypes, it has been noted that the more pronounced phenotypes and changes are seen in double mutants where two laccases have been knocked out as in the case of the lac4 and 17 double knockout in which the single mutants appear to be phenotypically wildtype whereas the double mutant are stunted and small (Berthet et al., 2011).

To add to the complexity, the laccases are tightly regulated in their activity by their corresponding microRNAs. The microRNAs regulate their laccases by binding to a specific target sequence within the laccases. A single microRNA can regulate multiple laccases such as in the case of miR408 controlling *LAC* 12, 13 and 8 or it can be a one to one regulation such as in the case of miR857 and *LAC* 7 (Fahlgren et al., 2007). The microRNAs and corresponding laccases that will be the focus for this section has been organized in Table 6.1. The effect of microRNAs has also been an area of study in which microRNA overexpressors were created to observe their effect on their corresponding laccases and phenotypes. A study done in poplar looked the effect overexpression of miR397a would have on its laccases and found that there was a clear repression in laccase activity but also an up regulation in peroxidase activity (Lu et al., 2013). Another study in *Arabidopsis* looked at the overexpression and knockout of microRNA 857 and observed that while the overexpressor looks phenotypically like the wildtype, further
observations show that it has a weaker stem and smaller xylem vessels whereas the opposite effect is seen in the knockout (Zhao et al., 2015).

In order to better understand the complex regulation that is occurring to ultimately produce the lignin in plant tissue, we took a two-part approach to further investigate the laccases and their corresponding microRNA. In order to do so, we created silent mutations within our laccases of interest so that their corresponding microRNA would not be able to recognize the laccases, thereby uncoupling the regulation of the microRNA to their designated laccases. In addition, we also created microRNA overexpressor lines in which the microRNAs were under the control of the double 35S promoter.

6.3 Results

6.3.1 Construct design and reasoning

MicroRNAs recognize their target based on the homology in the sequence of the target to the mature miRNA sequence. In order to create laccase genes that would not be targeted by their corresponding miRNA, the target region of the laccase that contain the homologous sequence can be altered by the introduction of a silent mutation (Jones-Rhoades and Bartel, 2004). The laccases of interest have been modified so that they are no longer under the control of the miRNA. Full length laccase genes were amplified from the cDNA from Arabidopsis using gene specific primers. Then utilizing primers that contained the silent mutation, the laccases were amplified as 5’ and 3’ fragments to introduce the silent mutation and the fragments were digested with a restriction enzyme and religated using t4 DNA ligase to make the full length mutated laccases. The mutated laccases were then cloned into the pCXN1250 vector and the orientation of the genes were confirmed using the 2x35S forward primer and a gene specific reverse primer.

In order to examine the roles of individual laccases, each of the mutated (microRNA undetectable) laccases were overexpressed using a 35S overexpressor promoter. In addition, to examine the role of the microRNAs, they were also put under the control of the 35S promoter. In total, seven laccases were mutated and their three corresponding microRNAs were overexpressed.
6.3.2 Phenotypic characterization

All of the measurements and studies were conducted on independent lines (T2 or T3 generation) that were PCR verified by the presence or absence of the 35S promoter as being homozygous.

As mentioned earlier, due to the understanding that lignin is an important component in efficient water transport though the plant, we considered if an unregulated increase of the laccases alter the plants’ drought/dehydration tolerance? In order to systematically evaluate the effects of the mutated laccases and the microRNAs, we first started with microRNA 857 that only had laccase 7 under its control. In order to investigate, polyethylene glycol (PEG) dehydration assays were conducted on both germinating seeds and on root growth on the four mutated laccase lines (from two independent lines). PEG assays are often utilized to determine if plants are more or less tolerant to desiccation stress by altering the water potential of the medium on which they are placed (Verslues et al., 2006). We hypothesized that if there were to be unregulated activity of the laccase, the plants may be more tolerant of water stress. Four different water potentials (-0.25 MPa, -0.5 MPa, -0.7 MPa and -1.25 MPa) were tested. However, none of the four lines were significantly more tolerant of the dehydration stress as compared to wildtype. This outcome was true for both the seedling as well as the root length assay (data not shown).

In looking at mature plant height in the independent lines of lac7 mutants, there was no significant difference in height as compared to wildtype (Figure 6.1). We also paid close attention to the other homozygous lines from the other sets of microRNA and laccases to assess whether there would be an appreciable difference in their phenotype. In addition, we wondered whether an alteration in the laccases or microRNAs could alter herbicide tolerance. To investigate, we plated all homozygous lines from the mutated laccases as well as the microRNA overexpressor lines onto plates with 3nM and 5nM concentrations of isoxaben, a known cellulose biosynthesis inhibitor. However, none of the lines appeared to be any more tolerant of the CBI than the wildtype control, suggesting that the alteration in laccases or microRNA expression does neither make the plants more or less tolerant of the herbicide. Although the majority of the lines had no apparent differences, a line from the mutated lac17 (line 22) group did display a significant reduction
in mature plant height as well as a more “bushier” phenotype that may be a subject for further investigation in the future (Figure 6.1).

6.3.3 Alternative areas of potential interest

We also assessed the possibility that perhaps there may be an alteration in the seed coat due to the abundance of flavonoids and proanthocyanidins that accumulate within the seed coat (Lepiniec et al., 2006; David et al., 2014). In addition, one of the well-studied transparent testa (tt) mutants, tt10 exhibited a delayed browning of the seed coat (testa) that was found to be caused by a mutation in an oxidase responsible for the polymerization of flavonoids that was concluded to function as a laccase like flavinoid oxidase (Pource et al., 2005). Publications have reported that alteration in proanthocyanidin and anthocyanin content in the seed increased permeability of the seeds to tetrazolium salt (Debeaujon et al., 2000). We wondered if an alteration in the laccases could result in a change in the seed coat composition due to the change in abundance of the products from the branch of the phenylpropanoid pathway. We tested all homozygous lines by incubating the seeds in the tetrazolium salt solution overnight, and examined under the light microscope. Normal wildtype seeds are impermeable to the salt solution therefore there would be no change in seed color. However, seeds that have a compromised seed coat allows for the solution to infiltrate, turning the seeds a red color. Only one of the lines, mutated lac4 had a semi permeable seed coat which resulted in stained seeds when examined (Figure 6.2).

6.4 Discussion

Ultimately there were no trends in major phenotypic alterations that could be deduced from the overexpression of the mutated laccases nor from the overexpression of their corresponding microRNAs.

It may have been too high of an expectation to hope that the overexpression of either the laccase or the microRNA would cause a drastic change in phenotype even with our approach of the mutated. Earlier studies looking at the effect of mutated (knockouts) of laccases have reported that single laccase mutants were phenotypically like the wild type, and that only when the mutants were crossed, did they start to see phenotypic
alterations, even then, the alterations were only slight and under severely stressful conditions (such as 24-hour light) (Berthet et al., 2011). A triple knockout of lac4,17 and 11 have been reported to have a significant reduction in lignin (Zhao et al., 2013). This may suggest that an alteration (in our case, and overexpression) of one laccase gene may not be enough to cause any change as the phenotypes and alterations are only seen with multiple compounding alterations. In addition, it was also reported that the disruption of LAC17 caused a redirection of the phenylpropanoid pathway (Berthet et al., 2011). Which could suggest that there are regulatory checks and balances in addition to the microRNA regulation that we may not be completely aware of. If noticeable alterations can only be seen with such drastic knockdowns, as a triple mutant, it may also be possible that the opposite (an overexpression) of a single laccase would have had very little effect.

In addition, there have been studies that have investigated the importance of copper in microRNA and laccase regulation (Abdel-Ghany and Pilon, 2008; Lu et al., 2013). As copper is a critical component of the laccases, it has been shown that in low copper concentrations, microRNA concentration increases while the laccase concentration decreases (Abdel-Ghany and Pilon, 2008) which would also collaborate with the findings that in a high copper concentration environment (could lead to toxicity) plants accumulate peroxides, antioxidants as well as peroxidases (also an enzyme that can facilitate the last step in lignin polymerization) (Lu et al., 2013). In such high copper environments, it was found to induce lignin accumulation (Robson et al., 1981; Chen et al., 2002; Zhang et al., 2007). Lu et al suggest that perhaps laccases and peroxidases are reciprocally regulated.

An alternative hypothesis as to why we did not observe a distinct/clear phenotype may lie in the redundancy of function with the peroxidases. The peroxidases as mentioned earlier also catalyze the last step in lignin biosynthesis by acting on the lignin monomer to polymerize. It could be very likely that in the case of the microRNA overexpressor lines, the peroxidases were able to take the place of the suppressed laccases thereby resulting in no outward change in the phenotype of the plants. This may be an interesting avenue to follow and perhaps looking at expression levels of the microRNAs and also the peroxidases in such plants could help to illuminate what kind of roles the enzymes are playing.
6.5 Materials and methods

6.5.1 Mutated laccases and microRNA overexpressor constructs

Full length laccase genes were amplified from the cDNA from Arabidopsis using gene specific primers. Then utilizing primers that contained the silent mutation, the laccases were amplified as 5’ and 3’ fragments to introduce the silent mutation and the fragments were digested with a restriction enzyme and religated using T4 DNA ligase to make the full length mutated laccases. The mutated laccases were then cloned into the pCXN1250 vector and the orientation of the genes were confirmed using the 2x35S forward primer and a gene specific reverse primer.

The pre-miRNA sequences were obtained from the mirbase database (http://www.mirbase.org/) and the sequence that span the miRNA and the stem loop were amplified using the Arabidopsis genomic DNA and cloned into pCXN1250 vector using TA cloning and the correct orientation of the insertion was verified using the double 35S promoter forward primer and the miRNA specific reverse primer.

6.5.2 Arabidopsis transformation

The plasmids that contain the mutated laccases and the microRNA overexpressors were transformed into Agrobacterium tumefaciens (strain: Gv3101) by electroporation and the transformed cells were utilized to transform Arabidopsis by the floral dip method (Clough and Bent, 1998). The resulting seeds from the transformed plants were collected and screened for positive transformants.

At least two independent lines (when possible) from each construct were selected and grown to the T2 line under hygromycin selection. At the T2 stage the presence of the transgene was verified using PCR- this shows that the plants were at least heterozygotes for the insertion. The T3 plants were grown on non-selective media and the 10 plants per line were genotyped to verify that all plants contain the transgene therefore, now assuming that they are homozygous. The 35S primers were utilized to verify for homozygosity.
6.5.3 PEG dehydration

Seeds were sterilized and vernalized at 4°C then plated directly onto the pre-prepared PEG plates. PEG plates were made as described in Verslues et al (2006). Briefly, half MS plates were made and once solidified, the overlay solution (half MS liquid+ PEG) was poured over the top in a 2:3 ratio and left in the hood overnight to equilibrate the solid medium to the correct water potential. The seedlings were grown upright in the growth chamber set at 16hours light and 8hours dark cycle.

6.5.4 PEG dehydration seedling assay

Seven day old seedlings that were grown on half MS media were transferred to their corresponding PEG plates. The plates were placed upright in the growth chamber on a 16-8hour light/ dark cycle and their rate of root elongation measured over the course of 48 hours.

6.5.5 Tetrazolium salt uptake

A 1% (w/v) solution of tetrazolium salt dissolved in water was used and seeds incubated at room temperature in the dark. The seeds were examined after 48hours incubation time.
6.6 Tables and figures

Table 6.1 Table of microRNAs with their corresponding laccase(s) under their control.

<table>
<thead>
<tr>
<th>MicroRNA 397a</th>
<th>LAC7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LAC14</td>
</tr>
<tr>
<td></td>
<td>LAC2</td>
</tr>
<tr>
<td>MicroRNA 408</td>
<td>LAC12</td>
</tr>
<tr>
<td></td>
<td>LAC13</td>
</tr>
<tr>
<td></td>
<td>LAC3</td>
</tr>
<tr>
<td>MicroRNA 857</td>
<td>LAC7</td>
</tr>
</tbody>
</table>
Table 6.2 PEG germination and dehydration assay media and overlay solution composition. Media and overlay solution had the same composition only differing in the inclusion/ exclusion of agar.

<table>
<thead>
<tr>
<th></th>
<th>Per Liter of Solid media</th>
<th>Per Liter of overlay solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS Salt</td>
<td>2.2g</td>
<td>2.2g</td>
</tr>
<tr>
<td>MES</td>
<td>1.2 g</td>
<td>1.2g</td>
</tr>
<tr>
<td>Agar</td>
<td>15g</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PEG 8000 Per Liter of overlay solution</th>
<th>Final water potential of plate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0g</td>
<td>-0.25</td>
</tr>
<tr>
<td>250g</td>
<td>-0.5</td>
</tr>
<tr>
<td>400g</td>
<td>-0.7</td>
</tr>
<tr>
<td>550g</td>
<td>-1.2</td>
</tr>
</tbody>
</table>
Figure 6.1 Mature plant height. Mature plant height of mutated laccase 7 homozygous independent lines (A) and mutated lac17 homozygous line (B). The bars are the mean of three biological replicates and error bars indicate standard deviation from the mean.
Figure 6.2 Tetrazolium salt uptake in seeds. A mutated *lac4* line showed staining that was in between the positive and negative controls. Scale bar = 100μm.
CHAPTER 7: CONCLUDING REMARKS

In this dissertation, the introductory chapter gave a broad overview of the plant cell wall starting with the major components of the wall - cellulose, lignin and hemicellulose as well as the difference in composition between the dicot (type I) and monocot (type II). The secondary cell wall was also discussed and of its necessity in cells that require additional reinforcement and structure. Moving into the molecular mechanisms of cell wall biosynthesis, the CSL gene family with a strong emphasis on the CESAs - who are responsible for cellulose biosynthesis were also discussed as well as the use of CBIs as a tool to further our understanding of cellulose biosynthesis.

From the introduction of the CESA genes, we moved forward to the identification of the CESA from the new C4 model species Setaria viridis (chapter 2). We were able to bioinformatically identify the CESAs based on their sequence and their defining characteristics of the QxxRW motif as well as the CxxC motif. We found that in Setaria there were ten putative CESA genes and their distribution between as the primary CESAs and the secondary CESAs were similar to that of Arabidopsis. From the information gained from the phylogenetic analysis, we focused on two Setaria CESAs, S2 (secondary CESA) and A2 (primary CESA) and found that their expression pattern through qRT-PCR as well as through heterologous expression in Arabidopsis that their promoter GUS fusion lines were consistent with what we had expected from the bioinformatic analysis. In addition, we also quantified the major cell wall components in Setaria. Moving on to a more established model crop, we identified a Sorghum mutant (RedforGreen, RG) with an altered cell wall composition (chapter 3). It had been initially identified by its hyper-accumulation of pigment but upon further investigation we found that the RG had an alteration in the phenylpropanoid pathway which resulted in improved saccharification efficiency of the stems.

In order to gain a better understanding of the molecular mechanism behind cell wall biosynthesis, we then turned our attention to the classic model plant Arabidopsis thaliana. We made efforts to characterize the HULK mutant that had been found from a heterologous expression line that had been created utilizing a Setaria CESA (chapter 4). Ultimately, it was determined that the mutant had been caused by an insertional event rather than by the heterologous expression of a gene and from the results of the ACC treatment experiment, it appears to be an ethylene mutant. However, the ability of the
mutant to grow significantly larger than the wildtype is still an attractive trait from a biofuel production perspective. In chapter 5, we returned to focus our attention on the CSL gene family, in particular in the CSLD family. Through a technique called phage display, we were able to identify multiple CESA interactors, one of which we focused on was CSLD1. In particular, we focused on the role of CSLD1 in seed coat biosynthesis and the adherence of the mucilage halo surround the seed once they are imbibed in water. We determined that CSLD1 is involved in the cell wall biosynthesis in the seed coat, and more specifically it appears to be involved in the cellulose biosynthesis in the seed coat cell wall.

In chapter 6 we focused our attention on to the other main component of the cell wall: lignin. Specifically, we looked at the group of enzymes that are responsible for the last step in lignin biosynthesis- the polymerization of the monolignols and the microRNAs that control their activity. Through the experiments, we realized the complex regulation involved in lignin biosynthesis and that a simple alteration in one enzyme may have very little to no effect on the outcome.

As a conclusion, through the studies conducted and discussed above we hope that we have been able to contribute to the field of cell wall biology from various aspects. From searching for and identifying the CESA genes in newly established model species to furthering our understanding of the other members of the CSL family, such as with the CSLD1. In addition, the characterization and the compositional analysis of various mutants as well as trying to unravel the complexity of lignin biosynthesis through new and unique routes.
REFERENCES


Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Cézard L, Le Bris P, Borrega N,


Voiniciuc C, Dean GH, Griffiths JS, Kirchsteiger K, Hwang YT, Gillett A, Dow G, Western TL, Estelle M, Haughn GW (2013) FLYING SAUCER1 is a transmembrane
RING E3 ubiquitin ligase that regulates the degree of pectin methylesterification in Arabidopsis seed mucilage. Plant Cell 25: 944–959


deposition and interferes with the trafficking of cellulose synthase complexes and their associated proteins KORRIGAN1 and POM2/CSI1. Plant Physiol 167: 381–393
VITA
1. Place of birth: Ami machi, Ibaraki, Japan
2. Education: Bachelor of Arts in Biology at Transylvania University, 2011
3. Professional positions: None held
4. Publications:


5. Mizuki Tateno