8-2016

Modification of carbohydrate active enzymes in switchgrass (*Panicum virgatum* L.) to improve saccharification and biomass yields for biofuels

Jonathan Duran Willis
University of Tennessee, Knoxville, jdwillis@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Jonathan Duran Willis entitled "Modification of carbohydrate active enzymes in switchgrass (*Panicum virgatum* L.) to improve saccharification and biomass yields for biofuels." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Plants, Soils, and Insects.

C. Neal Stewart, Major Professor

We have read this dissertation and recommend its acceptance:

Nicole Labbe, Feng Chen, Jay Chen

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Modification of carbohydrate active enzymes in switchgrass (*Panicum virgatum* L.) to improve saccharification and biomass yields for biofuels

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Jonathan Duran Willis
August 2016

Dedication

My dissertation is dedicated to my parents, James R. Willis and Pamela G. Willis, my wife Kazuyo Shimizu and to all the family friends who have given me advice, patience, and courage to complete this work.

Acknowledgments

I want to thank all of those who have been especially helpful during my graduate studies. A special thanks to Dr. C. Neal Stewart, Jr for giving me the chance to grow and develop in his lab. I have garnered a lot of new found knowledge and respect for the sciences from him and will always be grateful. I am very grateful for the time, guidance, and the commitment of my dissertation committee members, Dr. Nicole Labbe, Dr. Feng Chen, and Dr. Jay Chen.

I am grateful to all the assistance with writing and developing science projects with Mitra Mazarei. I appreciate all the support and friendship from Reginald Millwood during my time in (and out of) the lab. Thank-you to all my fellow my co-authors and Stewart lab members (past and present) who have helped me.

I would like to also thank the University of Tennessee Racheff Endowment, the US DOE Bioenergy Science Center (BESC), and the US DOE Advanced Research Projects Agency for Energy (ARPA-E) for funding.

Thank-you to my parents and lovely wife for all the time, patience, and never letting me give up.

Abstract

The natural recalcitrance of plant cell walls is a major commercial hurdle for plant biomass to be converted into a viable energy source as alternative to fossil fuels. To circumvent this hurdle manipulation of carbohydrate enzymes active in the cellulose and hemicellulose portions of the plant cell wall can be utilized to improve feedstocks. Production of cellulolytic enzymes by plants have been evaluated for reducing the cost associated with lignocellulosic biofuels. Plants have successfully served as bioreactors producing bacterial and fungal glycosyl hydrolases, which have altered plant growth to improve saccharification. A bioprospecting opportunity lies with the utilization of insect glycosyl hydrolases for transgenic production in plants. Lessons learned from microbial hydrolase expression can be applied to insect hydrolase expression along with gene stacking to develop autohydrolysis plant lines.

A step toward production of insect cellulases in plants was performed by insertion of the endo-glucanase TcEG1 gene, from *Tribolium castaneum*, into switchgrass. Transgenic lines overexpressing TcEG1 produced a functional enzyme with an optimal alkaline pH activity of 12.0. Recalcitrance was assayed by performing saccharification analysis, in which one line was superior over non-transgenic control; this line also had reduced 9% lignin content. Transgenic lines developed narrow stems, although biomass yield was unchanged due to increased tiller number and cell wall thickness.

Grasses contain a relatively high amount of glucoarabinoxylan in their cell walls, which cross links with lignin. By down-regulation of a uridine diphosphate arabinomutase (UAM) gene via

RNAi, it was hypothesized that attenuated production of this carbohydrate transferase would increase saccharification of switchgrass biomass from a disruption of cross linking. Transgenic events showed a reduction in arabinose content (up to 58%) and altered arabino-side chains, however saccharification was unchanged. UAM transgenic switchgrass showed a red node phenotype, which could be in response to increased lignin biosynthesis. A model of UAM-influenced cell wall interactions was proposed and will used to build hypotheses for future –omics research.

In summary, switchgrass saccharification and biomass yield can be increased by introduction of carbohydrate active enzymes. Combination of presented transgenic lines with low-lignin germplasm could be utilized to further improve saccharification yield.

Table of contents

List of tables

List of figures

List of abbreviations

AU = absorbance unit

AX = arabinoxoylan

AZCL-xylan = azo-crosslinked-xylan

AZCL= azo-crosslinked

BG = β-glucosidase

BMCC = bacterial microcrystalline cellulose

CaMV35S = cauliflower mosaic virus 35S promoter

CBH = cellobiohydrolase

CBM = carbohydrate binding module

CMC = carboxymethylcellulose

CMV = cucumber mosaic virus promoter

CWD = cell wall degrading enzyme

FLP = flippase recombinase

FU = fluorescence units

EG = endo-1,4-β-glucanase

ER = endoplasmic reticulum

GH = glycosyl hydrolase

GHF = glycosyl hydrolase family

GluB-1 = rice glutelin endosperm promoter

GmHSP = *Glycine max* heat shock protein

GusA = β-glucuronidase reporter gene

HEPES = 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

Hor2-4 = barley B1 hordein promoter

Mac = promoter combination of the Ti plasmid mannopine synthetase and CaMV 35S

MUC = 4-methylumbelliferyl-β-D-cellobioside

MUL = 4-methylumbelliferyl-β -D-lactopyranoside

n.d. = not determined

n.o. = not optimized

n.r. = not reported

pNPG = *p*-nitrophenyl β-D-glucopyranoside

PvUAM = *Panicum virgatum* UDP-arabinomutase

RBB-xylan = remazol brilliant blue R-D-xylan

RbcS-3C = Rubisco small sub-unit promoter

TcEG-1 = *Tribolium castaneum* endoglucanase 1

TRBO = tobacco mosaic virus RNA-based overexpression vector

U = unit

Xyl = xylanase

Zm-PEPc = maize phosphoenolpyruvate carboxylase

Chapter 1: Introduction

1.1 Lignocellulosic feedstocks for biofuels

Over the past decade, perpetual radical flux of the petroleum market has sparked renewed interest in biofuels. Biofuels have been divided into generations based on the type of biological material and energy source derived from their process. The first and second biofuel generations are categorized by the plant material used for fermentation to produce ethanol as a fuel source or fuel additive. The first biofuel generation derives ethanol from starch and sucrose rich biomass such as corn kernels and sugarcane extracts (Yuan et al., 2008). Amylose, glucose, and sucrose, glucose, and other carbohydrates are readily stored in plant reproductive organs (example kernels) or stems which are easily fermented by yeast into ethanol. Starch rich feedstocks are poorly suited for a sustainable market as these crops require high inputs including fertilizer and pesticides. These sorts of agronomic inputs are more economically feasible for food and feed than for bioenergy crops (Yuan et al., 2008).

The second biofuel generation utilizes lignocellulose biomass to produce fuels. Entire aboveground biomass from dedicated biomass crops or row crop residue, such as corn stover can be used to produce ethanol. Cellulosic ethanol, however, is expensive to produce because of the plant cell wall's natural recalcitrance to degradation (Wyman, 2007). Cellulose is the main targeted substrate for fermentation, but is tightly bound to hemicelluloses and lignin. Costly pre-treatment steps are required to free the cellulose and allow access for saccharification enzymes to convert cellulose into glucose in order to be utilized by yeast (Himmel et al., 2007). Pre-treatment steps and by-products from the hemicelluloses and lignin can carry fermentation inhibitors, which are also problematic for commercial production (Himmel et al., 2007).

Cellulosic feedstocks are considered to be more environmentally friendly over starch-rich feedstocks by providing a near net-zero carbon footprint (Yuan et al., 2008).

A leading second generation feedstock is the perennial C4 grass, switchgrass (*Panicum virgatum* L.), which is not used for human food and seldom used as animal forage. It can grow on marginal land that cannot support traditional row crops. C4 plants have higher water use efficiency than C3 plants. Furthermore, switchgrass requires much less fertilizer than maize and most other C4 cereal crops (Vogel, 2008; Wright and Turhollow, 2010; Wullschleger et al., 2010). Switchgrass can be readily harvested using current hay cutting and baling equipment. Utilization of contract plantings from biorefineries improves a farmer's option to plant switchgrass as a permanent installation with a set value over several years instead of being based on future's cost of decided annually on other crops (Griffith et al., 2014; Griffith et al., 2012).

Recent advances in plant transformation technologies for switchgrass have increased transformation and regeneration recovery rates making it a viable biotechnology crop to be utilized for cell wall studies (Burris et al., 2009; Li and Qu, 2011). Utilizing biotechnology strategies would provide new transgenic germplasm for commercial use and aid in understanding the biological mechanisms of cell wall development.

1.2 Plant cell wall's natural recalcitrance to degradation

The intertwining complexity of three main cell wall components (cellulose, hemicellulose, and lignin) is the overarching factor for cell wall recalcitrance to degradation. Lignin is the most studied cell wall component with regards to lignocellulosic recalcitrance. Along with lignin, the

plant cell wall consists of cellulose and hemicellulose. Cellulose is comprised of long chains of hexose sugars that result in crystalline structures in cell walls (Cosgrove, 2005). Hemicelluloses is the collective term for the non-cellulose polysaccharides such as pentose xylans, which are laid down in a more randomized designed (similar to lignin) (Scheller and Ulvskov, 2010; Rennie and Scheller, 2014). Hemicellulose crosslinking to cellulose fibrils and lignin increases the strength of cell wall to degradation (Cosgrove, 2005). The composition of plant cell walls is an evolutionary testament for structural support, nutrient and water distribution, and pest protection (Tavares et al., 2015). Specifically the factors of non-cellulose polysaccharide cross-linking, lignin content and composition, and cellulose micro-fibril crystallinity influence simple sugar release from plant materials (Tavares et al., 2015; Silveira et al., 2013; Chen and Dixon, 2007).

Lignin content and composition are highly correlated to plant cell wall recalcitrance, hence lignin biosynthesis was the first target for genetic manipulation in dedication biofuel crops (Chen and Dixon, 2007). Lignin's polypropanoid monolignol monomers are guaiacyl (G), syringyl, (S), and (H) *p*-hydroxyphenol, which are produced from manipulation of phenylalanine (Rencoret et al., 2011). Even though there are only three monolignols, the regulatory mechanism for the amounts and connections of each unit to form lignin polymers is poorly understood. The S and G units are the major concern in angiosperms when evaluating for recalcitrance as a higher G content is less favorable for sugar release. G units have more carbon-carbon bonds compared to the S units. The H unit is at minimal levels in dedicated angiosperm biofuel crops and not considered a recalcitrance target. Manipulation of the amount of S/G by disruption of key lignin

pathway genes has proven efficacious for improving ethanol yields (Chen and Dixon, 2007; Fu et al., 2011).

Hemicellulose is the non-cellulose polysaccharides bound to cellulose by various side chains, which may play a role in increasing cell wall recalcitrance (Scheller and Ulvskov, 2010). Hemicelluloses surround cellulose fibers like a sheath preventing accessibility for enzymes to degrade cellulose. Hemicellulose cross links with lignin polymers adding more protection from enzymatic interaction (Scheller and Ulvskov, 2010). Hemicelluloses are heavily branched with side chains, which prevent their own crystallization, but these branches increase their adherence to cellulose and are a potential target for reducing recalcitrance (Cosgrove, 2005).

The natural crystalline structure of cellulose plays a role in plant cell wall recalcitrance. Cellulose is a linear polymer of cellobiose units connected by β-1,4 linkages in a continuous chains (Cosgrove, 2005). Cellulose is formed out of the hexameric cellulose synthase complex whereas cellulose strands bind together forming a crystalline ribbon-like structure (Cosgrove, 2005; Park and Cosgrove, 2012). Disruption of the crystalline structure by either over-expression of natural plant glycosyl hydrolases or by introduction of foreign proteins has proven to reduce recalcitrance (Tavares et al., 2015; Takahashi et al., 2009). Each plant cell wall component acts in synergistic nature with a complex weaving to maintain plant structure and prevent cell wall degradation (Tavares et al., 2015; Silveira et al., 2013; Chen and Dixon, 2007).

1.3 Plant cell walls are not created equally

Overall the above sections stress the broad commonalities of the plant cell wall. However, there are reported areas where the plant cell wall differs across taxonomic clades highlighting the importance of utilizing the most suitable plant model to the commercial crop. Plant cell walls are conserved genetically and phenotypically within taxonomic clades, however there are differences in gene expression and cell architecture (Carpita and Vergara, 1998; Carpita, 2011). While the base materials are the same across clades, the known aspects of plant cell wall synthesis and arrangement are different. Postulations have been made based on sequence data that few orthologous maize and Arabidopsis cell genes exist (Bosch et al., 2011; Carpita and McCann, 2008). Dicots and some monocots have Type 1 cell walls which are roughly equal parts cellulose and hemicellulose with xyloglucans (Carpita, 1996). The grasses (or Poales) consist of Type 2 primary cell walls where the cellulose bundles are wrapped in glucouronarabinoxylans, high levels of hydroxycinnamates, and lower levels of pectins and structural proteins (Vogel, 2008). Monocots and dicots plant cell walls have similar cellulose polymer and architecture. For both monocots and dicots, cellulose microfibrils are putatively produced in the same fashion by synthesis from cellulose synthase complexes (Carpita, 1996). The main difference between these two groups lies at the content and composition of the lignin and hemicelluloses components of the cell walls.

There is conservation of monolignols between dicots and monocots, but their ratios differ. Dicot lignin is composed of equal amounts of G and S units with very few H units (Cesarino, 2012). Monocots contain higher levels of H lignin compared to dicots and roughly equal parts S and G.

The increased H level in monocots could be attributed to the abundance of ferulic acid and *p*-coumaric acid which have a polypropanoid structure (Vogel, 2008). The relatively small portion of H unit present in grasses poses a step where monocots and dicots differ.

The cross linkages of the hemicellulose and lignin are another cell wall distinction between monocots and dicots. Xylans make up the majority portion of hemicellulose. Monocot cell walls have higher occurrences of cross linking of xylans than dicots (Faik, 2010; Zeng et al., 2008). Dicots and some monocots have Type1 cell walls which are defined as utilizing xyloglucan as their main hemicellulose cross linker (Carpita, 1996). The majority of monocots have Type 2 cell walls defined as utilizing glucuronoarabinoxylans serving as the main cross linker (Carpita and McCann, 2008). The initiation points for biosynthesis of new xylan chains also differ between monocots and dicots (York and O'Neill, 2008). The xylan pathway is partially understood and it is unknown if a large complex (like cellulose synthase) is required for xylan formation (York and O'Neill, 2008; Faik, 2010). The greater abundance of heteroxylan mixtures in grass cells lends to the reduction in cellulose when compared to dicots at the secondary level of cell wall biosynthesis (Burton and Fincher, 2014).

Switchgrass lines from commercially available germplasm were selected to characterize biofuel traits, instead of model crops, to mitigate cell wall variability listed above. Switchgrass commercial seeds were selected for their suitability to tissue culture and genetic transformation and used for all subsequent experiments. Transgenic switchgrass germplasm generated from

these studies would be readily available for field evaluations and breeding with high-yield commercial varieties.

The overall goal of this dissertation was to characterize and demonstrate methods to develop genetically engineered improved biofuel feedstocks, with an emphasis on grasses. The first step was detailing the methodology of integrating cellulase enzymes into biofuel feedstocks. The review of past literature was further built upon by including the feasibility of using insect cellulases and strategies for improved transgenic protein yield. In addition to the review, an investigation into genetic modification of non-lignin cell wall pathways in switchgrass were performed. Targeting the cellulose pathway, we introduced the alkaline beetle cellulase, TcEG-1, from the red flour beetle (*Tribolium castaneum*) into switchgrass to evaluate this enzyme activity and effect on plant biomass yield and morphology. Taken together we have successfully generated and characterized two different transgenic switchgrass lines showing variable effects on plant cell wall growth. Targeting the hemicellulose pathway, we constructed an RNAi knockdown vector for a switchgrass UDP-arabinomutase (*PvUAM1*) involved in the decorating the xylan backbone with arabinofuranose side chains for cross linking to lignin. The hypothesis tested was that down-regulating *PvUAM1* would disrupt the cross linking of arabinoxylan to lignin and increase saccharification. Modification with these two transgenes provided insight into how the cell wall architecture was modified beyond those of the intended targets.

1.4 References

Bosch, M., Mayer, C.D., Cookson, A. and Donnison, I.S. (2011) Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. *J Exp Bot* **62**, 3545-3561.

Burris, J.N., Mann, D.G.J., Joyce, B.L. and Stewart, C.N. (2009) An improved tissue culture system for embryogenic callus production and plant regeneration in switchgrass (*Panicum virgatum* L.). *Bioenerg Res* **2**, 267-274.

Burton, R.A. and Fincher, G.B. (2014) Plant cell wall engineering: applications in biofuel production and improved human health. *Curr Opin Biotechnol* **26**, 79-84.

Carpita, N.C. (1996) Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Physiol Plant Molec Biol* **47**, 445-476

Carpita, N. and Vergara, C. (1998) A recipe for cellulose. *Science* **279**, 672-673.

Carpita, N.C. (2011) Update on mechanisms of plant cell wall biosynthesis: how plants make cellulose and other (1->4)-beta-D-glycans. *Plant Physiol* **155**, 171-184.

Carpita, N.C. and McCann, M.C. (2008) Maize and sorghum: genetic resources for bioenergy grasses. *Trends Plant Sci* **13**, 415-420.

Cesarino, I., Araujo, P., Junior, A.P.D., Mazzafera, P. (2012) An overview of lignin metabolism and its effec on biomass recalcitrance. *Brazilian J Bot* **35**, 303-311.

Chen, F. and Dixon, R.A. (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* **25**, 759-761.

Cosgrove, D.J. (2005) Growth of the plant cell wall. *Nat Rev Mol Cell Biol* **6**, 850-861.

Faik, A. (2010) Xylan biosynthesis: news from the grass. *Plant Physiol* **153**, 396-402.

Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M., Jr., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A. and Wang, Z.Y. (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc Natl Acad Sci U. S. A.* **108**, 3803-3808.

Griffith, A.P., Haque, M. and Epplin, F.M. (2014) Cost to produce and deliver cellulosic feedstock to a biorefinery: Switchgrass and forage sorghum. *Appl Energ* **127**, 44-54.

Griffith, A.P., Larson, J.A., English, B.C. and McLemore, D.L. (2012) Analysis of contracting alternatives for switchgrass as production alternative on an east Tennessee beef and crop farm. *AgBioForum* **15**, 206-216.

Guo, D., Chen, F., Inoue, K., Blount, J.W. and Dixon, R.A. (2001) Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa. impacts on lignin structure and implications for the biosynthesis of G and S lignin. *Plant Cell* **13**, 73-88.

Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W. and Foust, T.D. (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804-807.

Konishi, T., Ohnishi-Kameyama, M., Funane, K., Miyazaki, Y., Konishi, T. and Ishii, T. (2010) An arginyl residue in rice UDP-arabinopyranose mutase is required for catalytic activity and autoglycosylation. *Carbohydr Res* **345**, 787-791.

Konishi, T., Takeda, T., Miyazaki, Y., Ohnishi-Kameyama, M., Hayashi, T., O'Neill, M.A. and Ishii, T. (2007) A plant mutase that interconverts UDP-arabinofuranose and UDP-arabinopyranose. *Glycobiology* **17**, 345-354.

Li, R.Y. and Qu, R.D. (2011) High throughput *Agrobacterium*-mediated switchgrass transformation. *Biomass Bioenerg* **35**, 1046-1054.

Mitchell, R.A., Dupree, P. and Shewry, P.R. (2007) A novel bioinformatics approach identifies candidate genes for the synthesis and feruloylation of arabinoxylan. *Plant Physiol* **144**, 43-53.

Park, Y.B. and Cosgrove, D.J. (2012) A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant Physiol* **158**, 1933-1943.

Rencoret, J., Gutierrez, A., Nieto, L., Jimenez-Barbero, J., Faulds, C.B., Kim, H., Ralph, J., Martinez, A.T. and del Rio, J.C. (2011) Lignin composition and structure in young versus adult *Eucalyptus globulus* plants. *Plant Physiol* **155**, 667-682.

Rennie, E.A. and Scheller, H.V. (2014) Xylan biosynthesis. *Curr Opin Biotechnol* **26**, 100-107.

Sattler, S.E., Funnell-Harris, D.L. and Pedersen, J.F. (2010) Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. *Plant Sci* **178**, 229-238.

Scheller, H.V. and Ulvskov, P. (2010) Hemicelluloses. *Annu Rev Plant Biol* **61**, 263-289.

Silveira, R.L., Stoyanove, S.R., Gusarov, S., Skaf, M.S. and Kovalenko, A. (2013) Plant biomass recalcitracne: effect of hemicellulos composition on nanoscale forces that control cell wall stregth. *J Am Chem Soc* **51**, 19048-52

Takahashi, J., Rudsander, U.J., Hedenstrom, M., Banasiak, A., Harholt, J., Amelot, N., Immerzeel, P., Ryden, P., Endo, S., Ibatullin, F.M., Brumer, H., del Campillo, E., Master, E.R., Scheller, H.V., Sundberg, B., Teeri, T.T. and Mellerowicz, E.J. (2009)

KORRIGAN1 and its aspen homolog PttCel9A1 decrease cellulose crystallinity in Arabidopsis stems. *Plant Cell Physiol.* **50**, 1099-1115.

Tavares, E. Q. P., De Souza, A. P. and Buckeridge, M.S., (2015) How endogenouse plant cell-wall degradation mechanism can help achieve higher efficency in saccharifiation of biomass. J Exp Bot **66**, 4133-43.

Vogel, J. (2008) Unique aspects of the grass cell wall. *Curr Opin Plant Biol* **11**, 301-307.

Wright, L. and Turhollow, A. (2010) Switchgrass selection as a "model" bioenergy crop: A history of the process. *Biomass Bioenerg* **34**, 851-868.

Wullschleger, S.D., Davis, E.B., Borsuk, M.E., Gunderson, C.A. and Lynd, L.R. (2010) Biomass production in switchgrass across the United States: Database description and determinants of yield. *Agron J* **102**, 1158-1168.

Wyman, C.E. (2007) What is (and is not) vital to advancing cellulosic ethanol. *Trends Biotechnol* **25**, 153-157.

York, W.S. and O'Neill, M.A. (2008) Biochemical control of xylan biosynthesis - which end is up? *Curr Opin Plant Biol* **11**, 258-265.

Yuan, J.S., Tiller, K.H., Al-Ahmad, H., Stewart, N.R. and Stewart, C.N., Jr. (2008) Plants to power: bioenergy to fuel the future. *Trends Plant Sc.* **13**, 421-429.

Zeng, W., Chatterjee, M. and Faik, A. (2008) UDP-Xylose-stimulated glucuronyltransferase activity in wheat microsomal membranes: characterization and role in glucurono(arabino)xylan biosynthesis. *Plant Physiol* **147**, 78-91.

Chapter 2: Transgenic plant-produced hydrolytic enzymes and the potential of insect gut-derived hydrolases for biofuels

A version of this chapter has been published by Frontiers in Plant Sciences with following authors Jonathan D. Willis, Mitra Mazarei, and C. Neal Stewart Jr.
Jonathan D. Willis drafted the manuscript and Mitra Mazarei and C. Neal Stewart participated in the drafts and revisions. Journal article citation: Willis J, Mazarei M and Stewart N (2016). Transgenic plant-produced hydrolytic enzymes and the potential of insect gut-derived hydrolases for biofuels. *Front. Plant Sci.* **7**:675. doi: 10.3389/fpls.2016.00675

2.1 Abstract

Various perennial C4 grass species have tremendous potential for use as lignocellulosic biofuel feedstocks. Currently available grasses require costly pre-treatment and exogenous hydrolytic enzyme application to break down complex cell wall polymers into sugars that can then be fermented into ethanol. It has long been hypothesized that engineered feedstock production of cell wall degrading (CWD) enzymes would be an efficient production platform for exogenous hydrolytic enzymes. Most research has focused on plant overexpression of CWD enzyme-coding genes from free-living bacteria and fungi that naturally break down plant cell walls. Recently, it has been found that insect digestive tracts harbor novel sources of lignocellulolytic biocatalysts that might be exploited for biofuel production. These CWD enzyme genes can be located in the insect genomes or in symbiotic microbes. When CWD genes are transformed into plants, negative pleiotropic effects are possible such as unintended cell wall digestion. The use of codon optimization along with organelle and tissue specific targeting improves CWD enzyme

yields. The literature teaches several important lessons on strategic deployment of CWD genes in transgenic plants, which is the focus of this review.

2.2 Introduction

The natural recalcitrance of plant cell walls is a major commercial hurdle for cellulosic biofuel production. Two economic barriers of biofuel production are pretreatment and hydrolytic enzymes (Furtado et al., 2014; Wyman, 2007). It has been hypothesized that using a transgenic plant vehicle for the production of hydrolytic enzymes would improve cellulosic biofuel economics. Thus far, plant species transformed to overproduce hydrolytic enzymes are not bioenergy feedstock crops, but rather, easily transformed species such as *Arabidopsis thaliana*, *Nicotiana tabacum* (tobacco), and *Oryza sativa* (rice) (Taylor et al., 2008). All of these models are annual plants, and therefore may not be entirely predictive of enzyme activity perennial crops forecasted to be the most useful for cellulosic bioenergy.

Plant cell walls are composed of three major types of polymers: cellulose, hemicelluloses and lignin. Each of these polymers are cross-linked and/or intertwined with another type to provide stereochemical strength (Cosgrove, 2005). Plant cell walls have successfully been broken down using mechanical and chemical means, however these processes can release metabolites that can inhibit fermentation. Furfural is one such metabolite that inhibits the growth of yeast (Boyer et al., 1992; Zhang et al., 2011b). A mixture of mechanical, chemical, and enzymatic means might be required to reduce cell wall recalcitrance in lignocellulosic biomass processing. Enzymes represent one of the highest material costs in lignocellulosic processing (Klein-Marcuschamer et al., 2012; Taylor et al., 2008).

Synergistic action of multiple enzymes are required for complete cell wall degradation. Four-to-five classes of enzymes are required to break down cellulose into its smaller cellobiose sub-units (Bhat, 2000; Lopez-Casado et al., 2008), which is, in turn, composed of two glucose units connected by a β-1,4 linkage. The tightly-packed crystalline structure of cellulose prevents enzyme accessibility. Hemicelluloses have variable polymeric structures composed of pentose sugars and a variety of side chains, which necessitates a suite of enzymes for hydrolysis (Saha, 2003; Scheller and Ulvskov, 2010).

Before the advent of rapid DNA sequencing technologies, glycosyl hydrolase (GH) enzymes were categorized based on their substrate specificity. Substrate specificity categorization was inadequate with discovery of enzymes with multiple catalytic domains and/or multiple substrate adherence (Henrissat, 1991). The previous substrate specificity classification system was updated and GHs were re-grouped into families (Henrissat and Bairoch, 1993; Henrissat and Bairoch, 1996). The CAZy database (CAZy.org) provides tools for classification and defining novel cellulolytic enzymes (Lombard et al., 2014).

Enzymes used for industrial cell wall hydrolysis are prepared from mesophilic fungi and bacteria. Optimal enzyme activity in such cases is typically 50 °C and a pH range of 5.0-7.0 (Bhalla et al., 2013; Bruce and Palfreyman, 1998). Enzymatic degradation of cellulose to glucose requires three enzyme classes: endo-glucanases, cellobiohydrolases, and β-glucosidases (Box 1). Endo-glucanases (EG) act randomly on internal portions of the cellulose polymer creating smaller chains. Cellobiohydrolases (CBH) (also reported as exo-glucanases) cleave

these smaller units at the end portion into cellobiose units. β-glucosidases (BG) cleave cellobiose units into individual glucose monomers. For efficient degradation of cellulose, each of these enzymes are required (Kostylev and Wilson, 2012).

If transgenic plants are to be used to overproduce GHs, there are three important aspects to be considered for the host plant: 1) as a biocatalyst production unit, 2) altered cell walls, and 3) auto-hydrolysis of biomass (Damm, 2016; Furtado et al., 2014; Lambertz et al., 2014; Taylor et al., 2008). Utilizing plants as enzyme factories is an intriguing concept. It may not be practically feasible in this case, so using plant cell cultures for enzyme overproduction has been proposed (Hellwig et al., 2004). The highest potential goal is to produce bioenergy feedstocks with altered cell walls for high saccharification and their own saccharification enzymes in tandem. An autohydrolytic plant system would require enzymes to be inactive, induced, or sequestered until harvest in order to not negatively affect plant growth. Research on GHs produced in plants is in its relative infancy.

In this review we will highlight successes and challenges of GH production in plants. Here we compare past use of fungal and bacterial GHs produced in plants and focus on lessons that can inform on selection of appropriate expression systems. We focus especially on the types and properties of insect-derived GHs and assess their potential use in plants. Strategies and recommendations for conducting research and development of CWD enzymes in plants will be discussed with an eye toward improving bioenergy crops and integrated biofuel systems.

2.3 Endo-glucanases in transgenic plants

Heterologous expression of lignocellulosic enzymes *in planta* for improved production of biofuels was proposed in the 1990s, but relatively few enzymes have been tested (Furtado et al., 2014; Lambertz et al., 2014; Lopez-Casado et al., 2008). Mostly, bacterial and fungal enzymes have been overproduced in plants to assess efficacy of the concept (Lambertz et al., 2014; Taylor et al., 2008). Endo-glucanases (EG) have, by far, been the most abundant class of CWD enzymes produced in plants (Table 2-1). The E1 cellulase from the bacterium *Acidothermus cellulolyticus* has been the most frequent EG produced in transgenic plants (15 papers; Table 1). One reason behind the attention given to E1 is that it is a hyperthermophilic enzyme, which would be inactive during temperatures favoring plant growth.

The first report on E1 production in plants was in tobacco in which E1 was targeted to the chloroplast. E1 was active with an optimal temperature of 81 ºC and pH of 5.25 (Dai et al., 2000a). The cauliflower mosaic virus 35S promoter (*35S*) was used to drive E1 gene expression in *Arabidopsis thaliana* for apoplast targeting (Ziegler et al., 2000). In this study, E1 was accumulated up to 26% of total soluble protein in primary leaves, which is the highest reported for transgenic plant produced hydrolases (Taylor et al., 2008; Ziegler et al., 2000). Tobacco E1 expressing lines were developed for transport to the chloroplast and found that not all parts of the enzyme would translocate, however the catalytic domain was sufficient to maintain enzyme activity (Jin et al., 2003). In an attempt to optimize E1 expression in tobacco, the hybrid constitutive promoter *Mac* (a combination of the Ti plasmid mannopine synthetase and *35S* promoters) and the *RbcS-3C* (Rubisco small sub-unit) promoter were compared; the *RbcS-3C*

was ranked higher in performance. However, the highest total soluble protein was reported as only 0.25%. Apoplast-targeted E1 transgenic potato plants produced 2.6% E1 relative to total soluble protein (Dai et al., 2000b). Similar studies in maize, using an enhanced version of the *35S* promoter driving the E1 gene resulted in 2.1% (Biswas et al., 2006) and 1.13% (Ransom et al., 2007) E1 accumulation relative to total soluble protein. Even though none of the above studies reported any negative effects to the host plants, E1 transgenic maize (Ransom et al., 2007), and tobacco cell walls developed differently than their respective non-transgenic parents. In each of these two studies increased saccharification was reported. Wildtype and E1 maize expressing lines were exposed to exogenous application of E1 and showed to be equivalent. The proposed reason for this is due to E1 is active *in vivo* during plant growth causing small nicks within the cellulose polymers. Addition of extra exogenous E1 application of enzyme was ineffectual due to the enzyme already hydrolyzing available cellulose recognition sites (Brunecky et al., 2011). In the same study, Brunecky et al, (2011) examined the transgenic E1 corn and tobacco made by Ziegelhoffer et al., (2001) via histological examination and postulated E1 was having an effect on cell wall growth during development (Brunecky et al., 2011; Ziegelhoffer et al., 2001). The presence of biologically active E1 during cell wall synthesis could play a role in reducing cellulose crystallinity or by creating gaps in the cell wall for increased enzyme accessibility (Brunecky et al., 2011). Aforementioned experiments demonstrate the utility of plant produced endoglucanases as a method for cell wall architecture manipulation.

One goal has been to use transgenic plants as biocatalyst producers in a commodity, such as E1 accumulation in maize seeds. The first 40 codons of E1 were codon-optimized in attempt to increase expression and endow tissue specificity (Hood et al., 2007). E1 in maize seed was further targeted to the cell wall, ER, and the vacuole. In the Hood et al. (2007) study, the ER appeared to be the best subcellular localization target, but none of the localizations led to off-effects in maize. Though active, the E1 enzyme was found to be truncated from a ~70 kDa size down to ~40kDa. Furthermore, stable enzyme was observed after harvesting and drying (Hood et al., 2007). The researchers introgressed the transgenes into elite germplasm (Hood et al., 2012) for further field experiments (Garda et al., 2015). E1 accumulation increased up to seven-fold in the elite germplasm. Field produced E1 in maize seed appeared to be active, thus illustrating an economically-viable platform to overproduce and store E1 until needed.

In rice, E1 has been studied for production in vegetative biomass apoplasts (Chou, et al., 2011; Oraby et al., 2007) and seed endosperm (Zhang, et al., 2012). Apoplastic targeting appeared to not affect plant growth (Chou, et al., 2011; Oraby et al., 2007), whereas the endosperm production strategy led to dwarfing and early flowering (Zhang, et al., 2012). While introduction of E1 led to a suboptimal host plant phenotype, there are multiple variables which need to be addressed. In both cases, E1 was targeted to the apoplast, but used two different constitutive promoters and two different germplasms (Chou et al., 2011; Oraby et al., 2007). Chou et al. (2011) reported higher protein yield at 6.1% over Oraby et al. (2007) 4.9%. When E1 was targeted to the endosperm it was active in dried seeds, but seeds were smaller than wild type seed (Zhang et al., 2012b). The researchers speculated that lower seed weight is an effect of the

transformation process as transgenic rice expressing non-cellulase proteins have a similar issue (Zhang et al., 2012b). No other phenotype was reported in the endosperm-targeted rice. Amongst the three cases, only Zhang et al. (2012b) performed an optimization experiment to demonstrate that the E1-produced rice maintained its high thermophilic optimum of 81 ºC.

E1 was transformed into duckweed (*Lemna minor* 8627) an aquatic plant, for accumulation in the cytosol in which extraction buffers were tested for efficacy. Three extraction buffers, sodium citrate (50 mM pH 4.8), sodium acetate buffer (50 mM pH 5), and HEPES (50 mM pH 8), were used and evaluated for their effects on enzyme yield and activity (Sun et al., 2007). The HEPES buffer extraction yielded more total soluble protein compared to other experimental buffers. When the proteins were extracted with citrate buffer and brief heating at 65 ºC enzyme activity was increased, although total recovered protein yield was lower than HEPES buffer extraction (Sun et al., 2007).

Transient expression of E1 was performed in sunflower to study the effects of various promoters: *CaMV35S*, *CMVar* (cucumber mosaic virus advanced replicating), and *TRBO* (tobacco mosaic virus RNA-based overexpression vector). Neither the *CMVar* nor *TRBO* led to E1 production. Utilizing the *35S* promoter, it was shown that addition of methyl jasmonic acid led to a four-fold increase in production of active enzyme (Jung et al., 2014).

Thermobifida fusca (formerly named *Thermomonospora fusca*), a thermophilic bacterium, produces Cel6A (formerly E2), an EG that was transformed into alfalfa, tobacco and potato for

cytosolic production. Active Cel6A showed no negative phenotype on whole plants. Cel6A was evaluated for thermostability and was found to be still active from 60-65 °C, but became inactive at higher temperatures in the absence of substrate which was comparable to commercially produced Cel6A (Ziegelhoffer et al., 1999). In a later study, Cel6A was targeted to tobacco chloroplasts, and no morphological changes were observed in the host plant (Yu et al., 2007). Cel6A showed an improved total soluble protein yield 0.6-4%.

Tobacco cells have been used to express two other EGs from bacteria. The t-EGI from *Ruminococcus albus* was targeted to the cytosol of tobacco cells. The t-EGI was active at 30 °C. The t-EGI tobacco plants successfully autohydrolysed their cell walls (Kawazu et al., 1999). The EG SSO1354 from *Sulfolobus solfataricus* was successfully expressed in tobacco apoplasts and endoplasmic reticulum. SSO1354 produced in tobacco had highest activity at 90 °C and pH 4.5 and showed no growth differences relative to non-transgenic plants. SSO1354 tobacco was tested with ionic liquid pre-treatment solutions and found to still be active (Klose et al., 2012).

The TrCel5A EG from the fungus *Trichoderma reesei* was produced in tobacco under two construct parameters (Klose et al., 2013). The TrCel5A was active under both construct versions at 55 °C and pH 4.8 (Klose et al., 2013). Under the control of the *35S* promoter, TrCel5A produced dwarfed tobacco even with apoplast targeting. However, using an ethanol-inducible, *alcA*, promoter, TrCel5A was produced upon induction and the enzyme was indistinguishable from that in 35S-construct plants (Klose et al., 2013). Similar to the research described above in maize, tobacco progeny produced more TrCel5A than their parents (Klose et al., 2015). In this

study, TrCel5A was produced in apoplasts or the ER. In both cases, T2 progeny had wrinkled leaves with spotted necrosis, which would be considered important off-target effects (Klose et al., 2015).

The recent additions of synthetically creating genes by combing databases and preparing sequences *in silico* has been used to generate EG. A proprietary, synthetically-designed EG (psEG) was derived from a bacteria population and introduced into sugarcane. The psEG was targeted to chloroplast, ER, and vacuole with the chloroplast being the highest accumulator near 0.05% total soluble protein. The psEG was driven by a constitutive maize phosphoenolpyruvate carboxylase (*Zm-PEPc)* promoter which expressed psEG in the leaves. Unlike the observation in E1 in maize embryos, no truncation of the psEG protein was detected (Hood et al., 2007; Harrison et al., 2011).

2.4 Cellobiohydrolases in transgenic plants

Heterologous overproduction of cellobiohydrolases (CBHs) has not been explored as much as EG in plants, but CBHs are necessary for complete hydrolysis (Table 2-2). The first CBH evaluated in plants was Cel6B (formerly named E3), from the bacterium *Thermobifida fusca* in the cytosol of alfalfa, tobacco and potato (Ziegelhoffer et al., 1999). While Cel6B was detectable by western blotting and was not deleterious to plant growth, its activity was negligible and the authors declined further characterization (Ziegelhoffer et al., 1999). Although the *Mac* promoter was used for Cel6B relatively low yield was achieved. Reasoning could be the lack of a signal peptide to target to an organelle or the need to codon optimize the proteins which showed improvement in future CBH studies (Hood et al., 2007). In a later study, Cel6B were

targeted to tobacco chloroplasts, and no morphological changes were observed in the host plant (Yu et al., 2007). Cel6B showed had a total soluble protein yield of 0.6-4%. Enzyme activity was measured using crystalline cellulose as the substrate instead of a soluble cellulose derivative substrate (i.e. CMC or MUC) which could explain the low activity reported.

Initial experiments on CBH1 from the fungus *Trichoderma reesei* driven by the constitutive *35S* promoter, yielded 0.11 to 0.82% total soluble protein from tobacco leaves and callus and showed no deleterious effect on plant growth (Dai et al., 1999). Further experiments with CBH1 used the maize embryo promoter, globulin-1, and targeted to the cell walls, ER, and vacuoles for increased accumulation (Hood et al., 2007). The first 40 codons were optimized for improved maize protein synthesis. The combination of tissue, organelle, and codon optimization resulted in up to 16% total soluble protein, which is the highest yield reported for CBH1 (Hood et al., 2007). An issue noted while lines for CBH1 targeted to the ER had enzyme activity, they failed to be present on Western blots. However, cell wall targeted CBH1 was full length. Further evaluation demonstrated the CBH1 was truncated. In future cases, it may be necessary to fully codon optimize GHs to reduce the risk of truncation, especially in cases where the catalytic domain is near the C-terminus of the protein. Field experiments of maize progeny carrying CBH I, showed no yield or growth performance difference compared to wild type counterparts under field conditions (Garda et al., 2015).

Proprietary synthetic CBH1 and CBHII (psCBH1 and psCBHII) genes were generated using multiple bacterial and fungal sources as templates and were transformed into sugar cane

(Harrison et al., 2011). Each protein was targeted to either chloroplast, ER, or vacuole with none reporting any morphological defects. The highest activity was measured from senesced leaves of psCBH1 targeted to the vacuole (Harrison et al., 2011). Both proteins were successfully localized to the leaves by the *Zm-PEPc* promoter and neither of them were truncated (Harrison et al., 2011).

In another study, recombinant synthesis of an endogenous rice CBH (EXG1) resulted in increased glucose release in rice shoots (Nigorikawa et al., 2012). The EXG1 construct utilized its native signal peptide. Some transgenic events were not observably morphologically different than the non-transgenic parent, but three events had deformed, split leaves and extra lacunae. The authors hypothesized that the altered leaf phenotype was from weakened cell walls imbued by the EXG1. Of the 28 EXG1 rice events recovered, 12 were completely sterile and 14 were partially sterile (defined as <40 seeds produced) (Nigorikawa et al., 2012).

2.5 β-glucosidases in transgenic plants

β-glucosidases (BGs) are necessary for degradation of cellobiose into glucose, and various BG genes have been overexpressed in plants (Table 2-3). Heterologous expression of the *T. maritime* BglB enzyme resulted in no altered plant morphology in tobacco or Arabidopsis, and the recombinant protein had the characteristic thermophilic properties (Jung et al., 2010; Jung et al., 2013). Crude extract from BglB tobacco was applied to transgenic rice expressing Cel5A endoglucanase, which resulted in improved saccharification in an expected synergistic fashion (Jung et al., 2010). Over-expression of the endogenous BEG1 rice BG in transgenic rice also

showed no morphological effects, however no enzyme activity was reported (Nigorikawa et al., 2012). The lack of morphological effects on host plants were hypothesized to be the result of predominant cellobiose degradation and not complete cellulose chains (Jung et al., 2010; Nigorikawa et al., 2012). BGs are the least reported on of the essential CWDs and more investigation into their functional role for a complete autohydrolysis system will be required.

2.6 Bacterial and xylanases in transgenic plants

Xylans are the major carbohydrates in the hemicellulose portion of the plant cell wall. Xylanases break down the bonds of β-1,4-xylan reducing them to simpler pentose sugars which can be fermented for ethanol production. Dicots and non-commelinoid monocots have Type I cell walls predominately made up of xyloglucan, while grasses are made up of Type II cell walls which contain high levels of arabinoxylan (Scheller and Ulvskov, 2010). Due to this difference, careful design is needed to ensure effective xylanase usage for the different cell wall types in the host organism. Heterologously produced xylanases in plants have been performed on Arabidopsis, barley, potato, rice, sunflower, and tobacco and have maintained their thermophilic and mild acidic to neutral pH properties (Table 2-4). Overexpression of xylanases have not yielded any plants with observable morphological differences compared to controls.

A codon-optimized *xynA* gene from the fungus *Neocallimastix patriciarum* was expressed in barley under the control of two separate endosperm-specific promoters (*GluB-1* & *Hor2-4*) (Patel et al., 2000). The *GluB-1* version had twice the expression and activity of the other version in mature seed, with no expression observed in leaf, stem, or root tissues. Dried and stored seed

maintained xylanase activity, which would enable biofuel catalyst production or direct feeding to animals. Two *Clostridium thermocellum* xylanases have been produced in transgenic plants. The first was a truncated xylanase (XynZ) produced in tobacco apoplasts, in which thermostability was maintained (Herbers et al., 1995). XynZ was active against multiple xylan substrates at high temperature and slightly acidic pH. The second was XynA produced in rice, which accumulated in the cytoplasm. The native signal peptide was removed, but the catalytic domain was left intact. The recombinant enzyme had native-like activity at neutral pH and thermophilic temperature of 60 °C (Kimura et al., 2003). In these experiments, the enzyme accumulated in seed was active after dry storage. In Arabidopsis, two different xylanases from the fungus *Trichoderma reesei* were used to overproduce high levels of enzyme. The first, XYLII, accumulated in the cytosol, peroxisome, and chloroplast, depending on the experiment. Co-localization into chloroplast and peroxisomes resulted in the highest yield of XYLII (Hyunjong et al., 2006). The second xylanase tested in Arabidopsis was XYNII, which was targeted to either the chloroplast or the cytosol. Chloroplast-targeted XYNII yielded the higher amounts of protein compared with the non-targeted variant. Plant-derived XYNII was found to be comparable to a commercial xylanase in activity under lab conditions (Bae et al., 2008). Finally, *Streptomyces olivaceoviridis* XynB was overproduced in potato cytosol or apoplast. XynB was stable at 60 and 70 °C and subsequent generations of the transgenic potato produced greater amounts of XynB than the T0 generation (Yang et al., 2007).

2.7 Combining glycosyl hydrolase genes in transgenic plants for improved conversion performance

Cellulolytic enzymes have been combined to assess potential synergistic effects by gene stacking in transgenic plants (Table 2-5). Tobacco has been the most often-used host plant, but the first reported host for stacked GHs was barley (Lee et al., 2012; Mahadevan et al., 2011; Fan and Yuan, 2010; Jensen et al., 1996). The first experiments in barley were accomplished using protein fusions. The EII-hybrid construct is a plant codon-optimized fusion of two bacterial enzymes from *Bacillus spp.* glucanase and a *Bacillus spp.* α-amylase for expression in barley. EII-hybrid was expressed during endosperm germination using the high-pI-α-amylase signal peptide (Jensen et al., 1996). In another study, the same EII-hybrid construst was driven by *Hor3-1* endosperm promoter and the Hor3-1 native signal peptide for targeting to developing endosperms (Horvath et al., 2000). Neither EII-hybrid constructs showed any morphological changes to plant growth in either case (Horvath et al., 2000; Jensen et al., 1996). EII-hybrid was active at pH 7.4 and 65 °C (Horvath et al., 2000; Jensen et al., 1996). Another barley transgenic hydrolase named *cel-hyb1* is a fusion of an EG from the fungus *Neocallimastix patriciarum* and a 1,3-1,4-β-glucanase from *Piromyces sp*. (Xue et al., 2003). The Cel-hyb1 transgenic lines showed no observable morphological defects to plant health while maintaining equal enzyme activity for fresh or dried biomass. Cel-hyb1 had a lower temperature activity at 40 °C and slightly more acidic pH (Xue et al., 2003). Both cases demonstrated the use of barley as a functional bioreactor and could be used for production of transgenic cellulases or improved feedstocks for animals (Xue et al., 2003).

A hybridized xylanase-arabinase construct (Xyln-ara) was developed by fusing the catalytic domains of a xylanase from *C. thermocellum* and an arabinofuranosidase from *Geobacillus stearothermophilus* and introduced into tobacco. The fusion of the domains revealed novel enzyme character for having an improved activity at pH 6-9 when compared to the original enzyme. The fused enzymes demonstrated activity on carboxymethyl cellulose (CMC) substrate while the native forms did not. The Xyln-ara hybrid was targeted to the cytosol and showed no effect on plant growth (Fan and Yuan, 2010). Use of catalytic domains is advantageous by reducing the size of transformation constructs which improves multi-gene integration experiments and down-stream analysis (Fan and Yuan, 2010).

From *T. maritima,* the Cel5A endoglucanase and an engineered CBM6-Cel5A fusion hydrolase construct were introduced and compared in transgenic tobacco. CBM6 is a carbohydrate binding module from CBM family 6 of xylanase A in *C. stercorarium*. When the CBM6-Cel5A was first engineered and tested in *E. coli* it showed up to an 18-fold increase in activity over non-fused Cel5A and was selected for tobacco evaluation (Mahadevan et al., 2008). The CBM6-Cel5A was targeted to the chloroplast, apoplast, and cytosol where the chloroplast targeted enzyme had the highest accumulation and no morphological defect was reported (Mahadevan et al., 2011). Autohydrolysis of plant material was performed on biomass and found the enzymes were able to have increased release of glucose in plants with the hybrid and native gene (Mahadevan et al., 2011).

In tobacco, BglB, Cel5A, XylII or Cel6B were connected by the linker from the 2A oligopeptide from the foot-and-mouth disease virus to test a gene stack integration to improve hydrolysis. It was noted that selected hydrolase genes did have a pH dependent effect on hydrolysis. Combination of these enzymes worked well when on CMC substrate, however they showed low activity on filter paper (Lee et al., 2012). A synergistic effect was observed when BglB:Cel5A and Cel6B:Cel5A were co-expressed, but not observed for XylII:Cel5A.

In sunflower, E1 and the xylanase Xyn10A from *Acidothermus cellulolyticus*, were co-expressed to the apoplast and evaluated for biofuel use. Transient expression of both enzymes was achieved using the *CaMV35S* constitutive promoter, however yields were very low compared to previously reported plant expression systems (Jung et al., 2014; Taylor et al., 2008). Other transformation vectors tested using *CMVar* and *TRBO* promoters for both enzymes, however these transformation vectors failed to produce transient transgenic sunflowers biomass. Work in the non-model sunflower demonstrates again the necessity for proper promoter and transformation system in order to achieve meaningful transgene expression.

Future work on gene stacking GHs in plants should combine each enzyme grouping to constitute a complete autohydrolysis system. Selected GHs would need plant codon optimization for improved yield and prevent truncation (Hood et al., 2007). Organelle targeting to either the apoplast or the vacuole would be the best for high protein accumulation (Hood, et al., 2007; Ziegler, et al., 2000). Use of precision genome editing tools could knock out non-essential growth pathways to increase protein yield (Liu et al., 2014). For example, E1 targeted to maize

endosperm cell walls could be mutated to reduce starch storage in favor of storing E1. Adding in CBHs and BGs creates the enzymatic synergy needed for complete autohydrolysis.

2.8 Insect bioprospecting avenues for transgenic plant produced hydrolases

Insects are a relatively recent source for bioprospecting biocatalysts for biofuel production (Carroll and Somerville, 2009; Oppert et al., 2010). Isopteran (termite) digestive tracts are very efficient: 99% of ingested cellulose and 87% of hemicellulose is converted into usable sugars (Ohkuma, 2003). Researchers once believed that insects produced few-to-no endogenous CWD enzymes, but rather relied on those from symbiotic organisms in their guts to break down plant biomass. For some species this is still believed to be the case, however genomic and proteomic analyses have shown that insects produce endogenous enzymes (Tokuda and Watanabe, 2007; Watanabe et al., 1998). For example, termites can survive solely on cellulose, and thus would be a source of cell wall digesting enzymes for bioprospecting (Watanabe and Tokuda, 2010). Heterologous expression of insect cellulases in model organisms have shown them to be potential alternatives to bacterial and fungal-sourced enzymes. Many insect cellulases reported in the literature have temperature optima from 40 to 65 ºC and perform optimally at alkaline pH (Shi et al., 2013; Shi et al., 2011; Willis et al., 2011). High temperature and variable pH characteristics were discovered as researchers compared insect derived hydrolases to microbial hydrolases. These high temperature and variable pH could be from the native structure formed by GHs or from the symbiont ancestors which resided in the insect digestive system.

Insect cellulolytic enzymes fall into the GH families 1, 5, 9, 11, and 45. The majority of insect cellulases discovered are from the Coleoptera (beetle) and Isoptera (termite) taxonomic orders as

defined by Misof et al., (2014). In addition, Blattodea (roaches) Phasmatodea (stick insects), Lepidoptera (butterflies/moths), and Orthoptera (grasshopper/crickets) have demonstrated cellulolytic activity (Oppert et al., 2010; Shelomi et al., 2014). Endogenous and insect endosymbiont derived enzymes have similar characteristics to bacterial and fungal derived enzymes in terms of thermostability and variable pH (Oppert et al., 2010). Insect derived enzymes often have a more alkaline pH range than those from microbial sources, since the insect digestive systems can be up to a pH of 12 in the insect digestive systems (Brune, 2014; Dow, 1992).

The digestive systems of insects are broken into three distinct sections of foregut, midgut, and hindgut. The foregut includes the mouth, salivary glands, and the most anterior chamber of the digestive tract. Cellulolytic enzymes have been discovered in the salivary glands which along with the maceration from the mandibles begins initial degradation. The fore- and midgut contain mostly endogenous cellulases (Brune, 2014; Slaytor, 1992). Malpighian tubules, when present, separate the mid and hindgut. The majority of microbial symbionts have been discovered in the hindgut. Specifically, protist and bacterial cellulases have been isolated from the hindgut of Coleoptera, Isoptera, and Orthoptera insect orders (Brune, 2014; Watanabe et al., 1998). Both symbiont and endogenous insect GH systems have the potential for use as novel biocatalysts.

A major impediment to identifying and cloning insect CWD enzyme-coding genes is the lack of sequenced and assembled genomes among insect species, especially those that have a diet consisting of plant cell walls. *Apis melliferia, Drosophila melanogaster, Drosophila*

pseudoobscura, Tribolium castaneum, Nasonia spp., Acyrthosiphon pisum, and *Heliconius spp.* are taxa with the most complete genome data (Heliconius Genome, 2012; International Aphid Genomics, 2010; Richards et al., 2005; Tribolium Genome Sequencing et al., 2008; Weinstock et al., 2006; Werren et al., 2010). Of these, only *T. castaneum* and *A. pisum* and, *Heliconius spp*., are herbivores. There are efforts, Such as the "i5K" to sequence and assemble 5000 insect genomes, of which 28 pilot studies are underway and new phylogenetic trees are being designed (Behura, 2015; Misof et al., 2014; Robinson et al., 2011). To advance insect CWD gene discovery, herbivorous insects and their metagenomes should be sequenced to advance the field beyond its present state (Table 2-6).

Termites are the archetypical wood-consuming insects. The first endogenous insect cellulase was discovered from the genome of the termite, *Reticulitermes speratus* (Watanabe et al., 1998). Since then many other termite-derived hydrolases have been discovered and even evaluated for heterologous expression. Isoptera families are split into the lower termites (Mastotermitidae, Kalotermitidae, Hodotermitidae, Termopsidae, Rhinotermitidae, and Serritermitidae) and higher termites (Termitidae). The distinction between these families is based on presence of flagellated protist symbionts in the hindgut (lower termites) or their absence (higher termites) (Brune, 2014; Ni and Tokuda, 2013). The "lower vs higher" termite digestion distinction is the basis of evolutionary importance being that the higher termites evolved later (Ni and Tokuda, 2013; Slaytor, 1992; Watanabe and Tokuda, 2001). A prevailing theory on this change is due to ecological changes which caused Isoptera ancestors' loss of protists symbionts, change of habitat, and/or food source and possible reacquisition of different cellulase-harboring symbionts

(Ni and Tokuda, 2013; Slaytor, 1992; Watanabe and Tokuda, 2001). In either case of both higher and lower termites have provided cellulolytic enzymes which could be utilized for autohydrolysis plants.

GHs isolated from two of the lower termite families, Kalotermitidae and Rhinotermitidiae, have been evaluated in heterologous systems. From Kalotermitide, a *Neotermes koshunensis* BG (NkBG) was produced in *E. coli* and in *Aspergillus oryzae*. NkBG was functional in both systems, however it had a higher activity from *E. coli* (156.7 U/mg) compared to *A. oryzae* (12.4 U/mg). NkBG produced in *A. oryzae* had a unique property for maintaining 100% activity in the presence of 0.6 M glucose; most BG enzymes are inhibited by excess glucose (Uchima et al., 2011). Rhinotermitidae enzymes have been heterologously expressed from *Coptotermes formosanus, Reticulitermes flavipes, Reticulitermes santonensis*, and *Reticulitermes speratus*. From *C. formosanus* two EGs (CFP-EG1and CfEG5) were evaluated in *E. coli*. CFP-EG1 is from a protist symbiont and had a thermophilic activity at 70 °C with a CMC substrate activity of 105 U/mg (Inoue et al., 2005). Conversely, CfEG5 had three fold higher CMC activity at 325 U/mg, but was less thermally active with an optimum of 43 °C (Zhang et al., 2011a). A BG from *C. formosanus* was expressed in *E. coli* and demonstrated an activity of 462.6 U/mg on cellobiose (Zhang et al., 2012a). From *R. flavipes*, an EG (Cell-1) and a BG (RfBGluc-1) were expressed using the baculovirus expression system in which *Trichoplusia ni* larvae served as bioreactors. Hemolymph from baculovirus-injected larvae were extracted and analyzed for enzyme activity (Scharf et al., 2010). Cell-1 CMC activity of 1.4 U/mg was low while RfBGluc-1 was considerably more effective on cellobiose at 638 U/mg. An opposite trend was seen from

an *R. santonensis* BG which had a low activity of 0.441 U/mg against *p*-nitrophenol-β-D-glucopyranoside (pNPG) subtstrate (Matteotti et al., 2011). Possibly this was due to being produced in *E. coli* instead of *T. ni*. Also from *R. santonensis* is the xylanase, mXylB8, which has the highest reported beechwood xylan activity at 1837 U/mg for an insect derived xylanase with a mildly acidic range of pH 5.0 and temperature of 55 °C (Matteotti et al., 2012). From *R. speratus* two EGs (RsEG and RsSymEG1) were produced in *A. oryzae* both of which were active at 45 °C and a CMC activity of 1200 and 605 U/mg, respectively (Hirayama et al., 2010; Todaka et al., 2010). RsEG is an endogenous cellulase and RsSymEG1 is from a protist symbiont. It is interesting to note that the endogenous RsEG is roughly double the activity of the symbiont.

The higher termite family, Termitidae, has also provided useable GHs for plant biocatalysts engineering. The BG, bg1-gs1from *Globitermes sulphureus*, was produced in *E. coli* and demonstrated a thermophilic activity at 90 °C producing 110 U/mg from pNPG substrate (Wang et al., 2012). While not having as high a thermal stability, the MbmgBG1 from *Macrotermes barneyi*, had a higher activity of 206 U/mg on cellobiose with a more acidic pH of 5.0 at temperature of 45 °C (Wu et al., 2012). Another thermophilic insect derived BG is the G1mgNTBG from *Nasutitermes takasagoensis*, while optimum temperature is 65 °C, the activity was the lowest, 5.83 U/mg, when produced by *Pichia pastoris* (Uchima et al., 2012). Also from *N. takasagoensis*, is the EG, NtEG, which when produced in *A. oryzae* produced the highest CMC activity at 1392 U/mg at a slightly acidic pH of 6.0 at 65 °C (Hirayama et al., 2010). A xylanase from *Macrotermes annandalei*, was produced in *E. coli* and evaluated by digestion of beechwood xylan to produce 733 U/mg at a near neutral 7.5 pH and 55 °C (Liu et al., 2011).

Both higher and lower Isopteran species have proven to be a reliable source of GHs, which should be evaluated for their efficacy in plant based expression systems.

Coleopteran derived GHs are found predominantly in the larval stages. The wood-boring *Apriona germari* has provided three different EGs all with heterologous activity from Sf9 insect cells. When using CMC as the substrate, AgEG-I, AgEG-II, and AgEG-III produced activities of 992, 812, 1037 U/mg, respectively (Lee et al., 2004; Lee et al., 2005; Wei et al., 2006). A baculovirus transformation system into *Bombyx mori* larvae was used to heterologously express two other wood-boring beetle cellulases. Hemolymph extracted from transformed larvae carried the EGs when tested on CMC produced activity levels of 927 U/mg for *Batocera horsfieldi* EG and 319.22 ± 9.3 U/mg for *Anoplophora malasiaca* EG (Chang et al., 2012; Xia et al., 2013). Larvae from *Diabrotica virgifera virgifera* were examined as a biocatalyst source due to their diet of maize stems which could be used as a biofuel source. An EG from *D. virgifera* was isolated and produced in *E. coli* with a CMC activity of 0.166 U/mg (Valencia et al., 2013). Unlike the previous Coleoptera species, *Tribolium castaneum*, provided TcEG-1 which has a lower activity compared to other beetle species, but was active at very alkaline pHs (pH 9-11) in both S2 insect cells and yeast (Shirley et al., 2014; Willis et al., 2011). The Coleopteran cellulases were active at 45-50 °C similarly to that of mild-thermophilic microbial cellulases.

Two additional insect orders have received attention for their CWD enzymes: Lepidoptera and Orthoptera. *Spodoptera frugiperda'*s larval stage was used as the source of a β-glucosidase BG activity of 2.4 $mM^{-1}s^{-1}$ pNPG substrate and was effective when produced in *E. coli* (Marana et

al., 2004). Orthopteran species are known for their mandibular action for chewing through fibrous plant tissue. The symbiont *Klebsiella* sp. in the gut from an undisclosed grasshopper species was found to have BG with similar activity of 2.61 ± 0.75 $mM^{-1}s^{-1}$ on pNPG to that of *S. frugiperda* (Shi et al., 2011). From *Teleogryllus emma*, an EG from GH family 9 was produced in Sf9 cells showing an activity rate of 3118.4 (U/mg) against CMC (Kim et al., 2008).

Insect GHs have not yet been synthesized in transgenic plants. Under heterologous expression in microbial systems, the majority of insect derived GHs demonstrate similar thermal and acidic pH ranges. A few notable examples from Coleoptera have lower activity, but a uniquely alkaline pH range. Attributes, such as high alkaline affinity, learned from insect derived GHs could be utilized in site-directed mutagenesis studies to enhance non-insect biocatalyst research. Overall plant produced insect GHs are another viable avenue for an autohydrolytic plant systems.

2.9 Strategies for engineering autohydrolysis in plants

Constitutive gain-of-function hydrolase activity could have negative phenotypic effects in transgenic plants. While these negative growth effects seldom have been observed in plants harboring GHs with high temperature optima, we envisage host off-effects as a potential issue in heterologous production systems for non-thermophilic enzymes. Use of inducible promoters to attenuate hydrolase activity can reduce off-target effects and decrease chances of lethality (Schena et al., 1991;Taylor et al., 2008). Inducible systems could be utilized at several life cycle/growth stages to specifically induce the hydrolase activity to detect onset of different genes and pathways involved in cell wall biosynthesis. One such inducible approach has been used to

control CWD transgene expression in tobacco. An ethanol inducible promoter demonstrated effective cellulase production while preventing dwarfism when compared to constitutively produced cellulase in tobacco plants (Klose et al., 2013).

Protein modification can be used on GHs to control for off-target effects and increase protein yield. Sequestration of proteins by signal peptides to organelles, such as vacuoles, ER, or plastids improves production rate of transgenic proteins, but also isolates them from their intended substrates (Taylor et al., 2008; Hood et al., 2007). The use of hyperthermic proteins, where enzyme activation temperature is above that required for plant growth are useful during biorefinery steps to prevent undesired effects (Mir et al., 2014). A unique protein sequence known as an intein, can be inserted into the portions of enzymes and prevents proper protein folding. Inteins are heat labile and have been proven to successfully prevent undesired enzymatic action until placed in high temperature which induce intein excision (Shen et al., 2012).

Plants are attractive bioreactors for many recombinant proteins because their post-translational modification systems are analogous to other eukaryotes (Fischer et al., 2012; Hellwig et al., 2004; Zhang, 2015). Plant cell culture systems could also speed the screening process for functional GH production *in planta* prior to generation of time consuming whole transgenic plants (Hellwig et al., 2004; Shen et al., 2013). A multi-well system using a plant cell system for rapid GH stacking of multiple constructs would speed the screening process of GHs.

A current research problem is the variety of substrates used to measure activity and efficiency. Filter paper is reformed cellulose, but the different densities and size of the filter paper disc used during assays have effects on enzyme activity (Xiao et al., 2004). Carboxymethylcellulose is the most common substrate used because of its solubility in liquids, which is caused by methylating cellulose (Reese et al., 1950). Avicel (powdered crystalline cellulose) is a close relative to cellulose, but can be purchased at crystalline particle size making comparisons with the soluble form difficult (Yeh et al., 2010). Both of these (and other cellulose derivative) substrates can also be affixed with fluorophores which are activated upon hydrolysis (Megazyme Wicklow, Ireland). The units differ based on the method used to measure, some do a straight absorbance compared to a control, a reduction in viscosity of a solution, use of a standard curve, or percent of a commercial cellulase standard (Taylor et al., 2008; Willis et al., 2010).

An important decision to be made when developing commercial-ready autohydrolytic plants is the host plant. This issue is important because C4 grasses (leading biofuel crops) have different cell wall composition and genetic mechanisms than C3 plants. While current monocot model species, such as rice (*Oryza sativa*) and Brachypodium (*Brachypodium distachyon*), have extensive genomics toolsets; e.g., completed genomes and activation tagging systems; both are C3 grasses that are distantly related to C4 grasses (Ito et al., 2005; Jeon and An, 2001; Matsumoto et al., 2005; Vogel et al., 2010; Vogel et al., 2009). Thus, while rice and Brachypodium have proven to be powerful tools, they are suboptimal to dissect C4 grass metabolism and cell wall biosynthesis (Dal'Molin et al., 2010). This latter point is a crucial consideration for testing autohydrolytic biofuel crop lines. C4 grasses (e.g., maize, sorghum, and

switchgrass) possess divergent expression patterns of cell wall-related genes, and indeed likely use different genes, compared to C3 grasses. The specialized vascular anatomy of C4 plants relies on similar genes to C3 grasses, however, their regulation and expression differ, resulting in different structures; their secondary cell walls are simply not well modeled by C3 grasses (Nelson, 2011). The brown midrib (BMR) mutant phenotype in maize, sorghum, and pearl millet is a classic lignin alteration in cell walls, but no C3 plants have been reported with this cell wall phenotype. Putative BMR ortholog mutants in C3 plants seem to have a different basis; BMR rice mutant orthologs are known, but present different phenotypes if altered (Sattler et al., 2010). A digestibility assay comparing two C3 and two C4 grasses native to Australia showed that C3 grass leaves had more facile enzymatic hydrolysis compared to that of the C4 grasses, which was attributed to C3/C4 differences in leaf anatomy (Wilson and Hacker, 1987). *In vitro* evaluations using purified cellulose and lignin polymer models demonstrated that for lignin to affect cellulose saccharification it must be specifically cross linked to cellulose (Jung et al., 2011). Grass arabinoxylans cross link to lignin by ether bonds using ferulate and diferulate molecules (Burr and Fry, 2009). Dicot cell walls lack arabinoxylans for cross linking to lignin and therefore are inadequate to use as an improved digestibility model for C4 grass crops. Alternatively, dicot cell walls are primarily comprise xyloglucan hemicellulose for cross linking to lignin (Damm, 2016; Furtado et al., 2014; Scheller and Ulvskov, 2010). Utilization of C4 grass species is more appropriate for evaluation of CWD enzymes if the target feedstock is a C4 grass. The applied approach for evaluation of novel autohydrolytic lines should be used with recently improved transformation strategies of many C4 crops (e.g., maize, switchgrass,

sorghum) to reduce variability to real world marketable biomass (Brutnell et al., 2015; Brutnell et al., 2010; Ishida et al., 2007; Lambertz et al., 2014; Li and Qu, 2011; Liu and Godwin, 2012).

2.10 Conclusion

There are many economic pinch points in the production of cellulosic biofuels. Enzyme cost is one of them and is not trivial. Using recombinant systems, including the plant feedstock itself, is likely necessary to produce advanced biofuels in a sustainable manner (Gressel, 2008). Plants, including a few species that are relevant proxies for cellulosic feedstocks, have been used as hosts for overexpression of microbial GH genes. We envisage insect-derived GHs playing a role in reducing cell wall recalcitrance in transgenic plants owing to their high diversity of function and eukaryotic origin. Utilization of multiple enzyme classes should be used in concert in feedstocks, coupled with appropriate subcellular targeting as well as spatial and temporal control of synthesis. While there is much research needed to produce a commercial cellulosic feedstock with autocatalytic properties, field experiments with E1-producing transgenic maize demonstrates that research in this area should be worthwhile. The combination of improved biomass feedstocks and enzyme technologies is a step toward a renewable plant-based fuel system.

2.11 Acknowledgments

This work was supported by funding from the BioEnergy Science Center (DE-PS02-06ER64304). The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. We also thank the University of Tennessee, Ivan Racheff Endowment, and USDA HATCH funds for supporting the research program of CNS.

2.12 References

Bae, H.J., Kim, H.J. and Kim, Y.S. (2008) Production of a recombinant xylanase in plants and its potential for pulp biobleaching applications. *Bioresour. Technol.* 99, 3513-3519.

Behura, S.K. (2015). Insect phylogenomics. *Insect Mol. Biol.* 24**,** 403-411. doi: 10.1111/imb.12174.

Bhalla, A., Bansal, N., Kumar, S., Bischoff, K.M. and Sani, R.K. (2013) Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresour. Technol.* 128, 751-759.

Bhat, M.K. (2000) Cellulases and related enzymes in biotechnology. *Biotechnol. Adv.* 18, 355-383.

Biswas, G.C.G., Ransom, C. and Sticklen, M. (2006) Expression of biologically active *Acidothermus cellulolyticus* endoglucanase in transgenic maize plants. *Plant Sci.* 171, 617-623.

Borkhardt, B., Harholt, J., Ulvskov, P., Ahring, B.K., Jorgensen, B. and Brinch-Pedersen, H. (2010) Autohydrolysis of plant xylans by apoplastic expression of thermophilic bacterial endo-xylanases. *Plant Biotechnol. J.* **8**, 363-374. doi: 10.1111/j.1467-7652.2010.00506.x

Boyer, L.J., Vega, J.L., Klasson, K.T., Clausen, E.C. and Gaddy, J.L. (1992) The effects of furfural on ethanol-production by *Saccharomyces-cerevisiae* in batch culture. *Biomass Bioenerg.* 3, 41-48.

Bruce, A.M. and Palfreyman, J.W. (1998) *Forest Products Biotechnology*. London ; Bristol, PA:Taylor & Francis.

Brune, A. (2014) Symbiotic digestion of lignocellulose in termite guts. *Nat. Rev. Microbiol.* 12, 168-180.

Brunecky, R., Selig, M.J., Vinzant, T.B., Himmel, M.E., Lee, D., Blaylock, M.J. et al (2011) *In planta* expression of *A. cellulolyticus* Cel5A endocellulase reduces cell wall recalcitrance in tobacco and maize. *Biotechnol. Biofuels* 4, 1-11. doi: 10.1186/1754-6834-4-1

Brutnell, T.P., Bennetzen, J.L. and Vogel, J.P. (2015) *Brachypodium distachyon* and *Setaria viridis*: Model genetic systems for the grasses. *Annu. Rev. Plant Biol.* 66, 465-485.

Brutnell, T.P., Wang, L., Swartwood, K., Goldschmidt, A., Jackson, D., Zhu, X.G., Kellogg, E. and Van Eck, J. (2010) *Setaria viridis*: a model for C4 photosynthesis. *Plant Cell* 22, 2537-2544.

Burr, S.J. and Fry, S.C. (2009) Feruloylated arabinoxylans are oxidatively cross-linked by extracellular maize peroxidase but not by horseradish peroxidase. *Mol. Plant* 2, 883-892.

Carroll, A. and Somerville, C. (2009) Cellulosic biofuels. *Annu. Rev. Plant Biol.* 60, 165-182.

Chang, C.J., Wu, C.P., Lu, S.C., Chao, A.L., Ho, T.H., Yu, S.M. and Chao, Y.C. (2012) A novel exo-cellulase from white spotted longhorn beetle (*Anoplophora malasiaca*). *Insect Biochem. Mol. Bio.* 42, 629-636.

Chou, H.L., Dai, Z., Hsieh, C.W. and Ku, M.S. (2011) High level expression of *Acidothermus cellulolyticus* beta-1, 4-endoglucanase in transgenic rice enhances the hydrolysis of its straw by cultured cow gastric fluid. *Biotechnol. Biofuels* 4, 58.

Cosgrove, D.J. (2005) Growth of the plant cell wall. *Nat. Rev. Mol. Cell. Biol.* 6, 850-861.

Dai, Z., Hooker, B.S., Quesenberry, R.D., and Gao, J. (1999). Expression of *Trichoderma reesei* exo-cellobiohydrolase I in transgenic tobacco leaves and calli. *Appl. Biochem. Biotechnol.* 77-79**,** 689-699.

Dai, Z., Hooker, B.S., Anderson, D.B. and Thomas, S.R. (2000a) Expression of *Acidothermus cellulolyticu*s endoglucanase E1 in transgenic tobacco: biochemical characteristics and physiological effects. *Transgenic Res.* 9, 43-54.

Dai, Z.Y., Hooker, B.S., Anderson, D.B. and Thomas, S.R. (2000b) Improved plant-based production of E1 endoglucanase using potato: expression optimization and tissue targeting. *Mol. Breeding* 6, 277-285.

Dai, Z., Hooker, B.S., Quesenberry, R.D. and Thomas, S.R. (2005) Optimization of *Acidothermus cellulolyticus* endoglucanase (E1) production in transgenic tobacco plants by transcriptional, post-transcription and post-translational modification. *Transgenic Res.* 14, 627-643.

Dal'Molin, C.G., Quek, L.E., Palfreyman, R.W., Brumbley, S.M. and Nielsen, L.K. (2010) C4GEM, a genome-scale metabolic model to study C4 plant metabolism. *Plant Physiol.* 154, 1871-1885.

Damm, T., Commandeur, U., Fischer, R., Usadel, B., Klose, H. (2016) Improving the utilization of lignocellulosic biomass by polysaccharide modification. *Process Biochem.* 51, 288-296.

Dow, J.A. (1992) pH Gradients in Lepidopteran midgut. *J. Exp. Biol.* 172, 355-375.

Fan, Z. and Yuan, L. (2010) Production of multifunctional chimaeric enzymes in plants: a promising approach for degrading plant cell wall from within. *Plant Biotechnol. J.* 8, 308-315.

Fischer, R., Schillberg, S., Hellwig, S., Twyman, R.M. and Drossard, J. (2012) GMP issues for recombinant plant-derived pharmaceutical proteins. *Biotechnol. Adv.* 30, 434-439.

Furtado, A., Lupoi, J.S., Hoang, N.V., Healey, A., Singh, S., Simmons, B.A. and Henry, R.J. (2014) Modifying plants for biofuel and biomaterial production. *Plant Biotechnol. J.* 12, 1246-1258.

Garda, M., Devaiah, S.P., Vicuna Requesens, D., Chang, Y.K., Dabul, A., Hanson, C., Hood, K.R. and Hood, E.E. (2015) Assessment of field-grown cellulase-expressing corn. *Transgenic Res.* 24, 185-198.

Gressel, J. (2008) Transgenics are imperative for biofuel crops. *Plant Sci.* 174, 246-263.

Harrison, M.D., Geijskes, J., Coleman, H.D., Shand, K., et al (2011) Accumulation of recombinant cellobiohydrolase and endoglucanase in the leaves of mature transgenic sugar cane. *Plant Biotechnol. J.* 9, 884-896.

Heliconius Genome Consortium. (2012) Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487, 94-98.

Hellwig, S., Drossard, J., Twyman, R.M. and Fischer, R. (2004) Plant cell cultures for the production of recombinant proteins. *Nat. Biotechnol.* 22, 1415-1422.

Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280 (Pt 2), 309-316.

Henrissat, B. and Bairoch, A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 293 (Pt 3), 781-788.

Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316 (Pt 2), 695-696.

Herbers, K., Wilke, I. and Sonnewald, U. (1995) A thermostable xylanase from *Clostridium thermocellum* expressed at high-levels in the apoplast of transgenic tobacco has no detrimental effects and is easily purified. *Bio. Technol.* 13, 63-66.

Hirayama, K., Watanabe, H., Tokuda, G., Kitamoto, K. and Arioka, M. (2010) Purification and characterization of termite endogenous beta-1,4-endoglucanases produced in *Aspergillus oryzae*. *Biosci. Biotechnol. Biochem.* 74, 1680-1686.

Hood, E.E., Devaiah, S.P., Fake, G., Egelkrout, E., Teoh, K.T., Requesens, D.V., Hayden, C., Hood, K.R., Pappu, K.M., Carroll, J. and Howard, J.A. (2012) Manipulating corn germplasm to increase recombinant protein accumulation. *Plant Biotechnol. J.* 10, 20-30.

Hood, E.E., Love, R., Lane, J., Bray, J., Clough, R., Pappu, et al (2007) Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed. *Plant Biotechnol. J.* 5, 709-719.

Horvath, H., Huang, J., Wong, O., Kohl, E., Okita, T., Kannangara, et al. (2000) The production of recombinant proteins in transgenic barley grains. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1914-1919.

Hyunjong, B., Lee, D.S. and Hwang, I. (2006) Dual targeting of xylanase to chloroplasts and peroxisomes as a means to increase protein accumulation in plant cells. *J. Exper. Bot.* 57, 161-169.

Inoue, T., Moriya, S., Ohkuma, M. and Kudo, T. (2005) Molecular cloning and characterization of a cellulase gene from a symbiotic protist of the lower termite, *Coptotermes formosanus*. *Gene* 349, 67-75.

International Aphid Genomics Consortium. (2010) Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol.* 8, e1000313.

Ishida, Y., Hiei, Y. and Komari, T. (2007) *Agrobacterium*-mediated transformation of maize. *Nature Protoc.* 2, 1614-1621.

Ito, Y., Arikawa, K., Antonio, B.A., Ohta, I., Naito, S., Mukai, Y., et al (2005) Rice Annotation Database (RAD): a contig-oriented database for map-based rice genomics. *Nucleic Acids Res.* 33, D651-D655.

Jensen, L.G., Olsen, O., Kops, O., Wolf, N., Thomsen, K.K. and von Wettstein, D. (1996) Transgenic barley expressing a protein-engineered, thermostable (1,3-1,4)-beta-glucanase during germination. *Proc. Natl. Acad. Sci. U.S.A.* 93, 3487-3491.

Jeon, J.S. and An, G.H. (2001) Gene tagging in rice: a high throughput system for functional genomics. *Plant Sci.* 161, 211-219.

Jin, R., Richter, S., Zhong, R. and Lamppa, G.K. (2003) Expression and import of an active cellulase from a thermophilic bacterium into the chloroplast both in vitro and in vivo. *Plant Mol. Biol.* 51, 493-507.

Jung, H.G., Mertens, D.R. and Phillips, R.L. (2011) Effect of reduced ferulate-mediated lignin/arabinoxylan cross-linking in corn silage on feed intake, digestibility, and milk production. *J. Dairy Sci.* 94, 5124-5137.

Jung, S., Kim, S., Bae, H., Lim, H.S. and Bae, H.J. (2010) Expression of thermostable bacterial beta-glucosidase (BglB) in transgenic tobacco plants. *Bioresour. Technol.* 101, 7155-7161.

Jung, S., Lee, D.S., Kim, Y.O., Joshi, C.P. and Bae, H.J. (2013) Improved recombinant cellulase expression in chloroplast of tobacco through promoter engineering and 5' amplification promoting sequence. *Plant Mol. Biol.* 83, 317-328.

Jung, S.K., Lindenmuth, B.E., McDonald, K.A., Hwang, H., Bui, M.Q., Falk, B.W., Uratsu, S.L., Phu, M.L. and Dandekar, A.M. (2014) *Agrobacterium tumefaciens* mediated transient expression of plant cell wall-degrading enzymes in detached sunflower leaves. *Biotechnol. Prog.* 30, 905-915.

Kawazu, T., Sun, J.L., Shibata, M., Kimura, T., Sakka, K. and Ohmiya, K. (1999) Expression of a bacterial endoglucanase gene in tobacco increases digestibility of its cell wall fibers. *J. Biosci. Bioeng.* 88, 421-425.

Kim, N., Choo, Y.M., Lee, K.S., Hong, S.J., Seol, K.Y., Je, Y.H., Sohn, H.D. and Jin, B.R. (2008) Molecular cloning and characterization of a glycosyl hydrolase family 9 cellulase distributed throughout the digestive tract of the cricket *Teleogryllus emma*. Comp. *Biochem. Physiol. B, Biochem. Mol. Biol.* 150, 368-376.

Kimura, T., Mizutani, T., Tanaka, T., Koyama, T., Sakka, K. and Ohmiya, K. (2003) Molecular breeding of transgenic rice expressing a xylanase domain of the xynA gene from *Clostridium thermocellum*. *Appl. Microbiol. Biotechnol.* 62, 374-379.

Klein-Marcuschamer, D., Oleskowicz-Popiel, P., Simmons, B.A. and Blanch, H.W. (2012) The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnol. Bioeng.* 109, 1083-1087.

Klose, H., Roder, J., Girfoglio, M., Fischer, R. and Commandeur, U. (2012) Hyperthermophilic endoglucanase for *in planta* lignocellulose conversion. *Biotechnol. Biofuels* 5, 63.

Klose, H., Gunl, M., Usadel, B., Fischer, R. and Commandeur, U. (2013) Ethanol inducible expression of a mesophilic cellulase avoids adverse effects on plant development. *Biotechnol. Biofuels* 6, 53.

Klose, H., Gunl, M., Usadel, B., Fischer, R. and Commandeur, U. (2015) Cell wall modification in tobacco by differential targeting of recombinant endoglucanase from *Trichoderma reesei*. *BMC Plant Biol.* 15, 54-65.

Kostylev, M. and Wilson, D. (2012) Synergistic interactions in cellulose hydrolysis. *Biofuels* 3, 61-70.

Lambertz, C., Garvey, M., Klinger, J., Heesel, D., Klose, H., Fischer, R. and Commandeur, U. (2014) Challenges and advances in the heterologous expression of cellulolytic enzymes: a review. *Biotechnol. Biofuels* 7, 135-150.

Lee, D.S., Lee, K.H., Jung, S., Jo, E.J., Han, K.H. and Bae, H.J. (2012) Synergistic effects of 2A-mediated polyproteins on the production of lignocellulose degradation enzymes in tobacco plants. *J. Exp. Bot.* 63, 4797-4810.

Lee, S.J., Kim, S.R., Yoon, H.J., Kim, I., Lee, K.S., Je, Y.H., Lee, S.M., Seo, S.J., Dae Sohn, H. and Jin, B.R. (2004) cDNA cloning, expression, and enzymatic activity of a cellulase

from the mulberry longicorn beetle, *Apriona germari*. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol***.** 139, 107-116.

Lee, S.J., Lee, K.S., Kim, S.R., Gui, Z.Z., Kim, Y.S., Yoon, H.J., Kim, I., Kang, P.D., Sohn, H.D. and Jin, B.R. (2005) A novel cellulase gene from the mulberry longicorn beetle, *Apriona germari*: gene structure, expression, and enzymatic activity. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.* 140, 551-560.

Li, R.Y. and Qu, R.D. (2011) High throughput *Agrobacterium*-mediated switchgrass transformation. *Biomass. Bioenerg.* 35, 1046-1054.

Liu, G. and Godwin, I.D. (2012) Highly efficient sorghum transformation. *Plant Cell Rep.* 31, 999-1007.

Liu, N., Yan, X., Zhang, M., Xie, L., Wang, Q., Huang, Y., Zhou, X., Wang, S. and Zhou, Z. (2011) Microbiome of fungus-growing termites: a new reservoir for lignocellulase genes. *Appl. Environ. Microbiol.* 77, 48-56.

Liu, W.S., Yuan, J.S., and Stewart, C.N. Jr. (2014). Advanced genetic tools for plant biotechnology. *Nat. Rev. Genet.* 15, 781-973. doi: 10.1038/nrg3583.

Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M. and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490-D495.

Lopez-Casado, G., Urbanowicz, B.R., Damasceno, C.M. and Rose, J.K. (2008) Plant glycosyl hydrolases and biofuels: a natural marriage. *Curr. Opin. Plant Biol.* 11, 329-337.

Mahadevan, S.A., Wi, S.G., Kim, Y.O., Lee, K.H. and Bae, H.J. (2011) In planta differential targeting analysis of *Thermotoga maritima* Cel5A and CBM6-engineered Cel5A for autohydrolysis. *Transgenic Res.* 20, 877-886.

Mahadevan, S.A., Wi, S.G., Lee, D.S. and Bae, H.J. (2008) Site-directed mutagenesis and CBM engineering of Cel5A (*Thermotoga maritima*). *FEMS Microbiol. Lett.* 287, 205-211.

Marana, S.R., Andrade, E.H., Ferreira, C. and Terra, W.R. (2004) Investigation of the substrate specificity of a beta-glycosidase from *Spodoptera frugiperda* using site-directed mutagenesis and bioenergetics analysis. *Eur. J. Biochem.* 271, 4169-4177.

Matsumoto, T., Wu, J.Z., Kanamori, H., Katayose, Y., Fujisawa, M., Namiki, N., et al (2005) The map-based sequence of the rice genome. *Nature* 436, 793-800.

Matteotti, C., Bauwens, J., Brasseur, C., Tarayre, C., Thonart, P., Destain, J., et al (2012) Identification and characterization of a new xylanase from Gram-positive bacteria isolated from termite gut (*Reticulitermes santonensis*). *Protein Expr. Purif.* 83, 117-127.

Matteotti, C., Haubruge, E., Thonart, P., Francis, F., De Pauw, E., Portetelle, D. et al (2011) Characterization of a new beta-glucosidase/beta-xylosidase from the gut microbiota of the termite (*Reticulitermes santonensis*). *FEMS Microbiol. Lett.* 314, 147-157.

Mir, B.A., Mewalal, R., Mizrachi, E., Myburg, A.A. and Cowan, D.A. (2014) Recombinant hyperthermophilic enzyme expression in plants: a novel approach for lignocellulose digestion. *Trends Biotechnol.* 32, 281-289.

Misof, B., Liu, S.L., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., et al. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346**,** 763-767. doi: 10.1126/science.1257570.

Nelson, T. (2011) The grass leaf developmental gradient as a platform for a systems understanding of the anatomical specialization of C-4 leaves. *J. Exp. Bot.* 62, 3039-3048.

Ni, J. and Tokuda, G. (2013) Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. *Biotechnol. Adv.* 31, 838-850.

Ni, J.F., Tokuda, G., Takehara, M. and Watanabe, H. (2007) Heterologous expression and enzymatic characterization of beta-glucosidase from the drywood-eating termite, Neotermes koshunensis. *Appl. Entomol. Zool.* **42**, 457-463. doi: 10.1303/aez.2007.457

Nigorikawa, M., Watanabe, A., Furukawa, K., Sonoki, T. and Ito, Y. (2012) Enhanced saccharification of rice straw by overexpression of rice exo-glucanase. *Rice* 5, 14.

Ohkuma, M. (2003) Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Appl. Microbiol. Biotechnol.* 61, 1-9.

Oppert, C., Klingeman, W.E., Willis, J.D., Oppert, B. and Jurat-Fuentes, J.L. (2010) Prospecting for cellulolytic activity in insect digestive fluids. *Comp. Biochem Physiol. B. Biochem. Mol. Biol* 155, 145-154.

Oraby, H., Venkatesh, B., Dale, B., Ahmad, R., Ransom, C., Oehmke, J. et al (2007) Enhanced conversion of plant biomass into glucose using transgenic rice-produced endoglucanase for cellulosic ethanol. *Transgenic Res.* 16, 739-749.

Patel, M., Johnson, J.S., Brettell, R.I.S., Jacobsen, J. and Xue, G.P. (2000) Transgenic barley expressing a fungal xylanase gene in the endosperm of the developing grains. *Mol. Breed.* 6, 113-123.

Ransom, C., Balan, V., Biswas, G., Dale, B., Crockett, E. and Sticklen, M. (2007) Heterologous *Acidothermus cellulolyticus* 1,4-beta-endoglucanase E1 produced within the corn biomass converts corn stover into glucose. *Appl. Biochem. Biotechnol.* 137-140, 207-219.

Reese, E.T., Siu, R.G.H. and Levinson, H.S. (1950) The biological degradation of soluble cellulose derivatives and Its relationship to the mechanism of cellulose hydrolysis. *J. Bacteriol.* 59, 485-497.

Richards, S., Liu, Y., Bettencourt, B.R., Hradecky, P., Letovsky, S., Nielsen, R., et al., (2005) Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution. *Genome Res.* 15, 1-18.

Robinson, G.E., Hackett, K.J., Purcell-Miramontes, M., Brown, S.J., Evans, J.D., Goldsmith, M.R., et al (2011) Creating a buzz about insect genomes. *Science* 331, 1386-1386.

Saha, B.C. (2003) Hemicellulose bioconversion. *J. Ind. Microbiol. Biotechnol.* 30, 279-291.

Sattler, S.E., Funnell-Harris, D.L. and Pedersen, J.F. (2010) Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. *Plant Sci.* 178, 229-238.

Scharf, M.E., Kovaleva, E.S., Jadhao, S., Campbell, J.H., Buchman, G.W. and Boucias, D.G. (2010) Functional and translational analyses of a beta-glucosidase gene (glycosyl hydrolase family 1) isolated from the gut of the lower termite *Reticulitermes flavipes*. *Insect Biochem. Mol. Biol.* 40, 611-620.

Scheller, H.V. and Ulvskov, P. (2010) Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263-289.

Schena, M., Lloyd, A.M., and Davis, R.W. (1991). A steroid-inducible gene expression system for plant cells. *Proc. Natl. Acad. Sci. U.S.A.* 88**,** 10421-10425.

Shelomi, M., Watanabe, H. and Arakawa, G. (2014) Endogenous cellulase enzymes in the stick insect (Phasmatodea) gut. *J. Insect. Physiol.* 60, 25-30.

Shen, B., Sun, X., Zuo, X., Shilling, T., Apgar, J., Ross, M., et al (2012) Engineering a thermoregulated intein-modified xylanase into maize for consolidated lignocellulosic biomass processing. *Nat. Biotechnol.* 30, 1131-1136.

Shen, H., Mazarei, M., Hisano, H., Escamilla-Trevino, L., Fu, C., Pu, Y., Rudis, M.R., et al. (2013) A genomics approach to deciphering lignin biosynthesis in switchgrass. *Plant Cell* 25, 4342-4361.

Shi, W., Xie, S., Chen, X., Sun, S., Zhou, X., Liu, L., et al (2013) Comparative genomic analysis of the microbiome [corrected] of herbivorous insects reveals eco-environmental adaptations: biotechnology applications. *PLoS Genet.* 9, e1003131.

Shi, W.B., Ding, S.Y. and Yuan, J.S. (2011) Comparison of insect gut cellulase and xylanase activity across different insect species with distinct food sources. *Bioenerg Res.* 4, 1-10.

Shirley, D., Oppert, C., Reynolds, T.B., Miracle, B., Oppert, B., Klingeman, W.E. et al (2014) Expression of an endoglucanase from *Tribolium castaneum* (TcEG1) in *Saccharomyces cerevisiae*. *Insect Sci.* 21, 609-618.

Slaytor, M. (1992) Cellulose digestion in termites and cockroaches - What role do symbionts play? *Comp. Biochem. Phys. B* 103, 775-784.

Sun, Y., Cheng, J.J., Himmel, M.E., Skory, C.D., Adney, W.S., Thomas, S.R., et al (2007) Expression and characterization of *Acidothermus cellulolyticus* E1 endoglucanase in transgenic duckweed *Lemna minor* 8627. *Bioresour. Technol.* 98, 2866-2872.

Taylor, L.E., 2nd, Dai, Z., Decker, S.R., Brunecky, R., Adney, W.S., Ding, S.Y. and Himmel, M.E. (2008) Heterologous expression of glycosyl hydrolases *in planta*: a new departure for biofuels. *Trends Biotechnol.* 26, 413-424.

Todaka, N., Lopez, C.M., Inoue, T., Saita, K., Maruyama, J., Arioka, M., et al (2010) Heterologous expression and characterization of an endoglucanase from a symbiotic protist of the lower termite, *Reticulitermes speratus*. *Appl. Microbiol. Biotechnol.* 160, 1168-1178.

Tokuda, G. and Watanabe, H. (2007) Hidden cellulases in termites: revision of an old hypothesis. *Biol. Lett.* 3, 336-339.

Tribolium Genome Sequencing, C., Richards, S., Gibbs, R.A., Weinstock, G.M., Brown, S.J., Denell, R., er al (2008) The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 452, 949-955.

Uchima, C.A., Tokuda, G., Watanabe, H., Kitamoto, K. and Arioka, M. (2011) Heterologous expression and characterization of a glucose-stimulated beta-glucosidase from the termite *Neotermes koshunensis* in *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* 89, 1761-1771.

Uchima, C.A., Tokuda, G., Watanabe, H., Kitamoto, K. and Arioka, M. (2012) Heterologous expression in *Pichia pastoris* and characterization of an endogenous thermostable and high-glucose-tolerant beta-glucosidase from the termite *Nasutitermes takasagoensis*. *Appl. Environ. Microbiol.* 78, 4288-4293.

Valencia, A., Alves, A.P. and Siegfried, B.D. (2013) Molecular cloning and functional characterization of an endogenous endoglucanase belonging to GHF45 from the western corn rootworm, *Diabrotica virgifera virgifera*. *Gene* 513, 260-267.

Vogel, J.P., Garvin, D.F., Mockler, T.C., Schmutz, J., Rokhsar, D., Bevan, M.W., et al (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463, 763-768.

Vogel, J.P., Tuna, M., Budak, H., Huo, N.X., Gu, Y.Q. and Steinwand, M.A. (2009) Development of SSR markers and analysis of diversity in Turkish populations of *Brachypodium distachyon*. *BMC Plant Biol.* 9. 88-99.

Wang, Q., Qian, C., Zhang, X.Z., Liu, N., Yan, X. and Zhou, Z. (2012) Characterization of a novel thermostable beta-glucosidase from a metagenomic library of termite gut. *Enzyme Microb. Technol.* 51, 319-324.

Watanabe, H., Noda, H., Tokuda, G. and Lo, N. (1998) A cellulase gene of termite origin. *Nature* 394, 330-331.

Watanabe, H. and Tokuda, G. (2001) Animal cellulases. *Cell. Mol. Life. Sci.* 58, 1167-1178.

Watanabe, H. and Tokuda, G. (2010) Cellulolytic systems in insects. *Annu. Rev. Entomol.* 55, 609-632.

Wei, Y.D., Lee, K.S., Gui, Z.Z., Yoon, H.J., Kim, I., Zhang, G.Z., Guo, X., Sohn, H.D. and Jin, B.R. (2006) Molecular cloning, expression, and enzymatic activity of a novel endogenous cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comp. Biochem Physiol. B, Biochem. Mol. Biol.* 145, 220-229.

Weinstock, G.M., Robinson, G.E., Gibbs, R.A., Worley, K.C., Evans, J.D., Maleszka, R., et al (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443, 931-949.

Weng, X., Huang, Y., Hou, C. and Jiang, D. (2013) Effects of an exogenous xylanase gene expression on the growth of transgenic rice and the expression level of endogenous xylanase inhibitor gene RIXI. *J. Sci. Food. Agric.* **93**, 173-179. doi: 10.1002/jsfa.5746.

Werren, J.H., Richards, S., Desjardins, C.A., Niehuis, O., Gadau, J., Colbourne, J.K., et al (2010) Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science* 327, 343-348.

Willis, J.D., Oppert, B., Oppert, C., Klingeman, W.E. and Jurat-Fuentes, J.L. (2011) Identification, cloning, and expression of a GHF9 cellulase from *Tribolium castaneum* (Coleoptera: Tenebrionidae). *J. Insect Physiol.* 57, 300-306.

Willis, J.D., Oppert, C. and Jurat-Fuentes, J.L. (2010) Methods for discovery and characterization of cellulolytic enzymes from insects. *Insect Sci.* 17, 184-198.

Wilson, J.R. and Hacker, J.B. (1987) Comparative digestibility and anatomy of some sympatric C-3 and C-4 arid zone grasses. *Aust. J. Agr. Res.* 38, 287-295.

Wu, Y., Chi, S., Yun, C., Shen, Y., Tokuda, G. and Ni, J. (2012) Molecular cloning and characterization of an endogenous digestive beta-glucosidase from the midgut of the fungus-growing termite *Macrotermes barneyi*. *Insect Mol. Biol.* 21, 604-614.

Wyman, C.E. (2007) What is (and is not) vital to advancing cellulosic ethanol. *Trends Biotechnol.* 25, 153-157.

Xia, D., Wei, Y., Zhang, G., Zhao, Q., Zhang, Y., Xiang, Z. et al (2013) cDNA cloning, expression, and enzymatic activity of a novel endogenous cellulase from the beetle *Batocera horsfieldi*. *Gene* 514, 62-68.

Xiao, Z.Z., Storms, R. and Tsang, A. (2004) Microplate-based filter paper assay to measure total cellulase activity. *Biotechnol. Bioeng.* 88, 832-837.

Xue, G.P., Patel, M., Johnson, J.S., Smyth, D.J. and Vickers, C.E. (2003) Selectable marker-free transgenic barley producing a high level of cellulase (1,4-beta-glucanase) in developing grains. *Plant Cel. Rep.* 21, 1088-1094.

Yang, P.L., Wang, Y.R., Bai, Y.G., Meng, K., Luo, H.Y., Yuan, T.Z., Fan, Y.L. and Yao, B. (2007) Expression of xylanase with high specific activity from *Streptomyces olivaceoviridis* A1 in transgenic potato plants (*Solanum tuberosum* L.). *Biotechnol. Lett.* 29, 659-667.

Yeh, A.I., Huang, Y.C. and Chen, S.H. (2010) Effect of particle size on the rate of enzymatic hydrolysis of cellulose. *Carbohyd. Polym.* 79, 192-199.

Yu, L.X., Gray, B.N., Rutzke, C.J., Walker, L.P., Wilson, D.B. and Hanson, M.R. (2007) Expression of thermostable microbial cellulases in the chloroplasts of nicotine-free tobacco. *J. Biotechnol.* 131, 362-369.

Zhang, D., Lax, A.R., Bland, J.M. and Allen, A.B. (2011a) Characterization of a new endogenous endo-beta-1,4-glucanase of Formosan subterranean termite (*Coptotermes formosanus*). *Insect Biochem. Mol. Biol.* 41, 211-218.

Zhang, K., Agrawal, M., Harper, J., Chen, R. and Koros, W.J. (2011b) Removal of the fermentation inhibitor, Furfural, using activated carbon in cellulosic-ethanol production. *Ind. Eng. Chem. Res.* 50, 14055-14060.

Zhang, D., Allen, A.B. and Lax, A.R. (2012a) Functional analyses of the digestive beta-glucosidase of Formosan subterranean termites (*Coptotermes formosanus*). *J. Insect Physiol.* 58, 205-210.

Zhang, Q., Zhang, W., Lin, C.Y., Xu, X.L. and Shen, Z.C. (2012b) Expression of an *Acidothermus cellulolyticus* endoglucanase in transgenic rice seeds. *Protein Expr. Purif.* 82, 279-283.

Zhang, Y.H.P. (2015) Production of biofuels and biochemicals by in vitro synthetic biosystems: Opportunities and challenges. *Biotechnol. Adv.* 33, 1467-1483.

Zhou, X., Kovaleva, E.S., Wu-Scharf, D., Campbell, J.H., Buchman, G.W., Boucias, D.G. and Scharf, M.E. (2010) Production and characterization of a recombinant beta-1,4-endoglucanase (glycohydrolase family 9) from the termite *Reticulitermes flavipes*. *Arch. Insect Biochem. Physiol.* **74**, 147-162. doi: 10.1002/arch.20368

Ziegelhoffer, T., Raasch, J.A. and Austin-Phillips, S. (2001) Dramatic effects of truncation and sub-cellular targeting on the accumulation of recombinant microbial cellulase in tobacco. *Mol. Breeding* 8, 147-158.

Ziegelhoffer, T., Will, J. and Austin-Phillips, S. (1999) Expression of bacterial cellulase genes in transgenic alfalfa (*Medicago sativa* L.), potato (*Solanum tuberosum* L.) and tobacco (*Nicotiana tabacum* L.). *Mol. Breeding* 5, 309-318.

Ziegler, M.T., Thomas, S.R. and Danna, K.J. (2000) Accumulation of a thermostable endo-1,4-beta-D-glucanase in the apoplast of *Arabidopsis thaliana* leaves. *Mol Breeding* 6, 37-46.

2.13 Chapter 2 appendix

2.13.1 Definition box and tables

Endo- β -1,4-glucanase: Glycosyl hydrolase which has an open-ended catalytic side to randomly bind and cleave the internal β1-4 linkages of cellulose polymer chains producing reduced ends. Commonly called "cellulose."

Exo- β -1,4-glucanase (aka Cellobiohydrolase): Similar to its Endo- counterpart, except cellulose polymers are threaded through the enzyme to reach the catalytic site.

β-glucosidase (aka cellobiases): Glucosidase enzyme which hydrolyzes the β1-4 linkage cellobiose molecules releasing two glucoses.

Figure 2-1 Definition box: Classification of cellulase enzymes

Table 2-1: Endo- β -1,4-glucanase transgenic plant studies in which various host plant species and subcellular localization were targeted for enzyme overproduction. GHF represents glycosyl hydrolase family as defined by CAZy.org. The temperature and pH optimum are listed. Substrate and other abbreviations used in the table are listed below. ER = endoplasmic reticulum, CMC = carboxymethylcellulose, MUC = 4-methylumbelliferyl-β-D-cellobioside, BMCC = bacterial microcrystalline cellulose, n.o. (not optimized, used when reference reported single point activity instead of a range, TSP = total soluble protein, U = one unit is the release of 1µmol substrate per min, FU = Fluorescence units, ǂ = substrate was used only for qualitative stain for activity, ¥ = estimated values from published figure when exact numbers not provided, € = used to measure % relativity to a control within the experiment.

Enzyme Name	Source organism	Host plant	Organelle targeting	GHF	Temp. optimum	pH optimum	Substrate	Activity	Notes	Reference
E1	*Acidothermus cellulolyticus*	*Arabidopsis thaliana*	Apoplast	5	65 °C n.o.	5.5 n.o.	CMCǂ MUC	26% TSP		(Ziegler et al., 2000)
E1	*Acidothermus cellulolyticus*	*Lemna minor* 8627	Cytosol	5	80 °C	5.0	CMC	0.24 U/g fresh tissue		(Sun et al., 2007)
E1	*Acidothermus cellulolyticus*	*Nicotiana tabacum*	Chloroplast	5	81 °C	5.25	MUC	2756.3 pmol MU mg^{-1} TSP min^{-1}		(Dai et al., 2000a)
E1	*Acidothermus cellulolyticus*	*Nicotiana tabacum*	Apoplast	5	81 °C	5.25	MUC	18056 pmol MU mg^{-1} TSP min^{-1}		(Dai et al., 2005)
E1	*Acidothermus cellulolyticus*	*Nicotiana tabacum*	Chloroplast	5	65 °C n.o.	5.5 n.o.	MUC	2824 FU		(Jin et al., 2003)
E1	*Acidothermus cellulolyticus*	*Nicotiana tabacum*	Apoplast Chloroplast Cytosol	5	65 °C n.o.	5.5 n.o.	CMC MUC	1.6% TSP	Catalytic domain	(Ziegelhoffer et al., 2001)

Table 2-1 continued

Enzyme Name	Source organism	Host plant	Organelle targeting	GHF	Temp. optimum	pH optimum	Substrate	Activity	Notes	Reference
E1	*Acidothermus cellulolyticus*	*Nicotiana tabacum* *Zea mays*	Apoplast	5	50 °C n.o.	5.0 n.o.	MUC	0.3 ng/mg biomass	Tobacco provided by (Ziegelhoffer) corn provided by Biswall	(Brunecky et al., 2011)
E1	*Acidothermus cellulolyticus*	*Oryza sativa*	Apoplast	5	65 °C n.o.	5.5 n.o.	MUC	25000 pmol MU/mg protein/min		(Chou et al., 2011)
E1	*Acidothermus cellulolyticus*	*Oryza sativa*	Apoplast	5	65 °C n.o.	5.5 n.o.	MUC CMC Avicel	4.9% TSP 0.6 g/L glucose released 0.25 g/L glucose released		(Oraby et al., 2007)
E1	*Acidothermus cellulolyticus*	*Oryza sativa*	Endosperm	5	80 °C	5.0	CMC	830 U/g seed		(Zhang et al., 2012)
E1	*Acidothermus cellulolyticus*	*Zea mays*	Endosperm cell wall	5	50 °C n.o.	5.0 n.o.	MUC	16% TSP	First 40 codons optimized	(Hood et al., 2007)
E1	*Acidothermus cellulolyticus*	*Helianthus annuus*	Apoplast	5	65 °C n.o.	5.5 n.o.	MUC	2.5 mg/kg fresh tissue	Plant codon optimized	(Jung et al., 2014)
E1	*Acidothermus cellulolyticus*	*Solanum tuberosum*	Chloroplast	5	55 °C n.o.	5.5 n.o.	MUC	61000 pmol mg^{-1} min^{-1}		(Dai et al., 2000b)
E1	*Acidothermus cellulolyticus*	*Zea mays*	Apoplast	5	65 °C n.o.	5.5 n.o.	MUC	0.845 nmol/µg/min	Catalytic domain	(Biswas et al., 2006)
E1	*Acidothermus cellulolyticus*	*Zea mays*	Apoplast	5	50 °C n.o.	4.8 n.o.	MUC	1.13% TSP	Catalytic domain	(Ransom et al., 2007)

Table 2-1 continued

Enzyme Name	Source organism	Host plant	Organelle targeting	GHF	Temp. optimum	pH optimum	Substrate	Activity	Notes	Reference
Cel6A	*Thermobifida fusca*	*Medicago trunculata, Nicotiana tabacum, Solanum tuberosum*	Cytosol	6	55 °C n.o.	5.5 n.o.	Avicel CMC Corn stover	0.37 g/L 0.47 g/L 0.18 g/L		(Ziegelhoffer et al., 1999)
EG	Bacteria-like synthetic	*Saccharum officinarum*	Chloroplast ER Vacuole	5	40 °C n.o.	4.75 n.o.	CMC	23 nmol glucose/min/mg protein¥		(Harrison et al., 2011)
ENG1	*Oryza sativa*	*Oryza sativa*	Cytosol	9	n.d.	n.d.	MUC	n.d.	No transgenic plants regenerated	(Nigorikawa et al., 2012)
t-EGI	*Ruminococcus albus*	*Nicotiana tabacum*	Cytosol	5	30 °C n.o.	5.6 n.o.	CMC	1.7 µmol/mg protein	Truncated enzyme	(Kawazu et al., 1999)
Cel6A	*Thermobifida fusca*	*Nicotiana tabacum*	Chloroplast	6	50 °C n.o.	5.5	CMC BMCC	0.65 µm glucose released¥		(Yu et al., 2007)
sso1354	*Sulfolobus solfataricus*	*Nicotiana tabacum*	Apoplast ER	12	90 °C	4.5	CMC‡ Azo-CMC€ MUC Avicel	488 nmol 4MU min^{-1} mg^{-1} 0.1 g/L 5.5 mg L^{-1} h^{-1}		(Klose et al., 2012)
TrCel5A	*Trichoderma reesei*	*Nicotiana tabacum*	Apoplast	5	55 °C	4.8	CMC‡ Azo-CMC€ MUC	27 nmol 4MU min^{-1} mg^{-1}	Ethanol-inducible	(Klose et al., 2013)

Table 2-1 continued

TrCel5A	*Trichoderma reesei*	*Nicotiana tabacum*	Apoplast ER	5	55 °C n.o.	4.8 n.o.	Azo-CMC MUC	2.1 U mg^{-1} 46 nmol 4MU min^{-1} mg^{-1}		(Klose et al., 2015)

Table 2-2: Cellobiohydrolase transgenic plant studies in which various host plant species and subcellular localization was targeted for enzyme overproduction. GHF represents glycosyl hydrolase family as defined by CAZy.org. The tested temperature and pH are listed. Substrate abbreviations are listed below. CMC = carboxymethylcellulose, BMCC = bacterial microcrystalline cellulose, MUL = 4-methylumbelliferyl-β -D-lactopyranoside, MUC = 4-methylumbelliferyl-β-D-cellobioside, TSP = total soluble protein, ¥ = estimated values from published figure when exact numbers not provided.

Enzyme name	Source organism	Host plant	Subcellular localization	GHF	Temp	pH	Substrate	Activity	Notes	Reference
CBH I	*Trichoderma reesei*	*Zea mays*	Cell wall, ER, vacuole	7	50 °C	5.0	MUC	16% TSP	First 40 codons optimized	(Hood et al., 2007)
CBH I	Fungal-like synthetic	*Saccharum officinarum*	Chloroplast, ER, vacuole	7	40 °C	4.75	MUL	12.25 nmol 4-Mu /min/mg protein		(Harrison et al., 2011)
CBH II	Fungal-like synthetic	*Saccharum officinarum*	Chloroplast ER, vacuole	6	40 °C	4.75	Avicel PH-101	8.68 nmol reducing sugars/min/mg protein		(Harrison et al., 2011)
Cel6A	*Thermobifida fusca*	*Medicago trunculata, Nicotiana tabacum, Solanum tuberosum*	Cytosol	6	55 °C	5.5	CMC	0.1% TSP		(Ziegelhoffer et al., 1999)
Cel6B	*Thermobifida fusca*	*Medicago trunculata, Nicotiana tabacum, Solanum tuberosum*	Cytosol	6	55 °C	5.5	CMC	0.02% TSP		(Ziegelhoffer et al., 1999)

Table 2-2 continued

Enzyme name	Source organism	Host plant	Subcellular localization	GHF	Temp	pH	Substrate	Activity	Notes	Reference
Cel6B	*Thermobifida fusca*	*Nicotiana tabacum*	Chloroplast	6	50 °C	5.5	BMCC	0.55 µm glucose released¥		(Yu et al., 2007)
EXG1	*Oryza sativa*	*Oryza sativa*	Native target	6	37 °C	7.0	MUC	45 µg/ml 4-MU concentration		(Nigorikawa et al., 2012)

Table 2-3: β-glucosidase (BG) transgenic plant studies in which various host plant species and subcellular localization was targeted for enzyme overproduction. GHF represents glycosyl hydrolase family as defined by CAZy.org. The temperatures and pH ranges tested and (optimum) are listed. Substrate abbreviations are listed below. pNPG = *p*-nitrophenyl β-D-glucopyranoside, ¥ = estimated values from published figure when exact numbers not provided.

Enzyme name	Source organism	Host plant	Organelle targeting	GHF	Temp.	pH	Substrate	Activity	Notes	Reference
BEG1	*Oryza sativa*	*Oryza sativa*	Native	1	n.d.	n.d.	n.d.	n.d.		(Nigorikawa et al., 2012)
BglB	*Thermotoga maritima*	*Nicotiana tabacum* & *Arabidopsis thaliana*	Chloroplast Cytosol	3	40-90 °C (80 °C)	3-6 (4.5)	*p*NPG	180 mmol/ml¥		(Jung et al., 2010)
BglB	*Thermotoga maritima*	*Nicotiana tabacum* & *Arabidopsis thaliana*	Chloroplast	3	80 °C	4.5	*p*NPG	2090.9 mg/ml	New chloroplast promoter	(Jung et al., 2013)

Table 2-4: Endo-1,4-β –xylanase transgenic plant studies in which various host plant species and subcellular localization was targeted for enzyme overproduction. GHF represents glycosyl hydrolase family as defined by CAZy.org. The temperature and pH are values enzyme activity was performed. Substrate and other abbreviations are listed below. AZCL-xylan = azo-crosslinked-xylan, RBB-xylan = remazol brilliant blue R-D-xylan, U = amount of enzyme liberating 1 µmol xylose per min, AU = absorbance unit, ¥ = estimated values from published figure when exact numbers not provided.

Enzyme name	Source organism	Host plant	Organelle targeting	GHF	Temp.	pH	Substrate	Activity	Notes	Reference
ATX	*Thermobifida fusca*	*Oryza sativa*	Cytosol	11	50 °C	5.0	Birchwood xylan	3.64 U g^{-1} fresh tissue		(Weng et al., 2013)
XynA	*Clostridium thermocellum*	*Oryza sativa*	Cytoplasm	11	60 °C	7.0	Birchwood xylan	0.475 Unit/mg protein¥	Catalytic domain without native signal peptide	(Kimura et al., 2003)
XynA	*Dictyoglomus thermophilum*	*Arabidopsis thaliana Nicotiana tabacum*	Apoplast	11	85 °C	6.5	Wheat arabinoxylan	8.25 µmol reducing ends/min/mg¥	Plant codon optimized	(Borkhardt et al., 2010)
XynA	*Dictyoglomus thermophilum*	*Arabidopsis thaliana Nicotiana tabacum*	Apoplast	11	85 °C	6.5	Wheat arabinoxylan	8.25 µmol reducing ends/min/mg¥	Plant codon optimized	(Borkhardt et al., 2010)
mXynA	*Neocallimastix patriciarum*	*Hordeum vulgare*	Cytoplasm	11	40 °C	6.0	AZCL-xylan	27.06 AU/h/grain		(Patel et al., 2000)

Table 2-4 continued

Enzyme name	Source organism	Host plant	Organelle targeting	GHF	Temp.	pH	Substrate	Activity	Notes	Reference
XylII	*Trichoderma reesei*	*Arabidopsis thaliana*	Peroxisome Cytosol Chloroplast	11	50 °C	6.8	Xylan solution	24 U mg^{-1}		(Hyunjong et al., 2006)
XynII	*Trichoderma reesei*	*Arabidopsis thaliana*	Cytosol Chloroplast	11	50 °C	6.8	Xylan solution	22.7 U/mg		(Bae et al., 2008)
XynB	*Streptomyces olivaceoviridis*	*Solanum tuberosum*	Cytosol	11	60 °C	5.2	Birchwood xylan	87 µmol min^{-1} g-1		(Yang et al., 2007)
XynZ	*Clostridium thermocellum*	*Nicotiana tabacum*	Apoplast	10	60 °C	5.4	RBB-Xylan Birchwood xylan Oat spelt	46.6 U/mg	Truncated	(Herbers et al., 1995)

Table 2-5: Transgenic plants producing gene stacked cellulolytic enzymes produced in plants as reported in the literature.
Enzyme name is the name given by the author, source organisms are where the enzymes originated from, host plant is the plant species which was transformed with the enzyme, GHF is the glycosyl hydrolase family number as defined by CAZy.org, temperature and pH are values enzyme activity was performed, substrate is the material used to measure enzyme activity followed by parenthesis of reported highest mean activity, organelle targeting is where enzyme was targeted towards in the plant cell, activity is the reported maximum activity provided by authors, notes list other key interests for the reference, reference is the referred paper information was collected.

CBM = carbohydrate binding module, CMC = carboxymethylcellulose, MUC = methylumbelliferyl-β-D-cellobioside, *p*NPG = *p*-nitrophenol, RBB-xylan = Remazol brilliant blue, AX = arabinoxoylan, AZCL= azo-crosslinked, tsp = total soluble protein, AU = absorbance unit, ǂ = substrate was used only for qualitative stain for activity, ¥ = estimated values from published figure when exact numbers not provided.

Enzyme name	**Source organisms**	**Host plant**	**GHF**	**Temp.**	**pH**	**Substrate**	**Organelle targeting**	**Notes**	**Reference**
EII-hybrid	*Bacillus spp.*	*Hordeum vulgare*	16	65 °C	7.4	Azobarley β-glucan (40 ng ±18 ng) Lichenanǂ	Endosperm	Codon optimized	(Jensen et al., 1996)
EII-hybrid	*Bacillus spp.*	*Hordeum vulgare*	16	65 °C	7.4	Azobarley β-glucan (Up to 40 µg*mg^{-1}) Lichenanǂ	Endosperm	Codon optimized	(Horvath et al., 2000)

Table 2-5 continued

Enzyme name	Source organisms	Host plant	GHF	Temp.	pH	Substrate	Organelle targeting	Notes	Reference
E1 & Xyn10A	*Acidothermus cellulolyticus*	*Helianthus annuus*	5 &10	70 °C	6.0	MUC (up to 2.5 mg/kg) Birchwood xylan (up to 4.1 mg/kg)¥	Apoplast	Plant codon optimized	(Jung et al., 2014)
BglB:Cel5A	*Thermotoga maritime*	*Nicotiana tabacum*	3 & 5	80 °C	5.0	CMC (0.375 mg/ml) ¥ *p*NPG (240 µmol/ml) ¥	Chloroplast		(Lee et al., 2012)
XylII:Cel5A	*Trichoderma reesei* & *Thermotoga maritime*	*Nicotiana tabacum*	11 & 5	50 °C	7.0	CMC (0.275 mg/ml) ¥ Oat spelt (0.25 mg/ml) ¥	Chloroplast		(Lee et al., 2012)
Cel6B:Cel5A	*Thermobifida fusca* & *Thermotoga maritime*	*Nicotiana tabacum*	5 & 6	50 °C	7.0	CMC (0.375 mg/ml) ¥ Filter paper (0.225 mg/ml) ¥	Chloroplast		(Lee et al., 2012)
Cel-Hyb1 (CelA & Cel6G)	*Neocallimastix patriciarum* & *Piromyces sp.*	*Hordeum vulgare*	& 6	40 °C	6.0	CMC‡ AZCL-glucan (8,447 ±2059 AU/h/g/ grain) AZCL-HE-Cellulose (4,041 ±657 AU/h/g/ grain) Lichenan (1585 ±195 U/g grain	Cytoplasm		(Xue et al., 2003)
TmCel5A & TmCel5A fused CMB6A	*Thermotoga maritime*	*Nicotiana tabacum*	5 & CBM 6	50 °C	6.0	CMC (up to 2.0 Unit/mg tsp)¥	Apoplast Chloroplast Cytosol	Fused with cellulose binding module	(Mahadevan et al., 2011)
Xyln-ara	*Clostridium thermocellum* & *Geobacillus stearothermophilus*	*Nicotiana tabacum*	10 & 51	60 °C	6.0	RBB-xylan (55.3 µmol dye/min/µg) Wheat-AX (22.5 µmol xylose/min/µg) ¥ Rye-AX (27.5 µmol xylose/min/µg)¥	Cytosol	Fusion of catalytic domains	(Fan and Yuan, 2010)

Table 2-6: Heterologously produced insect cellulolytic enzymes as reported in the literature

Insect taxanomic order followed by insect family. Genus species are insect source of enzyme, Enzyme and family represents glycosyl hydrolase and its corresponding family as defined by CAZy.org, temperature and pH are values reported for best performance, substrate is the material used to measure enzyme activity, heterologous expression system defines transformation system evaluated, reference is the referred paper information was collected.

EG = endo-1,4-β-glucanase, BG = β-glucosidase, Xyl = Xylanase, U = units of mMol of substrate released per minute, n.d. (not determined); n.r. (not reported); n.o. (not optimized, used when references reports activity, however was a single temperature or pH testing).

Order: **Family**	Genus species	Enzyme (Family)	Temp.	pH	Substrate	Activity	Expression system	Reference:
Coleoptera								
Cerambycidae	*Apriona germari*	EG (GH 45)	50 °C	6.0	CMC	992 (U/mg)	Sf9 insect cells	(Lee et al., 2004)
Cerambycidae	*Apriona germari*	EG (GH 45)	50 °C	6.0	CMC	812 (U/mg)	Sf9 insect cells	(Lee et al., 2005)
Cerambycidae	*Apriona germari*	EG (GH 5)	55 °C	6.0	CMC	1037 (U/mg)	Sf9 insect cells	(Wei et al., 2006)
Cerambycidae	*Batocera horsfieldi*	EG (GH 45)	50 °C	6.0	CMC	928 (U/mg)	baculovirus expression in *Bombyx mori*	(Xia et al., 2013)
Cerambycidae	*Anoplophora malasiaca*	EG (GH 45)	50 °C	4.0	CMC	319.22 ± 9.3 (U/mg)	baculovirus expression in *Bombyx mori*	(Chang et al., 2012)
Chrysomelidae	*Diabrotica virgifera virgifera*	EG (GH 45)	45 °C	6.0	CMC	166.67 nmol min^{-1} $mgprotein^{-1}$	*E. coli*	(Valencia et al., 2013)

Table 2-6 continued

Order: **Family**	Genus species	Enzyme (Family)	Temp.	pH	Substrate	Activity	Expression system	Reference:
Tenebrionidae	*Tribolium castaneum*	EG (GH 9)	50 °C	9.0	CMC	12.9 (U/mg)	S2 insect cells	(Willis et al., 2011)
Isoptera								
Kalotermitidae	*Neotermes koshunensis*	BG (GH1)	45 °C 40 °C	5.0 5.0	Cellobiose pNPG	156.7 (U/mg) 12.4 (U/mg)	*E. coli* *A.orizae*	(Ni et al., 2007) (Uchima et al., 2011)
Rhinotermitidae	*Coptotermes formosanus*	EG (GH5)	70 °C	6.0	CMC	105 (U/mg)	*E. coli*	(Inoue et al., 2005)
Rhinotermitidae	*Coptotermes formosanus*	EG (GH9)	43 °C	5.6	CMC Cellodextrin Filter paper	325 (U/mg)	*E. coli*	(Zhang et al., 2011)
Rhinotermitidae	*Coptotermes formosanus*	BG (GH1)	49 °C	5.6-6.2	Cellobiose	462.6 (U/mg)	*E. coli*	(Zhang et al., 2012)
Rhinotermitidae	*Reticulitermes flavipes*	EG (GH9)	50-60 °C	6.5-7.5	CMC	1.3 (U/mg)	baculovirus expression in *Trichoplusia ni*	(Zhou et al., 2010)
Rhinotermitidae	*Reticulitermes flavipes*	BG (GH1)	40 °C	6.5-7.0	Cellobiose	638.0±39.0 (U/mg)	baculovirus expression in *Trichoplusia ni*	(Scharf et al., 2010)
Rhinotermitidae	*Reticulitermes santonensis*	BG (GH1)	40 °C	6.0	pNPG	0.441 (U/mg)	*E.coli*	(Matteotti et al., 2011)
Rhinotermitidae	*Reticulitermes santonensis*	Xyl (GH11)	55 °C	5.0	AZCL-xylan Beechwood xylan	1837 (U/mg)	*E. coli*	(Matteotti et al., 2012)
Rhinotermitidae	*Reticulitermes speratus*	EG (GH9)	45 °C	5.5	CMC	1200 ± 92 (U/mg)	*A.oryzae*	(Hirayama et al., 2010)
Rhinotermitidae	*Reticulitermes speratus*	EG (GH7)	45 °C	6.5	CMC	603 (U/mg)	*A. oryzae*	(Todaka et al., 2010)
Termitidae	*Globitermes sulphureus*	BG (GH1)	90 °C	6.0	pNPG	110 (U/mg)	*E.coli*	(Wang et al., 2012)

Table 2-6 continued

Order: **Family**	Genus species	Enzyme (Family)	Temp.	pH	Substrate	Activity	Expression system	Reference:
Termitidae	*Macrotermes annandalei*	Xyl GH11)	55 °C	7.5	Beechwood xylan	733 (U/mg)	*E.coli*	(Liu et al., 2011)
Termitidae	*Macrotermes barneyi*	BG (GH1)	45 °C	5.0	Cellobiose pNPG	206 (U/mg)	*E.coli*	(Wu et al., 2012)
Termitidae	*Nasutitermes takasagoensis*	EG (GH9)	65 °C	6.0	CMC	1392 ± 57 (U/mg)	*A.orizae*	(Hirayama et al., 2010)
Termitidae	*Nasutitermes takasagoensis*	BG (GH1)	65 °C	5.5	pNPG	5.83 (U/mg)	*P. pastris*	(Uchima et al., 2012)
Lepidoptera								
Noctuidae	*Spodoptera frugiperda*	BG (GH1)	30 °C	6.0	*p*NPG	2.4 ($mM^{-1}s^{-1}$)	*E. coli*	(Marana et al., 2004)
Orthoptera								
Acrididae	Grasshopper (n.r)	BG (GH 1)	60 °C	9.0	*p*NPG	2.61 ± 0.75 ($mM^{-1}s^{-1}$)	*E. coli*	(Shi et al., 2011)
Gryllidae	*Teleogryllus emma*	EG (GH9)	40 °C	5.0	CMC	3118.4 (U/mg)	Sf9 cells	(Kim et al., 2008)

Chapter 3: Heterologous production of the TcEG-1 beetle (*Tribolium castaneum)* cellulase in switchgrass improves sugar release and alters plant cell wall architecture

A version of this chapter will be submitted to Scientific Reports by Dr. C. Neal Stewart, Jr. with following authors Jonathan D. Willis, Joshua N. Grant, Mitra Mazarei, Lindsey M. Kline, Caroline S. Rempe, A. Grace Collins, Geoffrey B. Turner, Stephen R. Decker, Robert W. Sykes, Mark F. Davis, Nicole Labbe, Juan L. Jurat-Fuentes, and C. Neal Stewart Jr. Jonathan D. Willis drafted the manuscript, carried out vector construction, generated the majority of the transgenic plants, performed the statistical analysis, performed qRT-PCR and enzyme analysis, and prepared plant samples for recalcitrance analysis. Joshua N. Grant prepared the histological slides, performed image analysis, and contributed to manuscript drafting. Mitra Mazarei participated in experimental design and data analysis, assisted with revisions of the manuscript and coordination of the study. CR assisted with developing the program for automated image analysis. Lindsey Kline and Nicole Labbe performed cellulose crystallinity index analysis. Geoffrey B. Turner, Stephen R. Decker, Robert W. Sykes, and Mark F. Davis assisted with performing lignin and sugar release assays. A. Grace Collins contributed to tissue culture and generation of transgenic plants and care of plants. Juan L. Jurat-Fuentes provided original TcEG-1 construct and edited the manuscript. C. Neal Stewart, JR conceived of the study and its design and coordination and assisted with revisions to the manuscript. All authors contributed to text and data analysis and interpretation. All authors read and approve final version of the manuscript.

3.1 Abstract

Genetically engineering dedicated biofuel crops, such as switchgrass, that produce their own cell wall digesting cellulase enzymes would reduce costs of cellulosic ethanol production. The majority of experiments testing plant produced fungal and bacterial cellulases has been

performed in model plants. Many herbivorous insects possess cellulases that have similar properties to cellulases from fungi and bacteria. Production of the TcEG-1 cellulase, from *Tribolium castaneum* (red flour beetle) has been demonstrated in *Escherichia coli* and *Saccharomyces cerevisiae* to be optimally active at 50 °C and pH 12.0. TcEG-1 was produced in transgenic switchgrass with a range of endoglucanase activity of 0.16 to 0.05 units (μM glucose release/min/mg) at 50 °C with pH 12.0 from fresh leaf tissue. TcEG-1 activity from air dried leaves was unchanged from that of green tissue, but when dried in a desiccant oven (46 °C), specific enzyme activity decreased by 60%. Saccharification was increased in one transgenic event by 28%, which had a concurrent decrease in lignin content of 9%. Histological analysis revealed an increase cell wall thickness with no change to cell area or perimeter. Transgenic plants produced more, albeit narrower, tillers with equivalent dry biomass as the control. This is the first study in which an insect cellulase has been produced in plants.

Keywords: Switchgrass, glycosyl hydrolase, β-1,4-endoglucanase, insect cellulase, biofuel

3.2 Introduction

Plant biomass polysaccharides can be converted into ethanol for use as a replacement of fossil fuels. Cellulosic feedstocks include crop residues such as maize stover or dedicated biomass crops such as switchgrass (*Panicum virgatum*). Dedicated biomass crops are attractive inasmuch as growing demands for fuel might be met by low-input bioenergy crops grown on marginal lands unsuitable for food crop production (Mitchell et al., 2016). Plants utilizing C4 photosynthesis, such as switchgrass, have increased water-use efficiency over C3 plants. Furthermore, switchgrass and other perennial grasses have lower nutrient fertilizer requirements compared with most C4 cereal crops (e.g., maize) (Vogel, 2008; Wright and Turhollow, 2010; Wullschleger et al., 2010). For bioenergy, switchgrass aboveground biomass would be harvested using standard forage baling equipment at the end of the growing season after the first frost. This timing allows the plant to remobilize N and other nutrients to belowground biomass, thereby endowing high nutrient use efficiency. Utilization of farmer contracts from biorefineries would allow farmers to 'permanently' install switchgrass at low risk (Griffith et al., 2014; Griffith et al., 2012). Production of perennial, dedicated cellulosic feedstocks on marginal lands will allow farmers to produce profitable and environmentally stable fuel source (Mitchell et al., 2016).

Enzyme cost is significant for cellulosic ethanol production. High titers of cellulase cocktails are required to convert recalcitrant plant cell walls into simple sugars for fermentation. Economic modelling has demonstrated that cellulosic ethanol refineries should use an integrated approach of producing their feedstock and cellulases on site to reduce total cost (Johnson, 2016). A

biofuel feedstock that simultaneously produces its own cocktail of cellulolytic enzymes has been one all-in-one model integrated system for reducing enzyme costs (Lopez-Casado et al., 2008; Li et al., 2014). There are several challenges to address in designing such a feedstock. Complete digestion of cellulose in the plant cell wall requires the synergistic actions of three types of cellulase enzymes: endoglucases, exoglucanases, and β-glucosidases, which are all glycosyl hydrolases (Bayer et al, 1997; Bhat, 2000). Internal cellulose bonds are broken by endoglucanases (Bayer et al, 1997; Bhat, 2000; Schulein, 2000). Unbound chain ends of cellulose are cleaved by exoglucanases (also called cellobiohydrolases), which release the base units of cellulose, cellobiose. Cellobiose consists of two inverted glucose units, which are broken into free glucose by β-glucosidases. Genetically engineered feedstocks would conceivably require the concerted synthesis of each type of enzyme in a manner that would also not decrease plant growth. Most experiments have tested one cellulase gene at a time. Such experiments have utilized easily-transformed non-feedstock plant species such as *Arabidopsis thaliana* (Arabidopsis), *Nicotiana tabacum* (tobacco), and *Oryza sativa* (rice). All cellulase genes engineered into plants thus far have been from either bacteria or fungi (Willis et al., 2016).

A potential source for bioprospecting biocatalysts for biofuels has been digestive enzymes from insect species (Willis et al. 2016). Researchers once believed that insects had none-to-few cell wall-degrading of their own, but rather relied on those from symbiotic organisms in their guts to digest plant biomass. For some species this opinion is predomininant, however new genomic and proteomic analyses have shown that insects produce endogenous enzymes (Tokuda and Watanabe, 2007; Watanabe et al., 1998). For example, termites can survive solely on cellulose,

and thus would be a source of cell wall digesting enzymes for bioprospecting (Watanabe and Tokuda, 2010). Heterologous expression of insect cellulases in model microbes has proven to be potential alternatives to bacterial and fungal-sourced enzymes. These insect cellulases have temperature optima from 40 to 65 Cº and perform optimally at alkaline pH (Shi et al., 2013; Shi et al., 2011; Willis et al., 2011). Therefore, there appears to be potential advantages to using insect cellulases in feedstocks for subsequent biofuel production.

Here we report on transgenic switchgrass that overexpresses a gene encoding TcEG-1, an endoglucanase from the digestive system of the red flour beetle (*Tribolium castaneum*). Our goal was to assess the potential of this switchgrass-beetle cellulase system on improving saccharification of the feedstock host.

3.3 Methods

3.3.1 Vector construction

The *TcEG-1* open reading frame sequence (Willis et al., 2011) was amplified by PCR and cloned into the pCR8 entry vector and further then Gateway® sub-cloned into the pANIC-10A plant expression vector (Mann et al., 2012b) to yield the pANIC-10A-TcEG-1 vector. The expression cassette containing *TcEG-1* was 5' flanked by the constitutive maize ubiquitin 1 promoter (ZmUbi1),and 3' flanked by the AcV5 epitope tag and the octopine synthase terminator (Figure 1). The pANIC-10A-TcEG-1 also contained cassettes that included a hygromycin selectable marker and an orange fluorescent protein (OFP) reporter gene from the hard coral *Porites porites*

(*pporRFP*) (Mann et al., 2012a). An epi-fluorescence microscope (Olympus stereo microscope SZX12, Olympus America, Centre Valley, PA) having a 535/30 nm excitation filter and 600/50 nm emissions filter was used to track OFP fluorescence during transgenic callus development and to identify individual putative transgenic lines *in vitro*.

3.3.2 Transgenic plant production

Seed derived callus of the 'Performer' switchgrass cultivar was used to generate Type II embryogenic callus (Burris et al., 2009), which was stably transformed using *Agrobacterium tumefaciens* strain EHA105 harboring the pANIC-10A-TcEG-1 expression vector. Transformed calli were grown in LP9 growth medium (Burris et al., 2009) supplemented with 400 mg/L timentin and 40 mg/L hygromycin for approximately two months. Subsequently the transgenic calli were transferred to regeneration medium (Wong et al., 2011), supplemented with 250 mg/L cefotaxime to stimulate regeneration (Danilova and Dolgikh, 2004). Ten transgenic plants were successfully regenerated, rooted and acclimated as previously described by Burris et al. (2009). Transgenic and non-transgenic control plants were grown in growth chambers under 16 h light/8 h dark cycles at 25°C until moved to a greenhouse. Fertilizer (0.02% solution of Peter's soluble 20-20-20) was applied twice per month.

3.3.3 RNA extraction and qRT-PCR analysis for TcEG-1 transcript abundance

Quantitative RT-PCR was performed to estimate *TcEG-1* transcript abundance in transgenic T0 and non-transgenic plants. Total RNA was isolated from stem internodes of triplicate tillers at

the R1 (reproductive) developmental stage (Hardin et al., 2013) per event using TRI Reagent following manufacturer's instructions (Sigma-Aldrich, St. Louis, MO). Purified RNA was treated with DNase-1 (Promega, Madison, WI) and 3 µg treated RNA was used to generate cDNA using oligo-dT and Superscript III according to manufacturer's instructions (Life Technologies, Carlsbad, CA). qRT-PCR analysis was performed with Power SYBR Green PCR master mix (Life Technologies) according to manufacturer's protocols for optimization of annealing temperature, primer concentration, and cDNA concentration. The optimized qRT-PCR protocol employed a dilution of cDNA 1:100 with thermal cycling at 95 °C for 3 min, and 40 cycle repeats of (95 °C for 10 s and 50.0 °C for 30 s). The *TcEG-1* primers were: *TcEG-1*_F 5'- CTGGATTACAATGCGGGATTTC -3' and *AcV5_R* 5'-AGACCAGCCGCTCGCATCTTTCCAAGA -3'. The relative levels of transcripts were normalized to switchgrass ubiquitin 1 (*PvUbi1*) as a reference gene (Shen et al., 2009) and primers were PvUbi1_F 5'- CAGCGAGGGCTCAATAATTCCA -3' and PvUbi1_R 5' - TCTGGCGGACTACAATATCCA - 3' (Xu et al., 2011). All experiments were carried out in triplicate technical replicates. The differential Ct method was used to measure transcript abundance after normalization to *PvUbi1* according to Schmittgen and Livak (2008). Statistical analysis was determined with triplicate stem internodes averaged from triplicate measuring using SAS® (Version 9.3 SAS Institute Inc., Cary, NC) programming of mixed model ANOVA and least significant difference (LSD) for all quantifiable data.

TcEG-1 protein sequence was aligned against the switchgrass proteome to determine homology which could have cause a change to native glycosyl hydrolases within switchgrass. A pBLAST

search (https://phytozome.jgi.doe.gov/pz/portal.html) revealed 61 targets with no greater than 46% identity match (Table S3-1).

3.3.4 Plant protein extraction

Protein extraction procedure was according to Oraby et al. (2007) with modification. Briefly, 100 mg leaf tissue from fresh triplicate R1 development stage tillers were ground under liquid nitrogen to a fine powder. For the dry biomass enzyme analysis, triplicate R1 development stage tillers were collected and either air dried in the greenhouse for two weeks or placed in a desiccant oven at 46 °C for three days as described by Hardin et al. (2013). A protein extraction buffer of 50 mM sodium acetate, pH 5.5, 100 mM NaCl, 10% glycerol, 0.5 ml disodium EDTA, 1 mM PMSF, and a 1:200 dilution of Sigma plant proteinase inhibitor added to the fine powder in a 2 ml centrifuge tube and vortexed for 30 seconds. Samples were centrifuged at 4 °C for 10 minutes at 10,000 x g and supernatant was transferred to fresh tube. A secondary centrifugation was used when excess extracellular debris was present. Protein concentrations were estimated via Bradford assay using the Pierce Coomassie Protein Assay Reagent (Thermo Fisher, Wilmington, DE) following the manufacturer's instructions with bovine serum albumin (BSA) as standard. Samples were stored at 4 °C until ready for downstream assays.

3.3.5 Endoglucanase activity

Endoglucanase activity of protein extracts from the method described above was determined using a modified dinitrosalicylic acid (DNSA) assay (Miller, 1959) with carboxymethyl cellulose (CMC) sodium salt (Sigma-Aldrich, St. Louis, MO) as substrate. Protein samples (10 μg) were

added in triplicate to substrate solutions (2% w/v in 50 mM sodium phosphate buffer, pH 12.0) and incubated for 1 h at 50 °C. A modified DNSA reagent containing Rochelle salt (Miller, 1959) was added to the samples to halt enzymatic activity, after which a color change developed at 100 °C for 15 min. Samples were centrifuged at 2,000 x g for two minutes to precipitate any remaining substrate. Supernatants were transferred to polystyrene microplates and spectral absorbance at 595 nm was read on a Synergy HT microplate reader (BioTek, Winooski, VT) using the KC4 software (v. 3.1). Background amounts of native sugars and any possible native cellulases from switchgrass leaves were corrected for by subtracting non-transgenic protein values from transgenic TcEG-1 protein values. One unit of cellulolytic activity was defined as the amount of enzyme that produces 1 μmol of reducing sugar (glucose equivalents) per minute at 50 °C at pH 12.0. Specific activities were reported as units per mg of protein and represented averages of three independent replicates. Statistical analysis was determined with triplicate measures of proteins extracted from triplicate fresh leaves using SAS® (Version 9.3 SAS Institute Inc.) programming of mixed model ANOVA and least significant difference (LSD) for all quantifiable data. The standard error of the mean was calculated and reported in data displays. *P*-values of ≤ 0.05 were considered to be statistically significant.

3.3.6 Cell wall sugar release and lignin content and composition

Switchgrass tillers were collected at the R1 developmental stage from triplicate greenhouse-grown plants and air dried for 3 weeks at room temperature before grinding to 1 mm (20 mesh) particle size. Sugar release efficiency was determined via NREL high-throughput sugar release assays on pre-treated extractive- and starch-free samples (Decker et al., 2012; Selig et al., 2010).

Glucose and xylose release was determined by colorimetric assays with total sugar release being the sum of glucose and xylose released. The lignin content and composition was determined by pyrolysis molecular beam mass spectrometry (py-MBMS) on extractive- and starch-free samples via NREL high-throughput assays (Sykes et al., 2009). Statistical analysis was determined with triplicate measures of biomass collected from triplicate pots using SAS® (Version 9.3 SAS Institute Inc.) programming of mixed model ANOVA and least significant difference (LSD) for all quantifiable data.

3.3.7 Cell wall histology and measurements

Stem samples were collected from second-to-basal internode from three tillers at the R1 developmental stage and immediately placed in 2 ml Eppendorf tubes containing FAA (50 ml 95% ethanol (EtOH), 5 ml glacial acetic acid, and 10 ml 40% formaldehyde, and 35 ml distilled H_2O). Internodes were incubated for four days in FAA on a shaker, after which the FAA was discarded and replaced with a 10% EtOH solution. After two hours of gentle shaking, the 10% EtOH was discarded and replaced with 20% EtOH. At two-hour intervals each, 30%, and 50% EtOH were used as serial replacements, followed by 75% EtOH for a four-hour incubation, which was subsequently replaced by 95% EtOH. A two-day 95% EtOH incubation was performed with a change of solution midway through the incubation. Infiltration of glycol methacrylate was performed using a JB-4 Embedding Kit (Sigma-Aldrich) following manufacturer's instructions. Infiltrated samples were placed in molds (Sigma-Aldrich) and embedded under a nitrogen vacuum until hardened. After hardening, stem samples were mounted and sectioned to 5 µM with a glass blade microtome (Sorvall Dupont JB-4 microtome,

Newtown, CT). Dark field staining was performed with Pontamine Fast Scarlet 4B, which binds specifically to cellulose (Thomas et al., 2013). Dark field staining of total cell wall components was performed with Calcofluor White (Hughes and McCully, 1975). After staining, bright field and dark field images at multiple objectives were taken on a Zeiss Axioplan 2 Compound Microscope (Carl Zeiss, Oberkochen, Germany). Slides stained with Pontamine Fast Scarlet 4B were observed under a 543 nm laser and images were obtained using a Leica Confocal Microscope. Images were analyzed using ImageJ (Schneider et al., 2012) software to measure cell area, perimeter, and cell wall thickness both by hand and with a custom built program using Python and Python Imaging Library. Hand measuring occurred for 100 cell wall segments on three slide sections. Program measuring was conducted for all cell walls on 20 slide sections. The custom program, Python Cell Wall Thickness (pyCWT), was developed for the batch determination of plant cell wall thickness from images (cross-sections of plant stem internodes with fluorescently labeled cell walls). This automated approach of approximating plant cell wall thickness was written in Python (Python Software Foundation, Python Language Reference, version 2.7, https://www.python.org) using functions from the Python Imaging Library (PIL, Secret Labs AB) and the Scientific Python (Scipy) libraries ndimage and misc (Jones et al., 2001) and includes a graphical user interface (GUI) to easily work with batches of files and adjust image processing parameters. Each image analyzed with pyCWT underwent a series of processing steps that converted the image to grayscale, normalized pixel brightness distribution using a histogram, smoothed with a Gaussian blur, and then converted to black and white pixels based on the mean pixel brightness of the current image. A stepwise example of how pyCWT is detailed in supplementary figure 1 (Figure S1). A binary opening function with a 3X3 matrix

over 2 iterations was then used to better differentiate dark and light objects. The image was segmented and objects labeled using the PIL function “measurements.label()”. Labeled pixels were mapped back to their coordinate values and binary erosion was used to get a border within each labeled object, which corresponds to the border of a plant cell. Centroids of labeled objects were found with the PIL function “measurementscenter_of_mass()”. The border coordinate values were used to calculate area, using an implementation of Green’s Theorum by Jamie Bull (Bull posted function 2012), and perimeter, by summing distances between adjacent border coordinates, of each object. A size cutoff of 200% of the average cell area and perimeter was implemented to restrict the program from counting large gaps as cells. The mode for cell wall thickness was the recorded value for each image.

Cell wall thickness is calculated by dilating each labeled object (presumably a plant cell) 1 pixel width at a time while keeping track of the total number of objects. When two objects merge, meaning the total object count decreases by one, the current pixel count is considered the thickness of that cell wall. A distribution of all cell wall thickness in pixels is plotted based on the number of dilations required for objects to merge. The mode of cell wall thickness was recorded and when these values were compared with average thickness from manual measurements with ImageJ, there was no significant difference when compared with a t-test at $p < 0.05$ (Figure S2).

Statistical analysis was performed on the pyCWT image rendered datausing SAS® (Version 9.3 SAS Institute Inc.) programming of mixed model ANOVA and least significant difference (LSD) reported.

3.3.8 Cellulose crystallinity index

Collected tillers at the R1 developmental stage were ground to ½ mm (40 mesh) particle size and the crystallinity index was measured by Fourier transform infrared (FTIR) spectra were collected using a diamond crystal of an attenuated total reflectance (ATR) accessory of a Perkin Elmer Spectrum One spectrometer (Waltham, MA). Spectra were collected over the range of 4000-650 cm^{-1} in the absorbance mode, with 1 cm-1 resolution and 8 scans per spectra. Ten spectra were collected for each sample. Data pretreatments were then ATR corrected and normalized in the Spectrum One software. The index of crystallinity was calculated by the intensity ratio between the bands at 1422 and 899 cm^{-1}, assigned to CH_2 bending mode and deformation of anomeric CH, respectively (Kataoka and Kondo, 1998). Statistical analysis was determined with triplicate measures of biomass collected using SAS® (Version 9.3 SAS Institute Inc.) as stated above.

3.3.9 Plant growth analysis

Transgenic T0 and non-transgenic control lines were split into triplicate, single-tiller replicates and placed in a random design in the greenhouse. Plants were grown until the R1 developmental stage and measured for tiller number. The five tallest tillers for each replicate were used as a determination of aboveground plant height. These tillers were also used for the determination of stem diameter taken at 10 cm above potting level at internodes. At the R1 stage, the aboveground biomass was harvested for each plant and air dried in the greenhouse for approximately two weeks and biomass tallied. Statistical analysis was performed with using SAS® (Version 9.3 SAS Institute Inc.) programming of mixed model ANOVA and least significant difference (LSD) for all quantifiable data.

3.4 Results

3.4.1 Production of TcEG-1 transgenic plants, expression and enzyme activity

Hygromycin-resistant and orange-fluorescent callus and shoots resulted in the recovery of ten independent transgenic events. Transcript abundance in tillers ranged between 70-fold (relative to *PvUbi1* gene) in event Tc-1 to two-fold in Tc-3 (Fig 3-1B). All transgenic plants had functionally-active TcEG-1 endoglucanase on CMC substrate from an increase of reduced sugars at 50 °C at pH 12.0 (Fig 3-2A). Event Tc-1 had the highest enzyme activity of 0.16 ± 0.02 U/mg while event Tc-3 had the lowest activity at 0.05 ± 0.02 U/mg (Fig 3-2A). In addition the TcEG-1 enzyme activity of event Tc-1 over a range of pH conditions was highest at pH 12 (increased by 193%; Fig 3-2B).

Since commercial switchgrass biomass would be harvested and air dried in the field, it was important to assay endoglucanase activity in dry switchgrass host biomass. We used a subset of transgenic events based on endoglucanase activity. Transgenic event Tc-1 maintained the highest enzymatic activity (0.23 ± 0.02 U/mg) among all air dried plants tested (Fig 3-3). Only transgenic event, Tc-1, had any discernable enzyme activity after oven drying, but was only 60% of the air dried biomass (Fig 3-3).

3.4.2 The effects of TcEG-1 production on lignin and sugar release

Only event Tc-6 showed a significant increase in glucose release (49% higher than control) (Fig 3-4A). There was no difference in xylose release between transgenics and non-transgenic

control (Fig 3-4B). Event Tc-6 had significantly higher (28% more than control) total sugar release (Fig 3-4C). While there is no a priori reason that TcEG-1 synthesis would affect lignin, we routinely perform lignin composition and content measurements in feedstocks since lignin is an important factor in cell wall recalcitrance (Chen and Dixon 2007). Lignin content was decreased by up to 9% in events Tc-1, Tc-2, Tc-3, Tc -4, Tc -5, Tc -6, Tc -12, whereas in events Tc-9, Tc-10, and Tc -11 was equivalent to control (Fig 3-5A). Event Tc-6 had an increased S/G ratio by up to 14%, whereas events Tc-1, Tc-2, Tc-5, and Tc-11 had a decreased S/G ratio by up to 7% compared to control. The S/G ratio was unchanged in events Tc-3, Tc-4, Tc-9, Tc-10, and Tc-12 compared to control (Fig 3-5B).

3.4.3 Cell wall architecture and cellulose crystallinity

Histological analysis of stem internode sections revealed no differences in cell wall areas or perimeters among plants (Fig 3-6A and B). Transgenic events Tc-1, Tc-2, Tc-5, Tc-9, Tc-11, and Tc-12 had increased cell wall thickness by up to 93% in event Tc-11 with an average overall increase of 37% over the control (Fig 3-6C). Cellulose crystallinity was increased by up to 18% in events Tc-3, Tc-5, Tc-9, and Tc-10, decreased by up to 10% in events Tc-2 and Tc-12, and was unchanged in events Tc-1, Tc-4, Tc-6, and Tc -11 relative to the control (Fig 3-7).

3.4.4 Plant morphology and growth was minimally effected by TcEG-1

The transgenic switchgrass events that were selected for the growth study were based on enzyme activity (Tc-1 Fig 3-2A) and sugar release efficiency (Tc-6 and Tc-12 Fig 3-4A). Most growth

characteristics of selected transgenic events were not different from one another nor from the control (Fig 3-8A). There were no differences in plant height or dry biomass among lines (Fig 3-8B and 3-8E). All transgenic stem diameters were less than that of the control (Fig 3-8C). Tiller number was increased by 71% for event Tc-1 while for Tc-6 and Tc-12 were not different than the control (Fig 3-8D).

3.5 Discussion

An engineered self-degrading feedstock would represent a significant step toward an integrated strategy for reducing costs and increasing biofuel production (Furtado et al., 2014; Johnson, 2016; Wyman, 2007; Lynd et al. 2008). The transgenic overexpression of microbial cellulase genes in plants has increased release of fermentable sugars (Furtado et al., 2014; Lambertz et al., 2014; Taylor et al., 2008; Willis et al., 2016). The majority of insect cellulase properties are similar to those of microbial cellulases (thermotolerant and acidic pH optima) rendering them as feasible candidates for their production in lignocellulosic feedstocks (Slaytor, 1992; Watanabe and Tokuda, 2010; Oppert et al., 2010; Willis et al., 2016). However, some insect gut system cellulases have been discovered to have an alkaline pH optima. The variability of insect cellulase pH range most likely arises from digestive system environments having a variable pH range of 4.0 to 12.0 (Brune, 2014 and Dow, 1992). Here we present evidence for the first alkaline insect cellulase produced in a dedicated biofuel crop; in this case switchgrass.

Transgenic switchgrass lines produced functional TcEG-1 with its alkaline pH 12.0 optimum and thermal activity of 50 °C (Fig 3-2). When TcEG-1 was produced in switchgrass it had a similar

enzyme activity profile that was observed in S2 insect cells and *Saccharomyces cerevisiae*: thermotolerant and optimal alkaline pH (Shirley et al., 2014; Willis et al., 2011). However, TcEG-1 endoglucanase activity from switchgrass was much lower than both previous heterologous microbial production systems, which might be caused by suboptimal plant expression conditions. For example TcEG-1 activity in our experiment was lower than a sugar cane-produced synthetic endoglucanase that was targeted to the chloroplast, endoplasmic reticulum (ER), or the vacuole. In the latter experiment, the approximate highest endoglucanase activity observed was 23 nmol/min/mg protein on a CMC substrate (Harrison et al., 2011). When compared with other putative insect cellulases, crude digestive protein extraction from *T. castaneum* was relatively low. The overexpression of an acidic insect cellulase gene in switchgrass could improve saccharification yields (Oppert et al., 2010). Improvement of TcEG-1 activity could be performed by intracellular targeting to specific organelles or even to specific tissues as has been reported in E1 and other cellulases produced in plants (Willis et al., 2016).

Extracted TcEG-1 was active from fresh and dried tissue, whereas enzyme activity from oven dried tissue was attenuated (Fig 3-2 and 3-3), which may have been caused by decreased moisture content of tissue. Moisture content has been shown to improve sugar release and cellulosic ethanol yields by up to 25% from rehydrated switchgrass and sugarcane biomass compared to air dried (Ewanick and Bura, 2011). However, when leaves from transgenic alfalfa producing the E1-catalytic domain were dried at 50 °C, they showed no change to enzyme activity when compared to enzymes extracted from fresh leaf material (Ziegelhoffer et al., 2001). Yet, the dried leaf extraction reported in Ziegelhoffer et al., (2001) was carried out under

different conditions from the fresh material with assistance of exogenous cellulase and pectinase mixture (Ziegelhoffer et al., 2001). The addition of exogenous cellulase could increase the E1 yield recovered from the dried material over that of the non-cellulase extraction method used for fresh leaf material which would show as increased enzyme activity. The overproduced heterologous cellulase from transgenic maize and rice seeds is active after drying (Biswas et al., 2006; Garda et al., 2015; Hood et al., 2007; Zhang et al., 2012). However, fresh seed was not tested to compare if drying could affect the enzyme activity.

Green switchgrass harvested mid-season under forage production systems that is field-dried for at least a week has biomass moisture content of ~25% (Venturi et al., 2004; Kumar and Sokhansanj, 2007), which we simulated by using air dried biomass. The drying experiment demonstrates the feasibility of generating TcEG-1 switchgrass lines in a field-like harvest setting with active enzyme.

Saccharification was only increased in event Tc-6 (Fig 3-4) which is linked closely to the lower lignin content and increased S/G ratio (Fig 3-5). Saccharification was increased up to 15% in E1 transgenic maize and tobacco at the E1's optimal pH 5.0 (Brunecky et al., 2011). Although the saccharification of TcEG-1 switchgrass was unchanged in most events, the saccharification experiments were carried out at pH 5.0 (Decker et al., 2012) in which TcEG-1 is minimally active (Fig 3-2B). TcEG-1 events could be also used as a crossing partner with other low-lignin switchgrass, such as COMT and MYB transgenic lines modified for decreased lignin and

increased sugar release efficiency (Baxter et al., 2014, Baxter et al., 2015; Fu et al., 2011; Shen et al., 2012) to further improve the saccharification efficiency by transgene stacks.

Since the production of any cellulase *in planta* might potentially have off effects to plant cells, we analyzed transgenic stem internode cell structure via histology. TcEG-1 switchgrass cell morphology was not different from the control in cell wall area or perimeter, however cell wall thickness was increased (Fig 3-6C). The majority of histological examinations of other hydrolase expressing plants has mostly been to determine proper organelle targeting of the enzymes (Brunecky et al., 2011; Jung et al., 2013; Klose et al., 2015; Lee et al., 2012). In a few cases, some phenotypic alterations have been observed. For example, rice plants that over-expressed a native exoglucanase, EXG1, had an extra lacunae which was not observed in the controls (Nigorikawa et al., 2012). Tobacco plants with constitutive expression of TrCel5A showed an increase in the number of small vessels in the stems (Klose et al., 2013). Morphology of TcEG-1 switchgrass stem internodes showed usual cellular growth patterns.

The increase in cell wall thickness (Fig 3-6C) in TcEG-1 switchgrass could be alteration to cellulose structure. Similar cell wall thickening has been observed in Arabidopsis overexpressing an endoglucanase from Aspen (*PttCel9A1*) in which there was decreased cellulose crystallinity (Takahashi et al., 2009). Cellulose crystallinity is an estimation of the crystalline structure compactness of cellulose polymer chains. A high cellulose crystallinity attributes to increased recalcitrance of cellulose break down to enzymatic action (Park et al., 2010). TcEG-1 switchgrass had a range of cellulose crystallinity with no correlation to

expression or production pattern (Fig 3-7). The increase cell wall thickness could have been caused by the increased presence of other cell wall components that were not examined, for example, tightly bound cell wall sugars unaccounted for during saccharification.

Transgenic TcEG-1 switchgrass plants had more tillers with narrower stem thickness, but these changes resulted in no effect on biomass yield (Fig 3-8). Negative pleiotropic effects observed in transgenic plants that produced GHs include reduced height, wrinkled leaves, and sterility (Klose et al., 2013; Nigorikawa et al., 2012; Taylor et al., 2008; Ziegelhoffer et al., 2001). Transgenic potato plants with production of E1 under the control of a constitutive promoter were deformed when grown at 35 °C and high light density (450 µmol quanta $m^{-2}s^{-1}$), but when the temperature was decreased to 25 °C with lower light density (200 µmol quanta $m^{-2}s^{-1}$), the plants grew normally (Taylor et al., 2008). When E1 was targeted to the chloroplast, no adverse growth was observed at 35 °C and high light density in potato (Taylor et al., 2008). E1 is a thermophilic enzyme whose activity was likely attenuated with the decreased temperature restoring normal phenotype. Possibly TcEG-1 activity is attenuated in switchgrass as the pH of plant cells is approximately neutral (Kurkdjian et al., 1978; Martiniere et al., 2013) where TcEG-1 activity is low thus preventing deleterious growth effects.

Transgenic tobacco producing the endoglucanase TrCel5A from the bacterium *Trichoderma reesei*, driven by constitutive CaMV 35S promoter were dwarfed and had active endoglucanase (Klose et al., 2013). When TrCel5A expression was controlled by an ethanol inducible promoter, transgenic plants produced active enzyme with no change in plant phenotype

compared to control (Klose et al., 2013). Furthermore, when TrCel5A was targeted to the apoplast or ER in tobacco, the plants showed a wrinkled and necrotic leaf phenotype. Apoplast targeted TrCel5A tobacco also produced shorter plants, whereas ER-targeted TrCel5A were equal to the controls (Klose et al., 2015). These studies indicate that organelle targeting might not be sufficient to eliminate pleiotropic effects on plant growth. Production of TcEG-1 in switchgrass showed a non-deleterious effect on plant growth. The thicker cell walls might be one reason for equivalent biomass relative to controls even through tillers were smaller (Fig 3-8).

3.6 Conclusion

This is the first study in which an active insect cellulase has been produced in any plant; in this case a dedicated bioenergy crop, switchgrass. TcEG-1 enzyme activity was observed in all ten independent transgenic events. However, the enzyme activity was decreased in oven-dried biomass compared to air-dried biomass. Sugar release was increased in one event accompanied with the lowest amount of lignin content. Cellulose crystallinity was altered with no correlation to saccharification or enzyme activity. Transgenic plants developed thinner tillers than control, but did have thicker cell walls and increase in tiller number which maintained equivalent dry biomass yield. Improving engineering strategies by plant codon optimization and organelle targeting could increase protein yield and efficacy as seen in other glycosyl hydrolase plant production systems as a feasible strategy for improved biofuels.

3.7 Acknowledgments

We acknowledge the technical assistance Garret Montgomery and Kelsey Harrell for help in maintaining plants and collection of protein extracts for enzyme analysis. We thank Angela Ziebell, Erica Gjersing, Crissa Doeppke, and Melvin Tucker of NREL for their assistance with the cell wall characterization and Susan Holladay for her assistance with data entry into LIMS. This work was supported by funding from the BioEnergy Science Center. The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

3.8 References

Bayer, E. A., Chanzy, H., Lamed, R. and Shoham, Y. (1998) Cellulose, cellulases, cellulosomes. *Curr Opin Struct Biol* **8,** 548-57.

Bhat, M.K. (2000) Cellulases and related enzymes in biotechnology. *Biotechnol Adv* **18**, 355-383.

Biswas, G.C.G., Ransom, C. and Sticklen, M. (2006) Expression of biologically active *Acidothermus cellulolyticus* endoglucanase in transgenic maize plants. *Plant Sci* **171**, 617-623.

Bouton, J.H. (2007) Molecular breeding of switchgrass for use as a biofuel crop. *Curr Opin Genet Dev* **17**, 553-558.

Brune, A. (2014) Symbiotic digestion of lignocellulose in termite guts. *Nat. Rev. Microbiol.* 12, 168-180.

Brunecky, R., Selig, M.J., Vinzant, T.B., Himmel, M.E., Lee, D., Blaylock, M.J. and Decker, S.R. (2011) *In planta* expression of *A. cellulolyticus* Cel5A endocellulase reduces cell wall recalcitrance in tobacco and maize. *Biotechnol Biofuels* **4**, 1.

Burris, J.N., Mann, D.G.J., Joyce, B.L. and Stewart, C.N. Jr. (2009) An improved tissue culture system for embryogenic callus production and plant regeneration in switchgrass (*Panicum virgatum* L.). *Bioenerg Res* **2**, 267-274.

Carroll, A. and Somerville, C. (2009) Cellulosic biofuels. *Annu Rev Plant Biol* **60**, 165-182.

Chen, F. and Dixon, R. A. (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat Biotechnol* **25**, 759-761

Danilova, S.A. and Dolgikh, Y.I. (2004) The stimulatory effect of the antibiotic cefotaxime on plant regeneration in maize tissue culture. *Russ J Plant Physiol* **51**, 559-562.

Decker, S.R., Carlile, M., Selig, M.J., Doeppke, C., Davis, M., Sykes, R., Turner, G. and Ziebell, A. (2012) Reducing the effect of variable starch levels in biomass recalcitrance screening. *Methods Mol Biol* **908**, 181-195.

Dow, J.A. (1992) pH Gradients in Lepidopteran midgut. *J. Exp. Biol.* 172, 355-375.

Ewanick, S. and Bura, R. (2011) The effect of biomass moisture content on bioethanol yields from steam pretreated switchgrass and sugarcane bagasse. *Bioresour Technol* **102**, 2651-2658.

Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M., Jr., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A. and Wang, Z.Y. (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc Natl Acad Sci USA* **108**, 3803-3808.

Furtado, A., Lupoi, J.S., Hoang, N.V., Healey, A., Singh, S., Simmons, B.A. and Henry, R.J. (2014) Modifying plants for biofuel and biomaterial production. *Plant Biotechnol J* **12**, 1246-1258.

Garda, M., Devaiah, S.P., Vicuna Requesens, D., Chang, Y.K., Dabul, A., Hanson, C., Hood, K.R. and Hood, E.E. (2015) Assessment of field-grown cellulase-expressing corn. *Transgenic Res* **24**, 185-198.

Griffith, A.P., Haque, M. and Epplin, F.M. (2014) Cost to produce and deliver cellulosic feedstock to a biorefinery: Switchgrass and forage sorghum. *Appl Energ* **127**, 44-54.

Griffith, A.P., Larson, J.A., English, B.C. and McLemore, D.L. (2012) Analysis of contracting alternatives for switchgrass as production alternative on an east Tennessee beef and crop farm. *AgBioForum* **15**, 206-216.

Harrison, M.D., Geijskes, J., Coleman, H.D., Shand, K., et al (2011) Accumulation of recombinant cellobiohydrolase and endoglucanase in the leaves of mature transgenic sugar cane. *Plant Biotechnol J* 9, 884-896

Hughes, J. and McCully, M.E. (1975) The use of an optical brightener in the study of plant structure. *Stain Technol* **50**, 319-329.

Johnson, E. (2016) Integrated enzyme production lowers the cost of cellulosic ethanol. *Biofuel Bioprod Bior* **10**, 164-174.

Jung, S., Lee, D.S., Kim, Y.O., Joshi, C.P. and Bae, H.J. (2013) Improved recombinant cellulase expression in chloroplast of tobacco through promoter engineering and 5' amplification promoting sequence. *Plant Mol Bio* **83**, 317-328.

Kataoka, Y. and Kondo, T. (1998) FT-IR microscopic analysis of changing cellulose crystalline structure during wood cell wall formation. *Macromolecules* **31**, 760-764.

Klose, H., Gunl, M., Usadel, B., Fischer, R. and Commandeur, U. (2013) Ethanol inducible expression of a mesophilic cellulase avoids adverse effects on plant development. *Biotechnol Biofuels* **6**, 53.

Klose, H., Gunl, M., Usadel, B., Fischer, R. and Commandeur, U. (2015) Cell wall modification in tobacco by differential targeting of recombinant endoglucanase from *Trichoderma reesei*. *BMC Plant Biol* **15**, 54.

Kumar, A. and Sokhansanj, S. (2007) Switchgrass (*Panicum vigratum*, L.) delivery to a biorefinery using integrated biomass supply analysis and logistics (IBSAL) model. *Bioresour Technol* **98**, 1033-1044.

Kurkdjian, A., Leguay, J.J. and Guern, J. (1978) Measurement of intracellular pH and aspects of its control in higher plant cells cultivated in liquid medium. *Respir Physiol* **33**, 75-89.

Lambertz, C., Garvey, M., Klinger, J., Heesel, D., Klose, H., Fischer, R. and Commandeur, U. (2014) Challenges and advances in the heterologous expression of cellulolytic enzymes: a review. *Biotechnol Biofuels* **7**.

Lee, D.S., Lee, K.H., Jung, S., Jo, E.J., Han, K.H. and Bae, H.J. (2012) Synergistic effects of 2A-mediated polyproteins on the production of lignocellulose degradation enzymes in tobacco plants. *J Exp Bot* **63**, 4797-4810.

Li, Q., Song, J., Peng, S., Wang, J.P., Qu, G.Z., Sederoff, R.R. and Chiang, V.L. (2014) Plant biotechnology for lignocellulosic biofuel production. *Plant Biotechnol J* **12**, 1174-1192.

Li, R.Y. and Qu, R.D. (2011) High throughput *Agrobacterium*-mediated switchgrass transformation. *Biomass Bioenerg* **35**, 1046-1054.

Lopez-Casado, G., Urbanowicz, B.R., Damasceno, C.M.B. and Rose, J.K.C. (2008) Plant glycosyl hydrolases and biofuels: a natural marriage. *Curr Opin Plant Biol* **11**, 329-337.

Lynd, L. R., Laser, M.S., Bransby, D., Dale, B.E., Davison, B., Hamilton, R., Himmel M., Keller, M., McMillan, J.D., Sheehan, J. and Wyman C.E. (2008) How biotech can transform biofuels. *Nat Biotechnol* **26**, 169-172.

Mann, D.G., Abercrombie, L.L., Rudis, M.R., Millwood, R.J., Dunlap, J.R. and Stewart, C.N., Jr. (2012a) Very bright orange fluorescent plants: endoplasmic reticulum targeting of

orange fluorescent proteins as visual reporters in transgenic plants. *BMC Biotechnol* **12**, 17.

Mann, D.G.J., LaFayette, P.R., Abercrombie, L.L., King, Z.R., Mazarei, M., Halter, M.C., Poovaiah, C.R., Baxter, H., Shen, H., Dixon, R.A., Parrott, W.A. and Stewart, C.N., Jr. (2012b) Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum* L.) and other monocot species. *Plant Biotechnol J* **10**, 226-236.

Martiniere, A., Bassil, E., Jublanc, E., Alcon, C., Reguera, M., Sentenac, H., Blumwald, E. and Paris, N. (2013) *In vivo* intracellular pH measurements in tobacco and Arabidopsis reveal an unexpected pH gradient in the endomembrane system. *Plant Cell* **25**, 4028-4043.

Miller, G.L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Chem* **31**, 426-428.

Mitchell, R.B., Schmer, M.R., Anderson, W.F., Jin. V., Balkcom, K.S., Kiniry, J., Coffin, A., and White, P. (2016) Dedicated energy crops and crop residues for bioenergy feedstocks in the central and eastern USA. *Bioenerg Res* **9,** 384-398

Nigorikawa, M., Watanabe, A., Furukawa, K., Sonoki, T. and Ito, Y. (2012) Enhanced saccharification of rice straw by overexpression of rice exo-glucanase. *Rice* **5**, 14.

Ohkuma, M. (2003) Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Appl Microbiol Biotechnol* **61**, 1-9.

Oraby, H., Venkatesh, B., Dale, B., Ahmad, R., Ransom, C., Oehmke, J. and Sticklen, M. (2007) Enhanced conversion of plant biomass into glucose using transgenic rice-produced endoglucanase for cellulosic ethanol. *Transgenic Res* **16**, 739-749.

Oppert, C., Klingeman, W.E., Willis, J.D., Oppert, B. and Jurat-Fuentes, J.L. (2010) Prospecting for cellulolytic activity in insect digestive fluids. *Comp Biochem Physiol B Biochem Mol Biol* **155**, 145-154.

Park, S., Baker, J.O., Himmel, M.E., Parilla, P.A. and Johnson, D.K. (2010) Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol Biofuels* **3**, 10.

Schmittgen, T.D. and Livak, K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nature Protoc* **3**, 1101-1108.

Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-675.

Schulein, M. (2000) Protein engineering of cellulases. *Biochim Biophys Acta* **1543**, 239-252.

Selig, M.J., M.P., T., Sykes, R.W., Reichel, K.L., Brunecky, R., Himmel, M.E., Davis, M.F. and Decker, S.R. (2010) Lignocellulose recalcitrance screening by integrated high-throughput hydrothermal pretreatment and enzymatic saccharification. *Ind Biotechnol* **6**, 104-111.

Shi, W., Xie, S., Chen, X., Sun, S., Zhou, X., Liu, L., Gao, P., Kyrpides, N.C., No, E.G. and Yuan, J.S. (2013) Comparative genomic analysis of the microbiome of herbivorous insects reveals eco-environmental adaptations: biotechnology applications. *PLoS Genet* **9**, e1003131.

Shi, W.B., Ding, S.Y. and Yuan, J.S. (2011) Comparison of insect gut cellulase and xylanase activity across different insect species with distinct food sources. *Bioenerg Res* **4**, 1-10.

Shirley, D., Oppert, C., Reynolds, T.B., Miracle, B., Oppert, B., Klingeman, W.E. and Jurat-Fuentes, J.L. (2014) Expression of an endoglucanase from *Tribolium castaneum* (TcEG1) in *Saccharomyces cerevisiae*. *Insect Sci* **21**, 609-618.

Slaytor, M. (1992) Cellulose digestion in termites and cockroaches - What role do symbionts play? *Comp Biochem Phys B* **103**, 775-784.

Sykes, R., Yung, M., Novaes, E., Kirst, M., Peter, G. and Davis, M. (2009) High-throughput screening of plant cell-wall composition using pyrolysis molecular beam mass spectroscopy. *Methods Mol Biol* **581**, 169-183.

Takahashi, J., Rudsander, U.J., Hedenstrom, M., Banasiak, A., Harholt, J., Amelot, N., Immerzeel, P., Ryden, P., Endo, S., Ibatullin, F.M., Brumer, H., del Campillo, E., Master, E.R., Scheller, H.V., Sundberg, B., Teeri, T.T. and Mellerowicz, E.J. (2009) KORRIGAN1 and its Aspen homolog *PttCel9A1* decrease cellulose crystallinity in arabidopsis stems. *Plant Cell Physiol* **50**, 1099-1115.

Taylor, L.E., 2nd, Dai, Z., Decker, S.R., Brunecky, R., Adney, W.S., Ding, S.Y. and Himmel, M.E. (2008) Heterologous expression of glycosyl hydrolases *in planta*: a new departure for biofuels. *Trends Biotechnol* **26**, 413-424.

Thomas, J., Ingerfeld, M., Nair, H., Chauhan, S. and Collings, D. (2013) Pontamine fast scarlet 4B: a new fluorescent dye for visualising cell wall organisation in radiata pine tracheids. *Wood Sci Technol* **47**, 59-75.

Tokuda, G. and Watanabe, H. (2007) Hidden cellulases in termites: revision of an old hypothesis. *Biol Lett* **3**, 336-339.

Venturi, P., Monti, A., Piani, I. and Venturi, G. (2004) Evaluation of harvesting and post-harvesting techniques for energy destination of switchgrass. Proceedings of the 2nd world conference on biomass for energy, industry, and climate protection, 10-14 May, 2004. Rome, Italy. Pp 234-236

Vogel, J. (2008) Unique aspects of the grass cell wall. *Curr Opin Plant Biol* **11**, 301-307.

Watanabe, H., Noda, H., Tokuda, G. and Lo, N. (1998) A cellulase gene of termite origin. *Nature* **394**, 330-331.

Watanabe, H. and Tokuda, G. (2001) Animal cellulases. *Cell Mol Life Sci* **58**, 1167-1178.

Watanabe, H. and Tokuda, G. (2010) Cellulolytic systems in insects. *Annu Rev Entomol* **55**, 609-632.

Willis J.D., Mazarei, M., Stewart., C. N., Jr. (2016) Transgenic plant-produced hydrolytic enzymes and the potential of insect gut-derived hydrolases for biofuels. *Front Plant Sci* **7**: 675.

Willis, J.D., Klingeman, W.E., Oppert, C., Oppert, B. and Jurat-Fuentes, J.L. (2010a) Characterization of cellulolytic activity from digestive fluids of *Dissosteira carolina* (Orthoptera: Acrididae). *Comp Biochem Physiol B Biochem Mol Biol* **157**, 267-272.

Willis, J.D., Oppert, B., Oppert, C., Klingeman, W.E. and Jurat-Fuentes, J.L. (2011) Identification, cloning, and expression of a GHF9 cellulase from *Tribolium castaneum* (Coleoptera: Tenebrionidae). *J Insect Physiol* **57**, 300-306.

Willis, J.D., Oppert, C. and Jurat-Fuentes, J.L. (2010b) Methods for discovery and characterization of cellulolytic enzymes from insects. *Insect Sci* **17**, 184-198.

Wong, D.W., Chan, V.J., Batt, S.B., Sarath, G. and Liao, H. (2011) Engineering *Saccharomyces cerevisiae* to produce feruloyl esterase for the release of ferulic acid from switchgrass. *J Ind Microbiol Biotechnol* **38**, 1961-1967.

Wright, L. and Turhollow, A. (2010) Switchgrass selection as a "model" bioenergy crop: A history of the process. *Biomass Bioenerg* **34**, 851-868.

Wullschleger, S.D., Davis, E.B., Borsuk, M.E., Gunderson, C.A. and Lynd, L.R. (2010) Biomass production in switchgrass across the United States: Database description and determinants of yield. *Agron J* **102**, 1158-1168.

Wyman, C.E. (2007) What is (and is not) vital to advancing cellulosic ethanol. *Trends Biotechnol* **25**, 153-157.

Zhang, Q., Zhang, W., Lin, C.Y., Xu, X.L. and Shen, Z.C. (2012) Expression of an *Acidothermus cellulolyticus* endoglucanase in transgenic rice seeds. *Protein Expr Purif* **82**, 279-283.

Ziegelhoffer, T., Raasch, J.A. and Austin-Phillips, S. (2001) Dramatic effects of truncation and sub-cellular targeting on the accumulation of recombinant microbial cellulase in tobacco. *Mol Breeding* **8**, 147-158.

3.9 Chapter 3 Appendix

3.9.1 Figures

Figure 3-1 Transformation vector map and relative transcript abundance of TcEG-1 in transgenic switchgrass. A) pANIC-10A-TcEG-1 vector used for expression of *TcEG-1* in transgenic switchgrass. LB = Left border, PvUbi2 =Switchgrass ubiquitin 2 promoter and intron, hph = hygromycin B phosphotransferase coding region, 35S T= 35S terminator sequence, PvUbi1 = Switchgrass ubiquitin 1 promoter and intron, pporRFP = *Porites porites* orange fluorescent protein coding region, NOS T = *Agrobacterium tumefaciens nos* terminator sequence, ZmUbi1 = Maize ubiquitin 1 promoter, R1 and R2 = *attR1* and *attR2* recombinase sites 1 and 2, *TcEG1* = *TcEG-1* cDNA open reading frame, AcV5 = epitope tag, RB = Right border, Kan^r = kanamycin resistance gene, ColE1 = origin of replication in *E. coli*, pVS1 = origin of replication in *A. tumefaciens,* OCS T= octopine synthase terminator sequence. B) Relative transcript abundance of *TcEG-1* in stem internodes from transgenic events (Tc-1 to Tc-12). Relative expression analysis were determined by qRT-PCR and normalized to switchgrass ubiquitin 1 (*PvUbi1*). Bars represent mean values of three replicates ± standard error. Bars represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$).

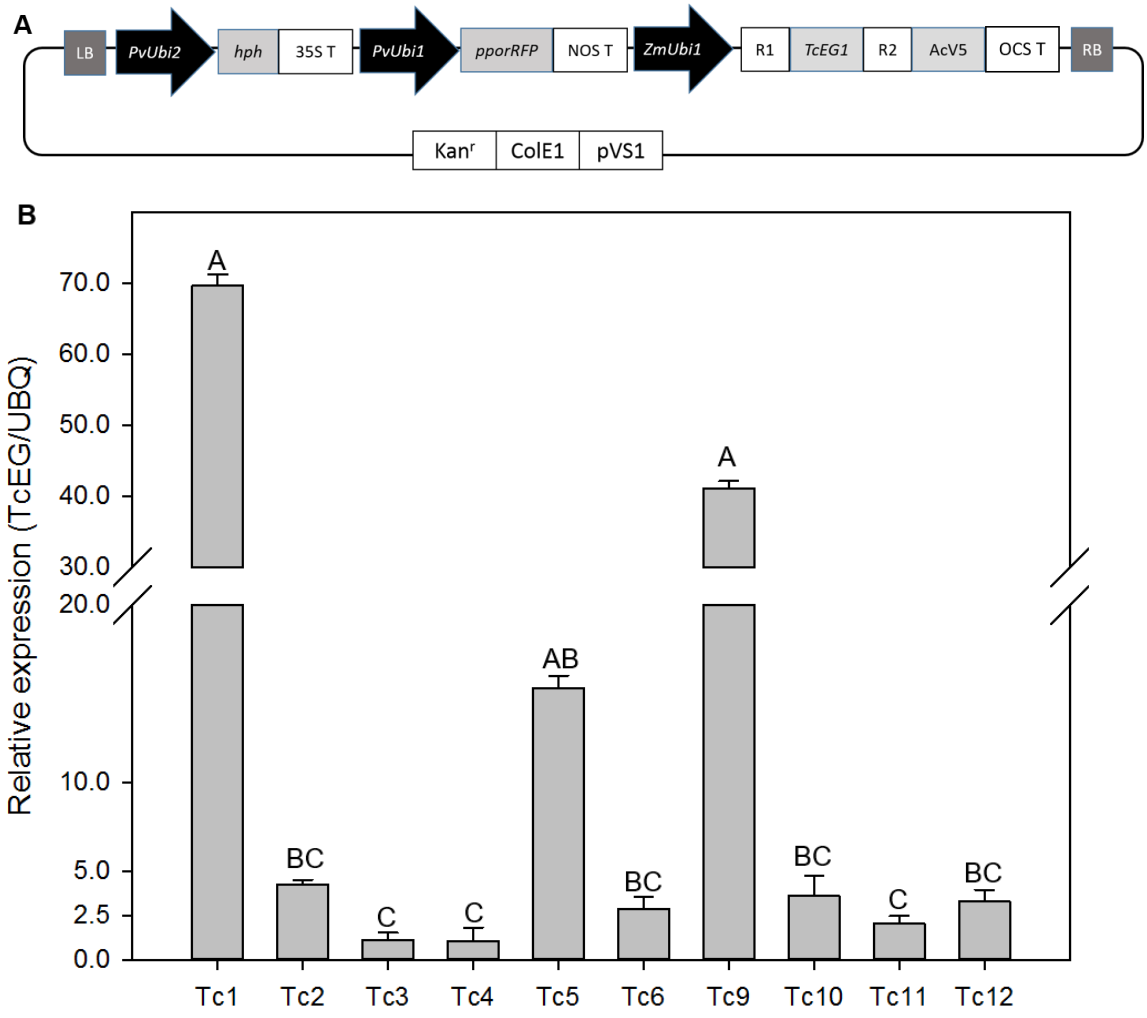

Figure 3-1 continued

Figure 3-2 Endoglucanase activity (units/mg of protein) from fresh leaves of transgenic TcEG-1 plants. A) Endoglucanase activity measurement using carboxymethyl cellulose (CMC) as substrate at pH 12.0 on TcEG-1 extracted from fresh leaves. Bars represent mean values of three replicates ± standard error. Bars represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$). B) Gradient pH measurement of endoglucanase activity of TcEG-1 extracted from fresh leaves of transgenic event Tc-1. Data points represent mean values of three replicates ± standard error. Data points represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$).

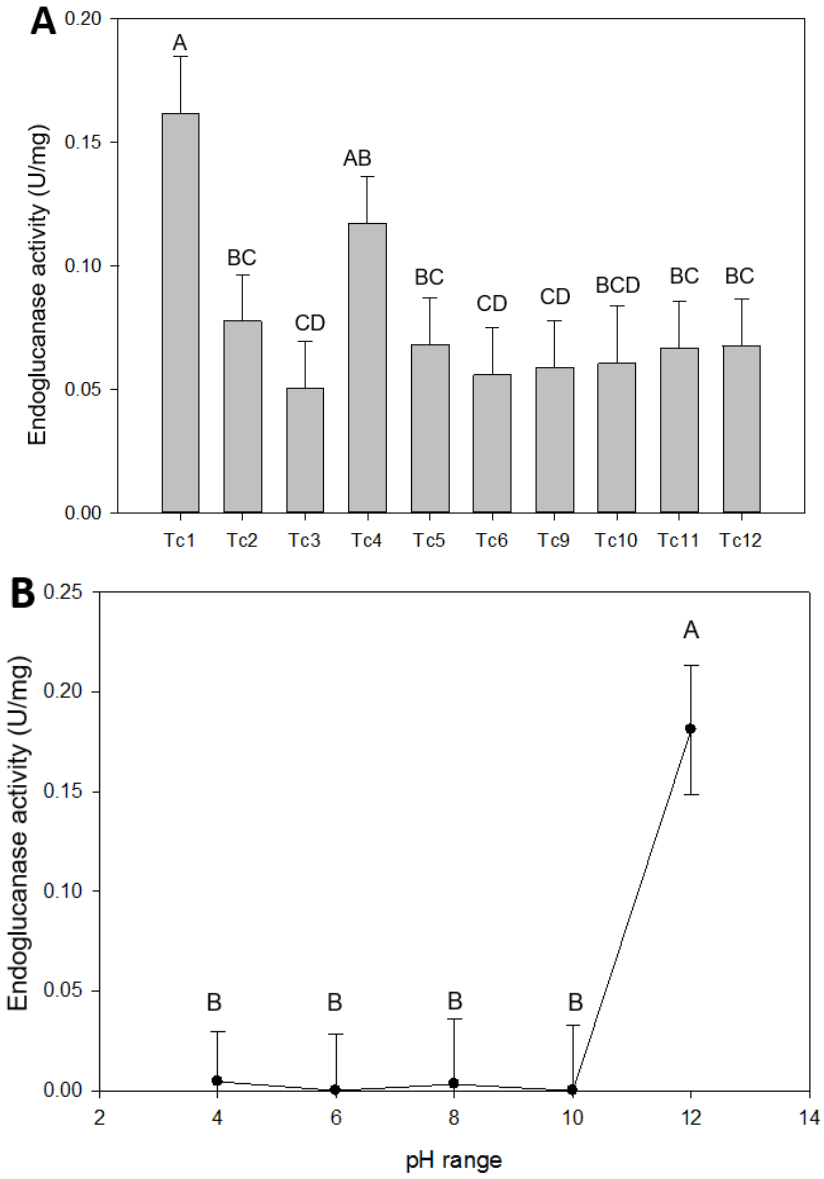

Figure 3-2 continued

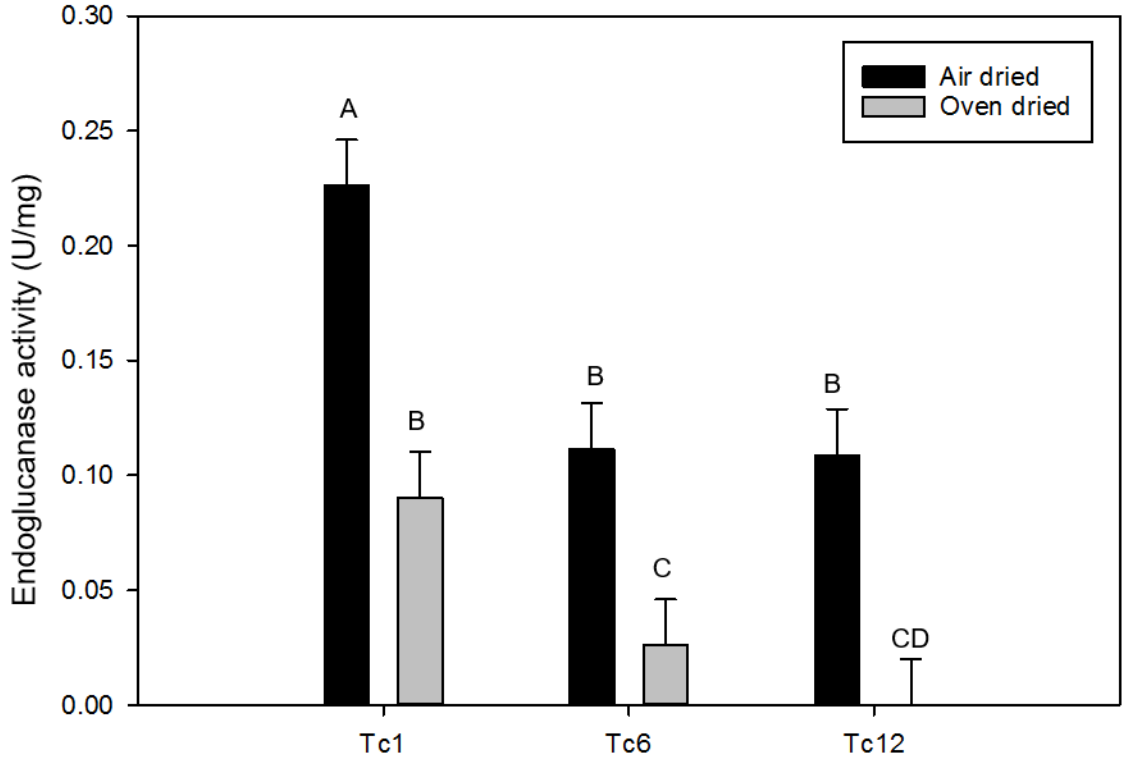

Figure 3-3 Endoglucanase activity (units/mg of protein) from leaves of transgenic TcEG-1 plants using carboxymethyl cellulose (CMC) as substrate at pH 12.0. Leaves were either air dried for two weeks in the greenhouse (black bars) or dried for three days in an oven at 46 °C (grey bars). Bars represent mean values of three replicates ± standard error. Bars represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$).

Figure 3-4 Glucose (A), xylose (B), and total sugar (C) release from transgenic TcEG-1 and non-transgenic (NT-Perf) tillers as determined by enzymatic hydrolysis. Bars represent mean values of three replicates ± standard error. Bars represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$).

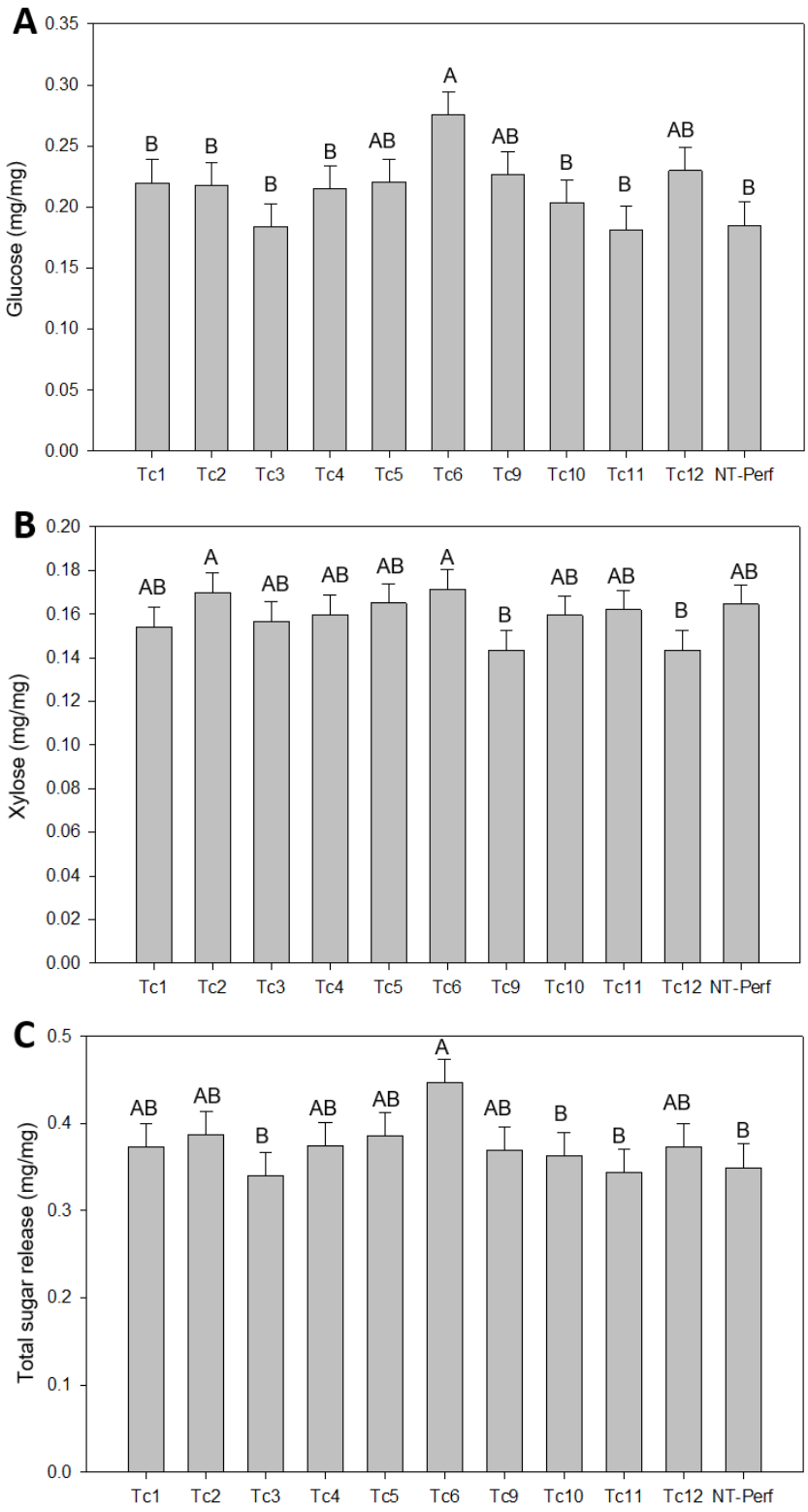

Figure 3-4 continued

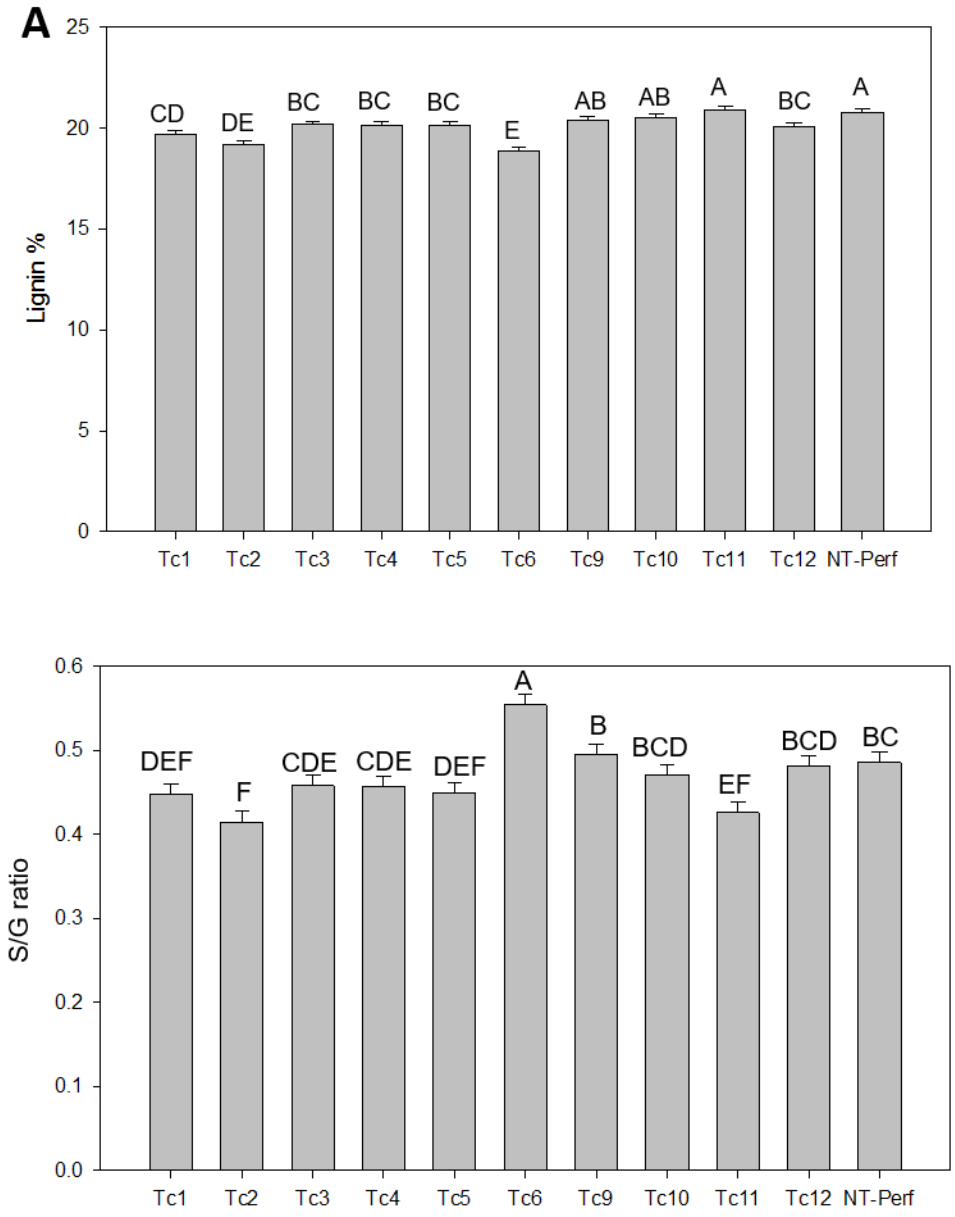

Figure 3-5 Lignin content (A) and S/G ratio (B) of transgenic TcEG-1 and non-transgenic (NT-Perf) tillers as determined by Py-MBMS. Bars represent mean values of three replicates ± standard error. Bars represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$).

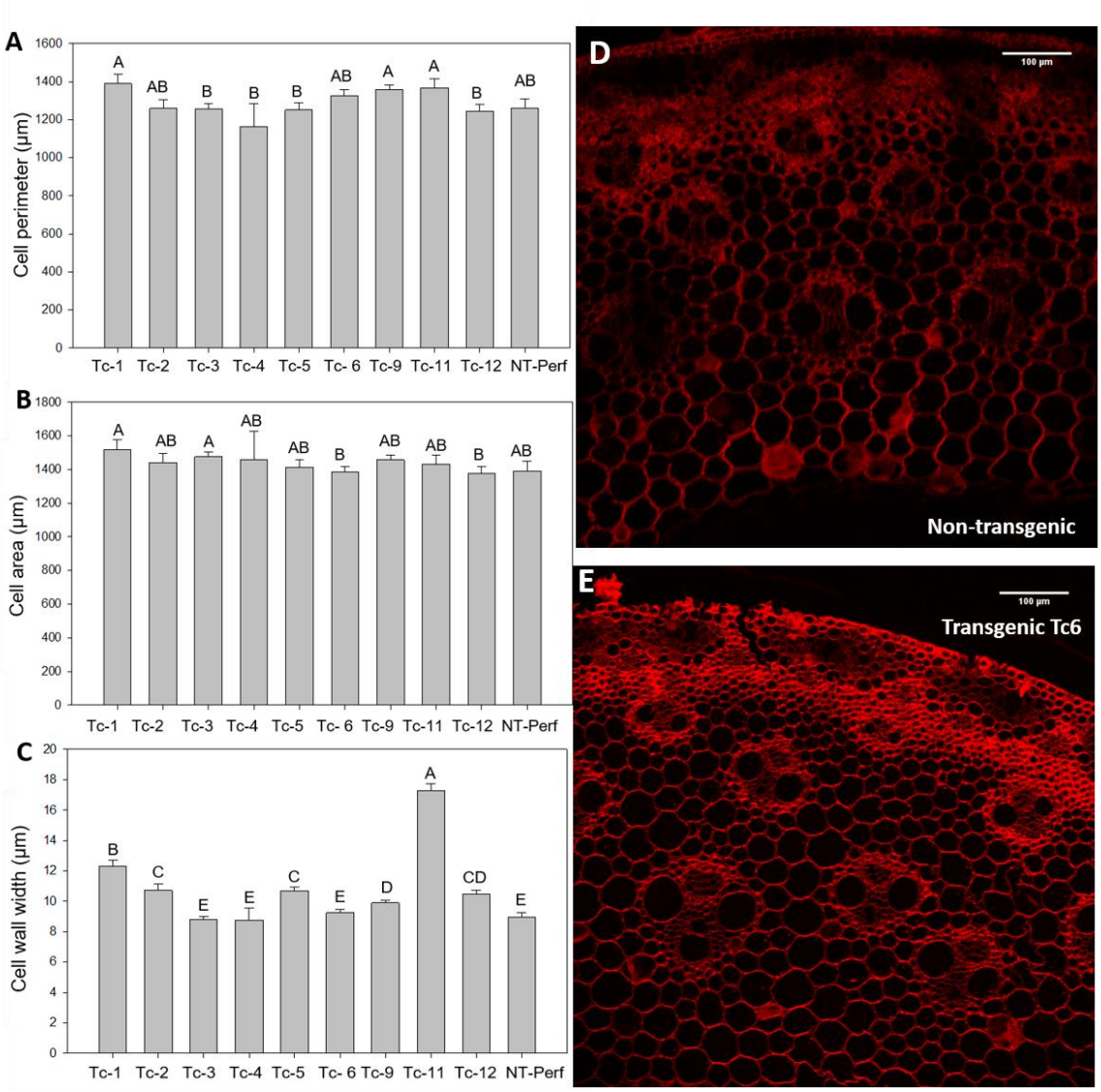

Figure 3-6 Cell wall measurements on histological analysis of stem internode sections of transgenic TcEG-1 and non-transgenic (NT-Perf) plants. Measurement of cell wall perimeters (A), cell wall thickness (B), and cell wall areas (C). Representative pictures of non-transgenic (D) and transgenic event Tc-6 (E) stem internodes stained with Pontamine Fast Scarlet. Bars represent mean value of replicates ± standard error. Bars represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$). Scale bar represents 100 µm.

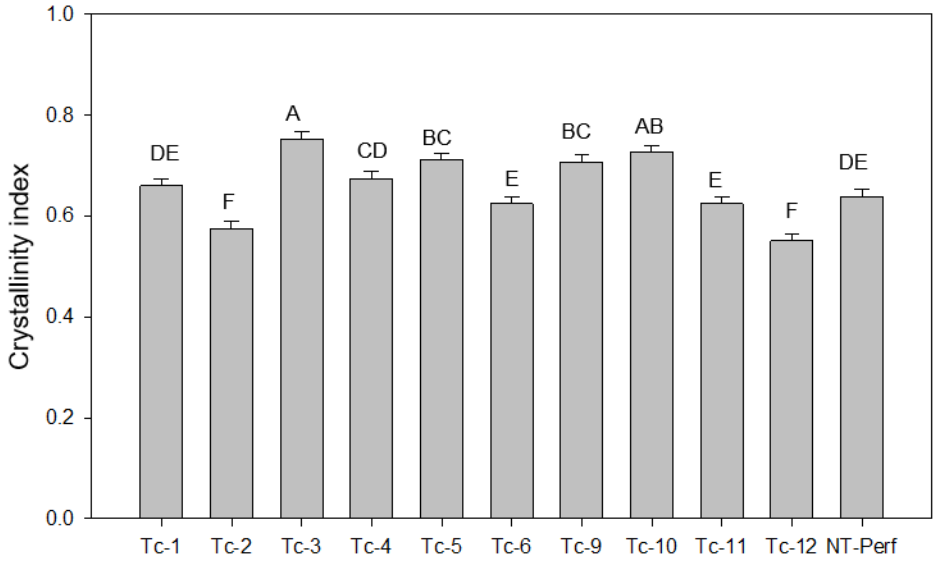

Figure 3-7 Cellulose crystallinity index measurements for transgenic TcEG-1 and non-transgenic (NT-Perf) plants. Bars represent mean values of three replicates ± standard error. Bars represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$).

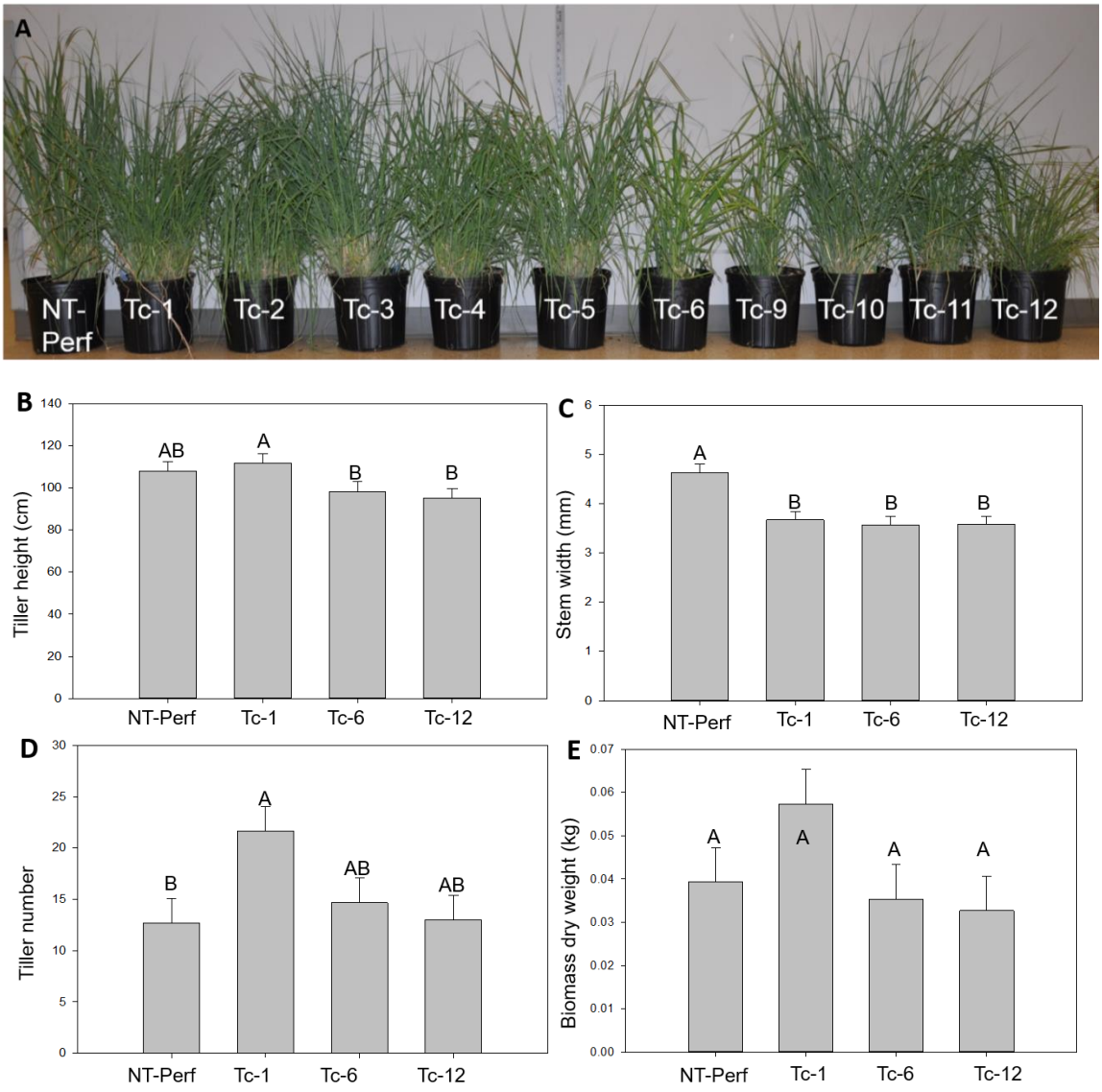

Figure 3-8 Plant morphology analysis of transgenic TcEG-1 and non-transgenic switchgrass plants. A) Representative transgenic *TcEG-1* and non-transgenic (NT-Perf) lines. Tiller height (B), stem width taken at 10 cm height above potting mixture (C), tiller number (D), and biomass dry weight (E) of transgenic *TcEG-1* and non-transgenic (NT-Perf) plants. Bars represent mean values of three replicates ± standard error. Bars represented by different letters are significantly different as calculated by LSD ($p \leq 0.05$).

3.9.2 Supplementary table and figures

Table 3-S1: Amino acid pBLAST of TcEG-1 protein to switchgrass proteome. Switchgrass protein is the protein matched by the search of the TcEG-1 amino acid against the switchgrass proteome database. %ID is the percent identical. Length is the amino acid alignment coverage of TcEG-1 the switchgrass protein. Mismatches represents the lack of match of two protein sequences. Gaps are the openings in the alignment represent where no matches occurred. E value represents initial score value determined by (https://phytozome.jgi.doe.gov/pz/portal.html). Bit score represents log transformed score based on (https://phytozome.jgi.doe.gov/pz/portal.html).

Switchgrass protein	%ID	Length	Mismatches	Gaps	E value	bit score
Pavir.J22089.1	38.05	452	252	10	3.00E-86	276
Pavir.Fa01623.1	38.16	456	237	9	4.00E-86	277
Pavir.Ea00142.1	37.13	474	247	13	4.00E-86	281
Pavir.Aa00742.1	39.57	460	228	13	1.00E-84	274
Pavir.Ca00745.1	36.23	461	251	8	3.00E-84	271
Pavir.Eb00189.1	36.79	473	248	11	4.00E-84	275
Pavir.Ga00244.1	38.84	466	236	13	3.00E-83	272
Pavir.Ab02908.1	37.96	461	247	13	2.00E-81	264
Pavir.Ab02962.1	38.36	464	228	12	3.00E-81	265
Pavir.Aa00400.1	37.74	461	248	12	5.00E-81	263
Pavir.Cb01721.1	36.81	470	254	11	1.00E-80	266
Pavir.Aa03559.1	36.74	479	256	10	4.00E-80	261
Pavir.J05126.1	37.25	459	248	11	4.00E-79	258
Pavir.Bb02664.2	38.53	462	242	12	5.00E-79	258
Pavir.Ab00278.1	36.4	489	254	12	5.00E-78	258
Pavir.Bb02664.1	38.66	463	240	12	2.00E-77	257
Pavir.Aa00303.1	36.82	459	250	11	9.00E-77	252
Pavir.J35890.1	35.65	474	265	10	1.00E-75	249
Pavir.Fa01564.1	36.47	447	241	12	3.00E-75	248
Pavir.Fb01149.1	36.64	453	244	12	4.00E-75	249
Pavir.Db01736.1	37.34	466	249	12	1.00E-74	247
Pavir.Ca00497.1	35.96	470	252	12	4.00E-74	248
Pavir.Bb02635.1	34.62	465	260	12	6.00E-74	245

Table 3-S1 continued

Switchgrass protein	%ID	Length	Mismatches	Gaps	E value	bit score
Pavir.Db00162.1	36.44	461	256	10	6.00E-73	243
Pavir.Gb00020.2	36.23	472	239	13	9.00E-73	247
Pavir.Gb00020.2	29.47	190	105	5	1.00E-14	77.8
Pavir.Gb00020.1	36.23	472	239	13	1.00E-72	247
Pavir.Gb00020.1	29.47	190	105	5	1.00E-14	77.8
Pavir.Ea01014.1	33.89	478	258	9	3.00E-71	238
Pavir.Ba01232.2	38.6	399	209	8	2.00E-67	229
Pavir.Ba01232.1	38.6	399	209	8	2.00E-67	229
Pavir.Fb00818.1	33.75	477	250	13	9.00E-66	223
Pavir.J09205.1	35.26	397	226	6	9.00E-65	219
Pavir.Da01211.1	36.43	420	214	11	2.00E-64	218
Pavir.Bb01897.1	46.77	263	138	2	6.00E-64	213
Pavir.Fa02209.1	32.83	463	270	10	8.00E-64	218
Pavir.J03155.1	35.85	463	243	13	9.00E-64	218
Pavir.J36442.1	35.84	413	219	11	6.00E-63	214
Pavir.J35848.1	32.55	467	272	11	1.00E-62	215
Pavir.J01061.1	31.75	463	274	10	1.00E-61	212
Pavir.Ia00651.1	34.02	485	257	16	4.00E-60	211
Pavir.Ib04298.1	33.81	485	258	16	2.00E-59	208
Pavir.Ia03435.1	34.77	463	232	16	1.00E-57	202
Pavir.Bb02664.3	44.57	258	141	2	1.00E-55	193
Pavir.Ga01268.2	33.2	485	254	16	3.00E-55	197
Pavir.Ga01268.1	33.2	485	254	16	5.00E-55	196
Pavir.Ga01270.1	31.79	497	249	19	9.00E-51	184
Pavir.Gb01448.1	31.83	487	261	17	1.00E-48	177
Pavir.Ia00651.2	32.95	440	237	14	5.00E-48	175
Pavir.J07817.1	43.95	223	120	3	1.00E-41	157
Pavir.J07817.1	27.67	159	93	5	5.00E-04	43.9
Pavir.Ib04298.2	37.65	247	133	6	6.00E-37	142
Pavir.Ib04298.3	37.65	247	133	6	8.00E-37	143
Pavir.J06476.1	30.4	454	250	12	2.00E-35	138
Pavir.J31800.1	38.87	247	130	7	1.00E-34	135
Pavir.J16274.1	33.9	351	179	12	2.00E-33	132
Pavir.J32907.1	48.7	115	59	0	1.00E-31	121
Pavir.Ib01536.1	43.64	110	57	2	4.00E-22	96.3
Pavir.Ga01737.1	26.69	281	161	11	3.00E-20	92
Pavir.Ba02164.1	31.18	93	54	4	1.00E-05	47.4
Pavir.Eb02381.1	38.78	49	23	2	0.033	35.8

Figure 3-S1. Stepwise demonstration of custom Python Cell Wall Thickness (pyCWT) program to identify sclerenchyma and parenchyma cells within a slide section specimen and determine output. A) Slide section images are converted to grey scale where cells are identified and marked with a red dot. A combined contours and binary erosion technique smooths out any imperfections in the image. During post-dilation, cells having their thickness labeled are colored a different shade of grey. B) Example output graph of pyCWT. pyCWT calculates the number of dilations required for two cells to overlap and records it. Line graph displays the number of cell walls imaged against the distribution in pixel sizes. The mode was chosen as the most representative cell wall thickness for the image. The mode for each image is converted to micrometers by a separate script which extracts the micrometer size of each image from a separate text file and then performs the operation to convert pixels into micrometers for each image individually. The final output is placed into a tab-separated values file, with the following columns: image name, mode (for entire image), cell identifier, cell wall thickness, area, perimeter, and pixel-to-micron factor.

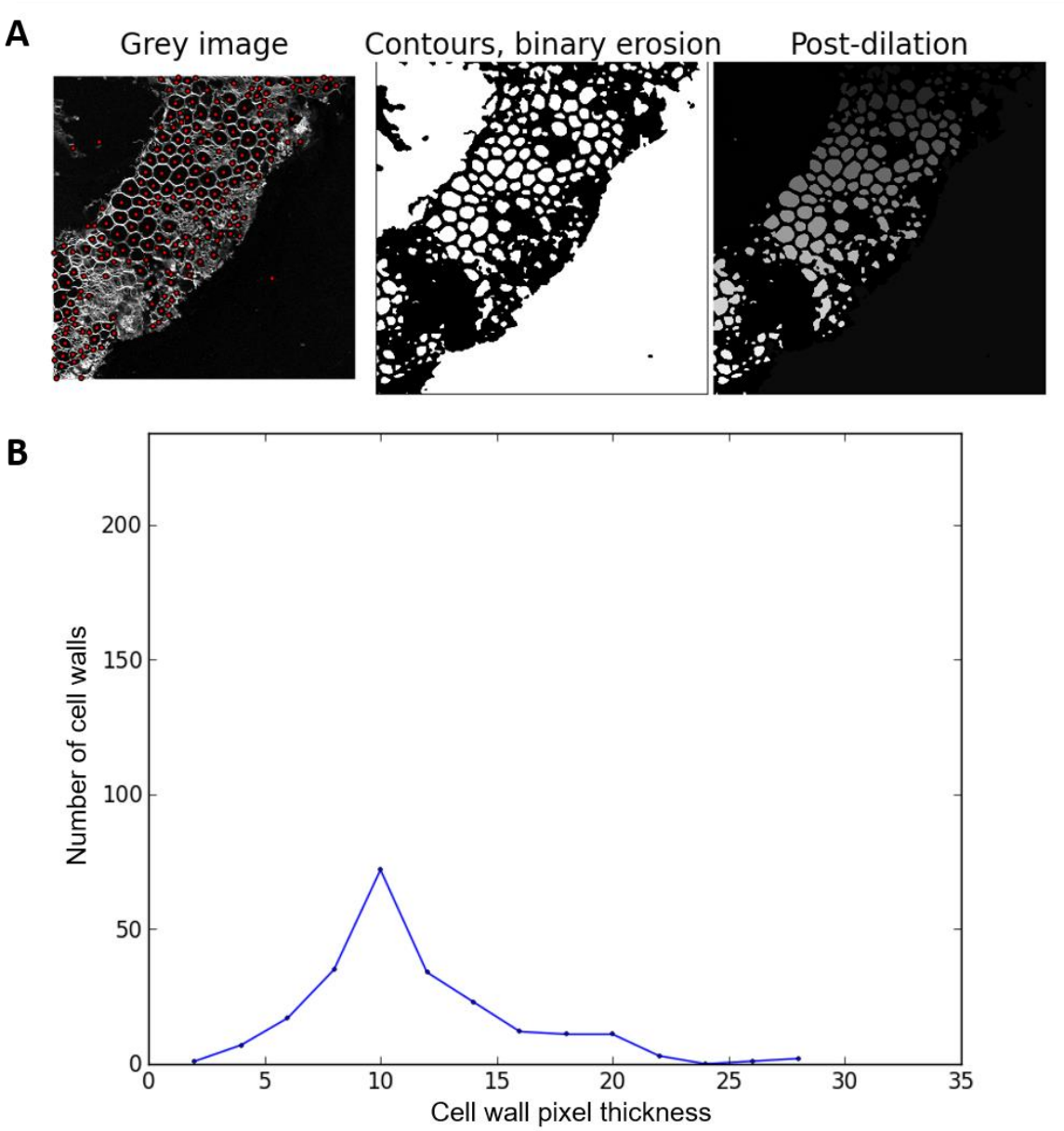

Figure 3-S1 continued

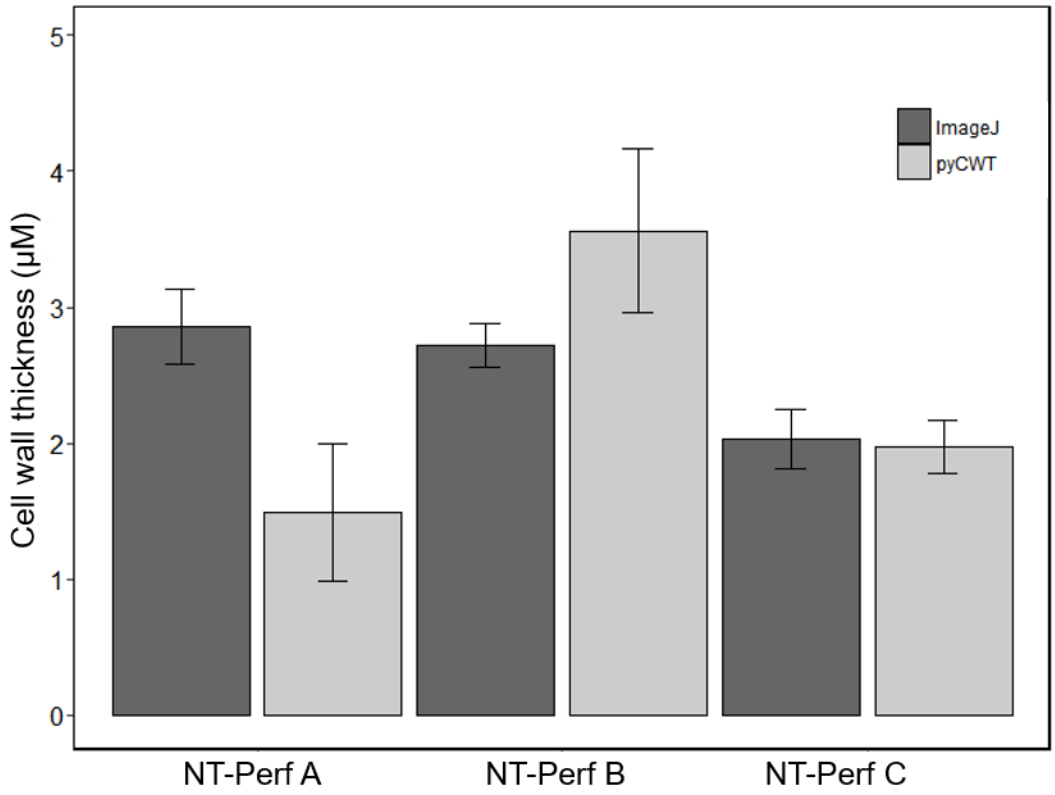

Figure 3-S2 Comparison ImageJ manual measuring vs Python Cell Wall Thickness (pyCWT) program measuring of cell wall thickness measurement methods. Three non-transgenic switchgrass stem internodes were imaged and cell wall thickness determined by either manual measuring of cells using ImageJ or pyCWT. Python program generated values did not differ from the hand measured values when compared via t-test at ($p < 0.05$). Standard error is shown.

Chapter 4: Downregulation of UDP-arabinomutase gene in switchgrass (*Panicum virgatum* L.) results in increased cell wall lignin and glucose with reduced arabinose

A version of this chapter will be submitted to Frontiers in Plant Sciences by Dr. C. Neal Stewart Jr. and Dr. Maor Bar-Peled with following authors Jonathan D. Willis, James A. Smith, Mitra Mazarei, Jiyi Zhang, Geoffrey B. Turner, Stephen R. Decker, Robert W. Sykes, Charleson R. Poovaiah, Holly L. Baxter, Dave G. J. Mann, Mark F. Davis, Michael K. Udvardi, Maria J. Pena, Jason Backe, Maor Bar-Peled and C. Neal Stewart Jr. Jonathan D. Willis drafted the manuscript, generated the majority of the transgenic plants, performed the statistical analysis, performed Southern blotting and qRT-PCR analysis, and prepared plant samples for recalcitrance analysis. James A. Smith produced the sugar profile data, collected samples for repeated recalcitrance analysis, ran NMR analysis, and contributed to manuscript drafting. Mitra Mazarei participated in experimental design and data analysis, assisted with revisions to the manuscript and coordination of the study. Jiyi Zhang and Michael K. Udvardi assisted with cloning of the target gene. Geoffrey B. Turner, Stephen R. Decker,Robert W. Sykes, and Mark F. Davis assisted with performing lignin and sugar release assays. Charelson R. Poovaiah and Holly L. Baxter contributed to tissue culture and generation of transgenic plants. David G.J. Mann contributed to initialization of the research and early oversight of the project. Maria J. Pena and Jason Backe assisted with NMR and sample preparation. Maor Bar-Peled and C. Neal Stewart conceived of the study and its design and coordination. All authors contributed to text and data analysis and interpretation. All authors read and approve final version of the manuscript

4.1 Abstract

Background: Switchgrass (*Panicum virgatum* L.) is a C_4 perennial prairie grass and a dedicated feedstock for lignocellulosic biofuels. Saccharification and biofuel yields are inhibited by the plant cell wall's natural recalcitrance against enzymatic degradation. Plant hemicellulose

polysaccharides such as arabinoxylans structurally support and cross link other cell wall polymers. Grasses predominately have Type II cell walls that are abundant in arabinoxylan, which comprise nearly 25% of aboveground biomass. A primary component of arabinoxylan synthesis is uridine diphosphate (UDP) linked to arabinofuranose (Ara*f*). A family of UDP-arabinopyranose mutase/reversible glycosylated polypeptides (UAM/RGPs) catalyze the interconversion between UDP-arabinopyranose (UDP-Ara*p*) and UDP-Ara*f*.

Results: The expression of a switchgrass arabinoxylan biosynthesis pathway gene, *PvUAM1*, was decreased via RNAi to investigate its role in cell wall recalcitrance in the feedstock. *PvUAM1* encodes a switchgrass homolog of UDP-arabinose mutase, which converts UDP-arabinopyranose to UDP-arabinofuranose. Southern blot analysis revealed each transgenic line contained between one to at least seven T-DNA insertions, resulting in some cases, a 95% reduction of native *PvUAM1* transcript in stem internodes. Transgenic plants had increased pigmentation in vascular tissues at nodes, but were otherwise similar in morphology to the non-transgenic control. Cell wall-associated arabinose was decreased in leaves and stems by over 50%, but there was an increase in cellulose, reflective of increased cell wall-associated glucose. In addition, there was a commensurate change in arabinose side chain extension. Cell wall lignin composition was altered with a concurrent increase in lignin content and transcript abundance of lignin biosynthetic genes in mature tillers. Enzymatic saccharification efficiency was unchanged in the transgenic plants relative to the control.

Conclusions: Plants with attenuated *PvUAM1* transcript had increased cell wall-associated glucose and lignin in cell walls. The increased cell wall-associated glucose in stems and leaves indicates that attenuation of *PvUAM1* expression might have downstream effects on cellulose and xyloglucan biosynthesis. A decrease in cell wall-associated arabinose was expected, which was likely caused by fewer Ara*f* residues in the arabinoxylan. The decrease in arabinoxylan may cause a compensation response to maintain cell wall integrity by increasing cellulose and lignin biosynthesis. In cases in which increased lignin is desired, e.g., feedstocks for carbon fiber production, down-regulated *UAM1* coupled with altered expression of other arabinoxylan biosynthesis genes might result in even higher production of lignin in biomass.
Keywords: Switchgrass, hemicellulose arabinoxylan, UDP-arabinopyranose mutase/reversible glycosylated polypeptide, biofuel, recalcitrance

4.2 Background

Switchgrass (*Panicum virgatum*) is a perennial grass species that is considered to be a lignocellulosic bioenergy feedstock with great potential, owing to its wide adaptations to various geographies and temperate climates. Recalcitrance, which is the inherent resistance of cell wall polysaccharides to be digested into fermentable sugars, is a sizeable economic barrier to lignocellulosic biofuel production. At the center of recalcitrance is the heterogeneous composition of plant cell walls, which are made of three main types of polymers: cellulose, lignin, and hemicellulose (Dixon, 2013). Feedstock genomics and biotechnology have enabled a better understanding of cell wall recalcitrance, including that for switchgrass (Casler et al., 2011; Chen et al., 2016). Relatively few studies in which hemicellulose has been manipulated were carried out to determine its role in cell wall recalcitrance in biofuel crops (Vega-Sanchez and Ronald, 2010).

In plants, hemicelluloses are comprised of non-cellulose cell wall polysaccharides, and share a sugar backbone composed of 1,4-linked β-D-glycoses and include xyloglucan, xylan, and glucomannan. Xylan itself constitutes a sub-grouping of polysaccharides whose members are distinguished from one another by the types of oligosaccharide side chains linked to the 1,4-linked β-D-xylopyranose backbone (Rennie and Scheller, 2014; York and O'Neill, 2008; Scheller and Ulvskov, 2010). The backbone of xylans isolated from grasses for example, is decorated by a large number of L-arabinose residues (found only in furanose form, Ara*f*), and hence referred to as arabinoxylans. The Ara*f* residues are 1α-3 and/or 1α-2 O-linked to various xylose residues in the backbone, and a portion of these side chains are extended with an additional (1,2)- α-linked

Ara*f*. Xylan may also contain non-carbohydrate modification of *O*-acetyl esters and methyl etherified sugars, as well as feruloyl, and *p*-coumaroyl moieties (Bar-Peled and O'Neill, 2011; Faik, 2010). For example, the aromatic residues (feruloyl and *p*-coumaroyl) can be ester-linked to O-5 of terminal arabinose residues of xylan, whereas the acetate can be O-2 or O-3 linked to xylose in the backbone. In grass species, ferulic acid is ester-linked to C5 of Ara*f* in arabinoxylan and in ether linkages of lignin monomers (Hartley and Ford, 1989; Scalbert et al., 1985). Although the role of feruloylation is not well understood, an increase in ferulic acid modification of arabinoxylan has been associated with cells that have stopped elongating (Carpita, 1986). Feruloylation has been hypothesized to prime polymerization of lignin thereby interconnecting a network of xylan and lignin (De O. Buanafina 2009; Iiyama et al., 1994). Additionally, adjacent arabinoxylan chains decorated with ferulic acid can dimerize through oxidative coupling, which may condense wall polymers into a tightly packed matrix enhancing the walls stability and resistance to degradation (Hatfield et al., 1999). Disruption of these ether linkages between arabinoxylan and lignin is an inviting target for improving cell wall degradation.

The diversity in xylan structures is known, but the functional role for such diversity is largely unknown. For example, it is not understood why xylan chemotypes differ among tissues in the same plants. It was proposed that xylan cross-links to cellulose and lignin, which serves to strengthen cell walls (Scheller and Ulvskov, 2010). Arabinoxylans comprise over 25% of the mass of grass cell walls (Faik, 2010; Konishi et al., 2011). The formation of arabinoxylan requires the building blocks UDP-xylose and UDP-arabinofuranose (UDP-Ara*f*). UDP-

arabinopyranose mutase (UAM) converts UDP-arabinose (pyranose-form, Ara*p*) to UDP-Ara*f* (Konishi et al., 2010; Konishi et al., 2007). UAM orthologs are found in some microalgae and land plants in which they comprise a small gene family (Kotani et al., 2013). Interestingly, UAM can also reversibly glycosylate itself in the presence of UDP-sugars, such as UDP-glucose, UDP-galactose and UDP-xylose (hence the name RGP) (Dhugga et al., 1991; Konishi et al., 2010; Konishi et al., 2007; Rautengarten et al., 2011). The role of UAM as an RGP is not well understood in the context of cell wall and glycan formation. It was hypothesized that the RGP function of UAM may regulate the internal balance of UDP-sugars in the cell or compete for the formation of UDP- Ara*f*, and this hypothesis has been explored in *Arabidopsis*, algae, *Brachypodium*, and rice (Konishi et al., 2011; Kotani et al., 2013; Rancour et al., 2015; Rautengarten et al., 2011). However, UAM's potential role in recalcitrance has never been examined nor manipulated in any bioenergy feedstock.

In this study, it was hypothesized that manipulation of the level of UDP-Ara*f* in cells will alter the amount of arabinoxylan in switchgrass cell walls, and potentially alter feedstock recalcitrance. In this study, a switchgrass *UAM1* homolog (*PvUAM1*), was down-regulated in independent transgenic lines of switchgrass, wherein cell wall composition, saccharification, and plant growth were analyzed.

4.3 Methods

4.3.1 PvUAM1 gene isolation, and RNAi construct

The amino acid sequence of switchgrass UAM protein was compared with UAM orthologues from eudicots and monocots: SiUAM1 (XP004982467.1), ZmUAM1 (NP001105598.1), SbUAM1 (XP002464260.1), OsUAM1 (XP006650286.1), BdUAM1 (XP003562308.1), TaUAM1 (CAA77237.1), SlUAM1 (NP001234554.1), VvUAM1 (XP002263490.1), BdUAM1 (XP003569874.1), GmUAM1 (XP003552602.1), BrUAM1 (XP009117866.1), AtUAM3 (AAM65020.1), AtUAM1 (AT3G02230.1), PtUAM1 (Potri.004G117800.1), MtUAM1 (Medtr5g046030.1), OsUAM1 (Q8H8T0.1), OsUAM3 (Q6Z4G3.1), OsUAM2 (Q7FAY6.10), AtUAM2 (NP197069.1), EgUAM1 (AGE46030.1), PdUAM1 (XP008811806.1). The sequences among UAM proteins were compared using alignment in the MUSCLE program (http://www.ebi.ac.uk/Tools/mas/muscle/) and alignment curated by Gblocks at the plhyologeny.fr (http://www.phylogeny.lirmm.fr) software program (Anisimova and Gascuel, 2006; Dereeper et al., 2008). The neighbor joining tree was generated using the MEGA 7.0 program (Tamura et al., 2013). The switchgrass *PvUAM1, PvUAM2, PvUAM3,* gene sequences were identified by BLASTN analysis of the switchgrass genome (https://phytozome.jgi.doe.gov/pz/portal.html) using the monocot UAM sequences from maize (GI: 542592), foxtail millet (GI: 101771463), and sorghum (GI: 8062976). The nucleotide coding sequence of the *PvUAM1* open reading frame was identified and a 193 bp target sequence was used to generate the RNAi plasmid construct (Fig S1). The target sequence was amplified by PCR and was cloned into the pCR8 entry vector and then Gateway® sub-cloned into the pANIC-

8A plant expression vector (Mann et al., 2012b) to yield the pANIC-8A-PvUAM1 construct (Fig S2A).

4.3.2 Transgenic plant production and growth analysis

Inflorescences of the 'Alamo' switchgrass 'ST1' genotype was used to generate Type II embryogenic callus (Burris et al., 2009). *Agrobacterium tumefaciens* strain EHA105 harboring the pANIC-8A-PvUAM1 vector was used for transformation. Transformed calli were grown on agar containing LP9 growth medium (Burris et al., 2009), supplemented with 400 mg/L Timentin and 40 mg/L hygromycin for approximately two months at 25 ° C in the dark. Subsequently the transgenic calli were transferred to regeneration medium as described by Li and Qu (2011) and was supplemented with 250 mg/L cefotaxime to stimulate regeneration (Danilova and Dolgikh, 2004). The T-DNA region of pANIC-8A-PvUAM1 plasmid also contains a cassette that constitutively expresses an orange fluorescence protein (OFP) reporter from the hard coral *Porites porites* (pporRFP) that is brightly fluorescent in transgenic plants (Mann et al., 2012a). Epi-fluorescence microscope having a 535/30 nm excitation filter and 600/50 nm emissions filter was used to track OFP fluorescence during transgenic callus development and to identify individual putative transgenic lines during growth on agar-plate. Regenerated transgenic plants were rooted and acclimated according to Burris et al. (2009).

T0 transgenic and non-transgenic ST1 control plants were grown in growth chambers under 16 h light/8h dark cycles at 25°C until moved to a greenhouse. Fertilizer (0.02% solution of Peter's soluble 20-20-20) was applied twice per month and plants were irrigated as needed. For growth

analysis, each transgenic plant line and the non-transgenic control was vegetatively divided into three clonal replicates. Each replicate, starting from a single-tiller, was grown in a 1 L pot that was randomly sited in the greenhouse. Plants were grown until the reproductive (R1) developmental stage as defined by Hardin et al. (2013) and tiller and panicle numbers were counted. The five tallest tillers for each replicate were used to estimate total plant height. The stem width at 10 cm from the potting surface of each of these tillers was measured with a digital caliper. Tillers were harvested and green fresh weight was recorded. Harvested tillers were placed into a drying oven at 42 °C for five days and dry weight was subsequently recorded. Hand sectioning was performed on fresh tillers and nodal sections to assess vascular phenotypes under a dissecting microscope and to depict deposition of pigment.

4.3.3 Southern blot analysis for T-DNA copy number

Approximately 100 mg of young (1 week-old from recently cut-back plants) fresh leaf tissue per plant was used to extract DNA (Freeling and Walbot, 1994). DNA quality was assessed using gel electrophoresis and quantified using a Nanodrop spectrofluorometer (Thermo Fisher, Wilmington, Del., USA). Twenty micrograms of DNA from each sample was digested with *NcoI,* which cuts once within the T-DNA. Digested DNA from transgenic plants and the control, was separated on a TAE-agarose gel, and transferred to a nylon membrane (Amersham Hybond™+ GE Healthcare, Pittsburgh, PA., USA). Blots were pre-hybridized with DIG easy hyb (Roche DIG kit, Nutley, NJ., USA) solution at 42 °C. The blots were then hybridized with the hygromycin DIG-PCR probe, washed, and probe was detected after the membrane was exposed to x-ray film according to manufacturer's instructions (Roche). The DNA probe (972

bp) used to detect the number of hygromycin *(hph*) gene cassette in DNA from transgene lines was amplified by PCR and labeled with digoxigenin (Roche DIG kit, Nutley, N.J., USA).

4.3.4 RNA extraction, qRT-PCR analysis of UAM's and lignin biosynthetic gene transcripts.

Quantitative RT-PCR was performed to estimate transcript abundance of *PvUAM* and lignin biosynthetic genes in transgenic *PvUAM1*-RNAi and non-transgenic plants. Total RNA was isolated from triplicate R1 tiller stem internodes and leaf cuttings using TRI Reagent following manufacturer's instructions (Sigma-Aldrich, St. Louis, MO). Purified RNA was treated with DNase-1 (Promega, Madison, WI, USA) and 3 µg treated RNA was used to generate cDNA using oligo-dT and Superscript III according to manufacturer's instructions (Life Technologies, Carlsbad, CA, USA). qRT-PCR analysis was performed with Power SYBR Green PCR master mix (Life Technologies, Carlsbad, CA, USA) according to manufacturer's protocols for optimization of annealing temperature, primer concentration, and cDNA concentration. Primers used for transcript analysis of *PvUAM* are listed in Table S1 and for lignin biosynthetic genes in Table S2. The optimized qRT-PCR protocol utilized a dilution of cDNA 1:100 with thermal cycling at 95 °C for 3 min, and 40 cycle repeats of (95 °C for 10 s and 60.0 °C for 30 s). The relative levels of transcripts were normalized to switchgrass ubiquitin 1 (*PvUbi1*) as a reference gene (Shen et al., 2009) using primer set: PvUBi1_F 5'- CAGCGAGGGCTCAATAATTCCA - 3' and PvUbi1 R 5' - TCTGGCGGACTACAATATCCA - 3' (Xu et al., 2011). All experiments were carried out in triplicate technical replicates. The differential Ct method was used to measure transcript abundance after normalization to *PvUbi1* (Schmittgen and Livak, 2008).

4.3.5 Total wall polysaccharide isolation and GLC analysis

Tillers at the R1 developmental stage were collected from a single plant grown in a greenhouse for approximately 6-8 weeks. The tillers were cut and divided into stem and leaf sections. A sample section was weighed, ground in liquid nitrogen, and washed as previously reported (Martinez et al., 2012) with slight modifications. Each 1 g sample was suspended in 10 mL 80% EtOH, vortexed for 2 min, then centrifuged (6,000 x g 5 min, 25 °C). The supernatant was removed, and the resulting cell pellet was washed two times each with 10 ml 95% EtOH, and then with 10 ml 100% EtOH. The cell pellet was resuspended in 10 mL chloroform:MeOH 1:1 (v/v) and mixed by tilting for 1 hr. Each cell pellet residue sample was filtered through Whatmann # 15 filter paper over vacuum and rinsed with acetone. Once dry, the alcohol-insoluble residue (AIR) samples were weighed and passed through a 0.5 mm mesh. The polysaccharides in each 1 mg AIR sample were hydrolyzed with 2N trifluoroacetic acid and the free monosaccharides were converted to their alditol acetate derivatives as previously described (York et al, 1986). All samples (including sugar standards) were supplemented with 50 µL 5 mM inositol as an internal standard. Alditol acetate sample or standard (1 µL) was separated on a Restek RTx-2330 fused silica column (0.25 mm I.D. x 30 m, 0.2 µm film thickness) using an Agilent 7890A GLC equipped with an FID as a detector. Relative molar percent content was calculated from the areas of sugar peaks identified by standard retention times and normalized to sample mass and internal standard.

4.3.6 Preparation, solubilization, and fractionation of wall polysaccharides and oligosaccharides

AIR sample (250 mg) was suspended in 25 mL buffer (0.1 M sodium acetate, 0.01% Thimerosal, pH 5.0) and supplemented with amylase mixture 30 µL Spirizyme Excel (Novozymes# NAPFM084) and 150 µL Liquozyme SC DS (Novozymes# AUP61163), as described by Decker, (2012). Starch digestion was carried out at 55 °C overnight and subsequently the slurry sample was filtered through a double filter layers (50 µm nylon mesh on top of Whatman Grade GF/A filter paper), washed with buffer, and destarched insoluble residue was reserved. To remove loosely bound pectin, the destarched AIR sample was resuspended in 25 ml oxalate solvent (0.5% ammonium oxalate, 0.01% Thimerosal, pH 5.0) and shaken overnight at room temperature. The slurry was filtered through a double filter layers, washed with oxalate, and the insoluble residue was reserved. Each oxalate-treated AIR sample residue was resuspended in 25 ml 1M base solution (1 M KOH, 1% $NaBH_4$) and shaken overnight at room temperature. The slurry was filtered through a double filter layers. The filtrate that contained soluble polysaccharide was reserved and the insoluble residue slurry on top of double layer filter was wash with 1M KOH and insoluble pellet was reserved. The filtrate and wash 1M KOH-solubilized wall polymer were combined, neutralized to pH 7 with glacial acetic acid (by pH paper); supplemented with a drop of octanol; and later dialyzed. The insoluble residue after 1 M KOH treatment was resuspended in 25 ml 4M base solution (4 M KOH, 1% $NaBH_4$), shaken overnight at room temperature, filtered, and the solubilized fraction was pH neutralized as described above. Both the 1 and 4M KOH solubilized soluble extracted polysaccharide fractions were dialyzed against deionized water (3500 MWCO) for 2 to 3 days. The dialyzed KOH

fractions were centrifuged (11,000 x g for 30 min, 25 °C), concentrated by Rotovac, lyophilized, and used for NMR and generation of oligosaccharides. The 4M-KOH insoluble residue (enriched in unfractionated cellulose) was stored at -20 °C for further analysis of cellulose.

4.3.7 Preparation of arabinoxylooligosaccharides and NMR analysis

Between 5–20 mg of 1 M KOH soluble and dialyzed fraction (above) was dissolved in 1-5 mL of 50 mM ammonium formate, pH 5.0. One unit of endoxylanase (from *Trichoderma viride*) was added and the solution incubated at 37 °C for 24 h. Hydrolase activity was terminated by boiling for 10 min in a water bath and the sample was centrifuged at 3,600 x g for 15 min at room temperature. The supernatant was transferred to a tube and lyophilized. A portion of freeze-dried arabinoxylooligomers (1-2 mg) was dissolved in 0.5 mL deuterium oxide (99.9 %; Cambridge Isotope Laboratories) and supplemented with 1 μL acetone that was used as an internal chemical shift reference. One-dimensional and two-dimensional ^{1}H NMR spectra were collected on a 600 MHz Varian Inova NMR spectrometer equipped with a 3-mm cold probe and a sample temperature of 25 °C. Data were processed with MestRe-C, (Universidad de Santiago de Compostela, Spain). All chemical shifts were calibrated relative to acetone (δ H 2.225 Hz/MHz).

4.3.8 Cellulose quantification

Thirty milligrams of unfractionated cellulose was weighed into a conical borosilicate tube with Teflon-lined screw cap, and 3 mL solvent (acetic acid/water/nitric acid, 8/2/1, v/v/v) was added (Updegraff, 1969). The sample was vortexed, heated in a boiling water bath for 30 min with

occasional mixing; cooled to room temperature and centrifuged (2,500 x g for 3 min). The pellet was re-suspended twice in 5 mL water, centrifuged and the supernatant discarded. The enriched cellulose pellet was treated with 2.5 mL of 72% sulfuric acid (Updegraff, 1969) and incubated at room temperature for 1 h while mixing every 10 min by vortex. Samples were then transferred to a 15 mL Falcon tube and water added to 10 mL. Ten microliters of solution was transferred to a new borosilicate tube and diluted to 400 µL with water. One milliliter of ice-cold anthrone reagent (0.2 g anthrone in 100 mL concentrated sulfuric acid (95–98%)) was added, and the mixture was heated in a boiling water bath for 15 min. Following the anthrone reaction, the amount of sugars in the cellulosic polysaccharide fraction of nitric acid-treated unfractionated cellulose, was determine by measuring absorbance at 620 nm with a DU 800 series spectrophotometer (Beckmann Coulter) using glucose from Avicel as standard.

4.3.9 Cell wall sugar release and lignin content and composition

Tillers were collected at the R1 developmental stage from greenhouse-grown plants and air dried for 3 weeks at room temperature before grinding to 1 mm (20 mesh) particle size. Sugar release efficiency was determined via NREL high-throughput sugar release assays on extractive- and starch-free samples using glycosyl hydrolases according to NREL protocol (Decker et al., 2012; Selig et al., 2010). Glucose and xylose release was determined by colorimetric assays with total sugar release being the sum of glucose and xylose released. Lignin analysis was performed on the same samples described above. The lignin content and composition was determined by high-throughput pyrolysis molecular beam mass spectrometry (py-MBMS) on starch-free samples (Sykes et al., 2009) at the National Renewable Energy Laboratory (NREL).

4.3.10 Statistical analysis

Statistical analysis was carried out with biological and technical replicates using SAS® (Version 9.3 SAS Institute Inc., Cary, NC) programming of mixed model ANOVA and least significant difference. This statistic analyses was performed on *PvUAM1* and *PvUAM1* homolog transcript abundance by qRT-PCR, growth analysis, cell wall-associated sugar content, NMR sugar side chain analysis, cellulose quantification, enzymatic sugar release, lignin content and composition, and lignin biosynthesis gene quantification by qRT-PCR. The standard error of the mean was calculated and displayed as error bars. *P*-values of ≤ 0.05 were considered to be statistically significant.

4.4 Results

4.4.1 Identification of PvUAM homologs

Orthologous of functional UDP-arabinomutase (UAM1) amino-acid sequences from monocot and eudicot plant species were used to identify the switchgrass PvUAM1 sequence (Fig 4-1). The UAM1 has additional function or reversibly glycosylated protein (RGP1), hence forward be named UAM1/RGP1 or UAM1. PvUAM1 has 93% and 86% amino-acid sequence similarity to the rice UAM1 and Arabidopsis UAM1/RGP1, respectively. Sequence relationships of UAM proteins from diverse plant species grouped into a central monocot cluster and a split eudicot grouping. PvUAM1 belongs to the monocot group. In addition to UAM1 two other UAM-homologs are known. PvUAM1 is 48% similar to PvUAM2 whereas PvUAM1 has 86% amino-acid sequence similarity to PvUAM3.

4.4.2 Molecular and phenotypic characterization of PvUAM-RNAi transgenic plants

To study the role of PvUAM1 in switchgrass in hemicellulose metabolic pathways RNAi-transgenic plants were generated. Seven independent transgenic lines regenerated from transformed callus were analyzed (Fig 4-2A). Southern blot analysis showed that each transgenic line carried at least one and up to seven T-DNA inserts (Fig 4-S2B). One transgenic line (270-3) did not survive and was removed from subsequent analysis. The *PvUAM1* transcript abundance was less than that of the control in each of six remaining transgenic lines in both stems and leaves. For example, *PvUAM1* transcript level in stems and leaves of the RNAi plant lines, decreased by 67-5% and by 77-98% relative to the non-transgenic control, respectively (Fig 4-2B). Gene expression analysis of *PvUAM* homologs (*PvUAM1, PvUAM2, and PvUAM3*) was performed on stem internode sections (Fig 4-2C). *PvUAM2* expression amongst all transgenic lines was not significantly different than the control. However and although unintended, the *PvUAM1* RNAi target sequence was similar enough to cause significant downregulation in both *PvUAM1* and *PvUAM3* homologs for lines 270-1 and 270-2. *PvUAM3* transcript was significantly reduced in lines 270-1 and 270-2 to 10% and 36% relative to the control, respectively. An opposite effect was observed for 270-4 in which *PvUAM3* was found to be upregulated 4.6 fold over the control. *PvUAM3* transcript abundance was unchanged in 270-5, 270-6, and 270-7 compared to the non-transgenic control. Interestingly, we found no apparent correlation between number of T-DNA insert and the reduced transcript abundance of UAM1 in these transgenes.

There were several instances of altered plant growth among the transgenic switchgrass lines (Table 1). Transgenic plant lines 270-1, 270-2, 270-5 and 270-7 had equivalent number of tillers compared to the control whereas lines 270-4 and 270-6 had significantly more tillers per plant. Plant lines 270-1, 270-2, and 270-6 were shorter, whereas 270-4 and 270-5 were equivalent to control. Line 270-7 was taller when compared with control line. Tiller stem width was significantly reduced up to 22% in lines 270-1, 270-2, and 270-6, but was increased in 270-7 by up to 6%, whereas the remainder of the lines had unchanged stem width from the control. Fresh weight was significantly increased from the control by up to 102% in lines 270-2, 270-4, 270-5, and 270-7, whereas lines 270-1 and 270-6 were equivalent to control. Dry biomass results were similar to that of fresh weight except line 270-6 was also higher than control. Panicle number was significantly increased in all transgenic lines. In addition to the above mentioned phenotypic differences between PvUAM-RNAi lines and control we interestingly found that line 270-1, 270-2, and 270-4 appeared to have an increased level of red pigment in the stem nodes when compared with the non-transgenic plants (Fig 3A). Cross sectioning of fresh stem nodes and internodes at the E3 (elongation) developmental stage on plant line 270-4 showed a dark pigmentation that was deposited in the vascular bundles and outer tissue of the nodes (Fig 3B).

4.4.3 The wall of PvUAM-RNAi transgenic plants has reduced arabinose and increased glucose

Following phenotypic analyses of PvUAM-RNAi transgenic we investigated that nature of wall polysaccharide content in these plant lines. When compared with wall polysaccharide of control, the leaves from transgenic PvUAM-RNAi lines showed up to 52% decreased arabinose content

(Fig 4-4A). Similarly the wall from transgenic stem showed up to 58% decreased arabinose relative to the control (Fig 4-4B).

In addition to a reduced level of arabinose content in the wall, most transgenic plant line had also reduced level of xylose (up to 25%) in leaves (Fig 4C). One transgenic line was not significantly different in xylose content from the control line. No specific trend for xylose content was observed in wall isolated from stems (Fig. 4D). Interestingly, the level of glucose in leaf cell wallwas significantly higher in all transgenic lines when compared with control (Fig 4E), and a similar trend was observed in stems (Fig 4F). The galactose content in stem cell wall was significantly lower in transgenic plant line 270-1, 270-2, and 270-5 while lines 270-4, 270-6, and 270-7 were equal to the control (Table S3). In leaves, on the other hand, the amount of galactose in the wall was unchanged compared to the control (Table S4). The rhamnose content in the walls from leaves was similar among most transgenic plants, with exception of plant line 270-2 that had a significant increase rhamnose (up to 82%) (Table S4) when compared with control. In stems, three lines (270-1, 270-4, and 270-5) had reduced levels of cell wall rhamnose (up to 48%) (Table S3). Cell wall mannose levels in leaves and stems were not different among transgenic and control plants (Table S3 and S4). The amount of cellulose in both stem internodes and leaves was determined as well. An increase in cellulose level was observed in stems of all PvUAM-RNAi transgenic lines (Fig 5A), and in leaves, transgenic lines 270-1 and 270-6 showed a significantly higher amount cellulose when compared to the control (Fig 5B). The cellulose level in in leaves for lines 270-2, 270-4, 270-5, and 270-7 were equal to the control (Fig 5B).

Because total sugar analysis is insufficient to identify gross changes in polysaccharide structure and organization, NMR analysis of extracted arabinoxylan was performed.

4.4.4 Arabinoxylan has altered side chains in PvUAM-RNAi mutants

Analysis of individual sugars in the wall does not provide adequate assessment to indicate what is/are the specific polysaccharides whose synthesis was altered in PvUAM-RNAi transgenic plants. Therefore, we decided to analyze soluble polysaccharides; extracted and digested arabinoxylooligomers, from the stem internodes and in the leaves by NMR. This method for example, should distinguish in principle an arabinose in the furanose form from a pyranose form and should provide linkage anomeric configurations (α- or β-form) as well as linkage positions (xylan 1-4 and any branching 1-2, 1-3, etc). In addition NMR is an excellent method to quantify resolved sugar signals in mixtures of polymers. To characterize structural differences in heteroxylans, cell wall-extracted material from stem internodes and leaves was fractionated to enrich for arabinoxylan, digested into short oligomers and analyzed by ^{1}H-NMR.

The 1M KOH wall-soluble polysaccharides mixture isolated form stems and analyzed by proton –NMR showed clearly that the anomeric (H1) signals for α-xyl and β-xyl were higher in plant lines 270-1, 270-6, and 270-7 compared with the non-transgenic control. Because the reducing end xyl of the digested xylooligimer fragments in solution can be both α and β, both α-xyl and β-xyl are detected, with their relative increase suggesting less decorated polysaccharide. On the other hand, plant lines 270-2, 270-4, and 270-5 had equivalent (H1) xylose signals to the control (Table 4-2). NMR anomeric (H1) signals showed two clearly resolved arabinose signals. One signal depicts an α-arabinose that is linked 1-2 to a xylose. Another anomeric H1 signal

identifies terminal arabinose that is α-linked 1-2 to a xylose or to another arabinose. The relative amount of anomeric signal for a primary branch arabinose decorating the xylan backbone (2-α-ara) in stems was increased in lines 270-1 and 270-2 over control and was found to be equivalent for 270-4, 270-5, 270-6, and 270-7. In stems the relative anomeric signal for terminal arabinose residues (T- α-ara) was increased in lines 270-5, 270-6, and 270-7 over control while lines 270-1, 270-2, and 270-4 were equivalent (Table 4-2). In stems there was a relative increase in α-4-GlcA anomeric signal in lines 270-1 and 270-2 while the remaining transgenic lines were equivalent to the control. The 4-MeGlcA signal represents side-chain decoration of xylan. Because T-α-ara signal is representative of H1 of arabinofuranose for both monomeric and extended branches, these data suggest stem arabinoxylan in lines 270-1 and 270-2 has relatively more extended branches (2-α-ara) than monomer. Additionally, line 270-1 and 270-2 have less overall arabinose branching compared to non-transgenic primarily due to the monomer terminal arabinose signal dilution still appearing equivalent to non-transgenic (Fig 4-S3). Additionally, in these lines there appears to be increased GlcA decoration of the xylan backbone. In leaves the α-xyl and β-xyl were unchanged from control (Table 4-3). Leaf signal for 2-α-ara was increased in 270-1 and 270-2 compared to control and equivalent for the other lines. Leaf signal for T- α-ara was reduced in lines 270-4 and 270-5 compared to control. From leaves the α-4-GlcA signal was unchanged amongst all transgenic lines (Table 4-3). These results largely indicate that leaf arabinoxylan structure in lines 270-1 and 270-2 has been altered similar to stems; relatively less overall branching of the xylan backbone with more extended branches than monomers.

4.4.5 Saccharification of PvUAM-RNAi lines unchanged for total sugars

PvUAM-RNAi plant cell wall sugars were analyzed for polysaccharide enzymatic release from dried R1 tillers. Enzymatic sugar release is one indicator for the level of recalcitrance of the plant cell wall against enzymatic degradation. Enzymatic glucose release was increased up to 13% for 270-4, 270-6 whereas lines 270-1, 270-2, 270-5, and 270-7 were equal to control (Fig 4-6A). Enzymatic release of xylose was significantly increased only in line 270-5 by 17% while the other transgenic lines were equivalent to control (Fig 4-6B). When data for glucose and xylose released were added, there was no apparent change amongst transgenic plants and the control for total combined sugar release upon saccharification (Fig 4-6C).

4.4.6 Lignin biosynthesis: gene expression, lignin content and composition in tillers

Lignin amount was up to 13% higher in transgenic plants compared with the control (Fig 4-7A). Interestingly however, the relative ratio of the monolignol components syringyl and guaiacyl (also known as S/G ratio) was unchanged. On the other hand, the S/G ratio in transgenic plant line 270-2 and 270-7, was increased by 9% and 14%, respectively (Fig 4-7B) compared to control.

The finding that transgenic PvUAM-RNAi lead to increase in lignin prompt us to examine selected genes involved in lignin biosynthesis. The relative amount of gene transcript was determined by qRT-PCR. There was increased expression of *PAL*, *F5H*, *4CL*, *C4H*, and *CAD* genes in PvUAM-RNAi transgenes when compared to control (Fig 4-8). Expression levels of *COMT*, *C3H*, *CCR*, and *HCT* genes were unchanged compared to control.

4.5 Discussion

Arabinoxylans, which comprise a relatively large portion of cell walls in grass species, likely play an important role in recalcitrance in feedstocks such as switchgrass. Arabinoxylans strengthen cell walls through cross linkages with other cell wall polysaccharides and lignin (Faik, 2010; Rennie and Scheller, 2014; Scheller and Ulvskov, 2010; Tan et al., 2013). The xylan precursor UDP-Ara*f* has been identified as a common side chain unit of the xylan backbone and plays an integral part in cross linking to other cell wall components (Anders et al., 2012). Currently the UAM class of plant proteins is the sole candidate known to convert UDP-Ara*p* to UDP-Ara*f* (Konishi et al., 2011). It has been hypothesized that decreasing the pool of available UDP-Ara*f* would, in turn, change how arabinoxylans interact with cellulose microfibrils: the reduction of numbers of cross linkages would increase the solubility of arabinoxylans and lignin (Sternemalm et al., 2008). Hence, we propose that a reduction of *PvUAM1* would reduce arabinose side chains used for cross linkages among cell wall components, ergo, reducing recalcitrance. Our data partially supports the proposition, as the putative cross linkages of arabinoxylan to lignin appeared to be decreased, however we suggest that recalcitrance was unchanged by a compensation-like interaction of lignin and hemicellulose. There were reduced arabinose levels, but inexplicably, there was a concurrent increase in lignin in some of our transgenic plants.

4.5.1 PvUAM downregulation affects plant growth

Two of the shorter plants (transgenic lines 270-1 and 270-2) had decreased expression of both *PvUAM1* and *PvUAM3*. A similar double knockdown was seen in some transgenic rice due to

the close homology of the *OsUAM1* and *OsUAM3* resulting in down-regulation of both homologs (Konishi et al., 2011). Transgenic switchgrass with the double knockdown showed the significant differences in cell wall-associated arabinoxylan side chains in both stems and leaves. Downregulation of only *PvUAM1* did not result in a significant change to the side chains. The transgenic switchgrass exclusively downregulated for *PvUAM1* were taller; rice *OsUAM1*-RNAi plants were shorter (Konishi et al., 2011).

4.5.2 PvUAM downregulation alters cell wall-associated sugars with no change to sugar release

The phenotype we observed of decreased cell wall-associated arabinose is congruent with prior research in rice, *Arabidopsis* and *Brachypodium* (Konishi et al., 2011; Rancour et al., 2015; Rautengarten et al., 2011). In leaves, the *PvUAM1* transcript and cell wall-associated arabinose was significantly decreased in all lines

As arabinose is a portion of arabinoxylan, NMR was employed to deeper characterize arabinoxylan side chains. While a large amount of structural information can be deduced from the NMR spectra, there are regions of overlap in the anomeric signals of the xylo-oligos (Balazs et al., 2013). Specifically, the T-Ara signals for single 3/2-α-L-Ara*f* side chains and that for the di-substituted β-D-Xyl*p*-(1,2)-α-L-Ara*f* are unresolved as are the 2-Ara signals for 2-α-L-Ara*f*-(1,3)-α-D-Xyl*p* and α-L-Ara*f*-(1,2)-α-L-Ara*f* side chains. Overall, the data (Table 2 and 3) suggest that the switchgrass *PvUAM*-RNAi lines with knockdown of both *PvUAM1* and *PvUAM3* have an altered arabinoxylan structure, and that the reduction of available UDP-Ara*f*

causes either 1) reduced Ara*f* branching, or 2) reduced decoration of substituted side chains with a terminal Ara*f* residue (Rancour et al., 2015; Rancour et al., 2012). Evidence from 270-1 and 270-2 stems demonstrates there is a concomitant increase in glucuronate and 4-Me-glucuronate signals, which may be due to increased substitution with glucuronic acid (GlcA) to make glucuronoarabinoxylan. Arabinoxylan sidechains on the other lines with only *PvUAM1* knockdown were not significantly different from the non-transgenic control. With NMR analysis we can detect the changes in terminal and secondary sidechain substitution, however we cannot be certain of the identity of the residues.

Cell wall-associated glucose levels were increased in several transgenic lines (Fig 4-4E and 4-4F), which was suspected to be incorporated into cellulose. The majority of the transgenic lines also showed an increase in cell wall-associated cellulose in both stems and leaves (Fig 4-4A and 4-4B). In transgenic rice and *Arabidopsis* in which the relevant *UAM1* homolog had decreased expression, there was no significant change to cell wall-associated glucose (Konishi et al., 2011; Rautengarten et al., 2011). Cell wall-associated xylose was reduced in the leaves of all transgenic lines, while in stem, it was varied among the transgenic lines (Fig 4-4C and 4-4D). In transgenic *AtUAM1*-RNAi *Arabidopsis* cell wall-associated xylose was increased, but in *OsUAM1*-RNAi rice in was unchanged (Konishi et al., 2011; Rautengarten et al., 2011). Transgenic *Brachypodium* with *BdUAM1* knocked down by RNAi showed an increase of cell wall-associated xylose in some lines and a decrease in others (Rancour et al., 2015).

Even though an increase in cell wall-associated glucose and cellulose content was detected, no

change in enzymatic saccharification was found (Fig 4-6A and 4-6C). Line 270-5 had increased enzymatic xylose release, which might be attributed to the increase in available stem cell wall-associated xylose (Fig 4-4D and 4-6B). Transgenic *Brachypodium* with *BdUAM1* knocked down by RNAi had slightly increased enzymatic glucose release from stems, but significantly lower release from leaves (Rancour et al., 2015). We did not analyze saccharification by organ—only whole tillers.

4.5.3 PvUAM down-regulation increases lignin content and composition

Attenuated UAM switchgrass plants had higher lignin in tillers. When *BdUAM1* was knocked down in *Brachypodium* lignin was increased in the leaves, but was found to be unchanged in sheath/stem portions of tillers (Rancour et al., 2015). Lignin was not analyzed in the rice and *Arabidopsis* studies involving *UAM1* knockdown (Konishi et al., 2011; Rautengarten et al., 2011). Even though there was an increase in cell wall lignin in our study, saccharification was mostly unchanged, which contrasts to the similar study in *Brachypodium*, in which saccharification increased (Fig 4-7A) (Rancour et al., 2015). For line 270-5, *PvUAM1* expression was reduced to 56% of native PvUAM1 transcript, lignin content and composition was unchanged, while enzymatic xylose release increased. Further knockdown of *PvUAM1* expression (as seen in other lines) caused an increase in lignin production without reducing enzymatic sugar release. One might speculate that altered ferulated xylan formation affects lignification and in order to maintain proper plant growth and development compensation is made by increasing lignin content which has been reported in *Brachypodium* (Rancour et al., 2015) and now switchgrass. Ferulation is suggested to enable cross-linking of these

polysaccharides to each other as well as to lignin (Molinari et al., 2013). These cross-links are believed to strengthen the cell wall and in part, contribute to the enhanced rigidity of the walls.

Examination of the converse phenotype, where lignin is down-regulated shows evidence of a potential cell wall compensation mechanism as hemicellulose is increased to replace missing lignin. In the maize brown-midrib lignin mutants (*bm3)*, cell wall-associated xylose content was discovered to be equivalent or higher in certain lines while arabinose, rhamnose, and xylose-substitutions decreased (Guillaumie et al., 2008). In transgenic switchgrass using RNAi to downregulate COMT lignin biosynthetic gene, an increase in hemicellulose, xylan, and arabinan were observed (Baxter et al., 2014). In-depth characterization of cell wall polysaccharides in cell wall mutants might reveal the interactions amongst cell wall biosynthesis pathways.

We propose a model of the interaction of hemicellulose and lignin in light of our study (Fig 4-9). From glycolysis, glucose-6-phosphate converted sucrose or shunted towards phenylpropanoid production starting with shikimic acid pathway. Shikimic acid is the precursor for phenylalanine, which is at the top of the lignin biosynthesis pathway (Whetten and Sederoff, 1995). Sucrose is converted to UDP-glucose which is either up taken by cellulose synthase (Ces) complex to form cellulose or is converted to UDP-GlcA by UDP-glucose dehydrogenase (UGD). UDP-GlcA is converted by UDP-D-xylose synthase (UXS) to UDP-D-xylose. Excess UDP-D-xylose in the Golgi stack can inhibit UGD and UXS preventing further buildup of UDP-D-xylose (Harper and Bar-Peled, 2002). UDP-D-xylose is either converted to 1,4,-B-D-xylan by a xylan synthase (XS) or into the arabinose precursor UDP-L-Ara*p* (Rennie and Scheller, 2014). UDP-L-

Ara*p* is converted to UPD-L-Ara*f* by UAM and then recruited into arabinose or arabinoxylan (Konishi et al., 2011; Kotani et al., 2013; Rancour et al., 2015; Rautengarten et al., 2011; Rennie and Scheller, 2014). We propose that the reduction in available UPD-L-Ara*f* caused by *PvUAM*-RNAi results in an increase of UDP-D-xylose with a corresponding reduction in D-xylan and therefore arabinoxylan. The possible reduction in available arabinoxylan, which is normally ferulated by an unknown feruloyl transferase, causes an increase of feruloyl-CoA. Excess ferulic and caffeic acid accumulation is shunted to lignin biosynthesis. This model is supported by the decrease in xylan, arabinose, and arabinose-furanose side chains (Fig 4-6 and Tables 4-2 and 4-3) found in the *PvUAM*-RNAi lines. The increase in lignin content and S/G ratio along with upregulation of lignin biosynthetic genes (Fig 4-7 and 4-8) supports the probability of increased synaptic acid levels being generated by increased F5H transcript. Further testing of hypotheses inferred by this model could be accomplished by making combinatorial knockdowns of genes that code enzymes in arabinoxylan and lignin biosynthetic pathways. Pleiotropic perturbations in gene expression and metabolic flux would be informative. Identification and characterization of the suspected feruloyl transferase would aid in discerning the complete molecular mechanism for how the cross linking between arabinoxylan and lignin occurs.

Interference with arabinofuranose metabolism has impacted cross-linking and lignin, with no evident influence on sugar release in switchgrass. The increase in lignin of *PvUAM*-RNAi plants might seem unfavorable for lignocellulosic ethanol production, however, the sugar release efficiency was not affected by the increase in lignin content. Yet, the increase in cell wall-associated glucose content is advantageous for cellulosic ethanol production. While *PvUAM*-

RNAi might not be directly suitable for bioenergy feedstocks, they might be employed as a crossing partner with other switchgrass low in lignin. Particular examples are *COMT* and *MYB4* transgenic lines modified for decreased lignin and increased sugar release efficiency, which might complement the increase in glucose, lignin, and biomass of *PvUAM*-RNAi transgenics in transgene stacks (Baxter et al., 2014, Baxter et al., 2015; Fu et al., 2011; Shen et al., 2012). Additional switchgrass lines that might be useful to cross with UAM1 knockdown transgenics are *MYB4* and *miR156* overexpressers, both of which were dwarfed but yielded relatively high biofuel production (Baxter et al., 2015; Fu et al., 2012; Shen et al., 2012). Additionally, any feedstock that produces inordinately high amounts of lignin might be useful for co-products, such as carbon fiber and bio-plastics (Lindsey et al., 2013; Ragauskas et al., 2014).

4.6 Conclusion

We have identified UAM in switchgrass and the downregulated *PvUAM*1 switchgrass plants have altered cell wall-associated sugar content and side chains. Downregulation of *PvUAM1* produced a decrease in cell wall-associated arabinose content with concurrent increase in cell wall-associated glucose and lignin content in transgenic switchgrass. We propose a model in which the decrease of available arabinoxylan caused an increase in lignin content due to excess metabolites not being used for arabinoxylan-lignin cross linking. Enzymatic saccharification was not negatively affected by the increase in lignin content possibly because there was more available cell wall-associated glucose in transgenic leaves and stems. Some transgenic *PvUAM1* were larger, which would be useful for commercial biomass and carbon sequestration platforms as well as a lignin feedstock.

4.7 Acknowledgements

We acknowledge the technical assistance of Ellen Haynes for providing switchgrass callus tissue for transformation and A. Grace Collins for help in maintaining plants and assisting with DNA extractions for Southern blot analysis. We thank Angela Ziebell, Erica Gjersing, Crissa Doeppke, and Melvin Tucker of NREL for their assistance with the cell wall characterization and Susan Holladay for her assistance with data entry into LIMS. This work was supported by funding from the BioEnergy Science Center. The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

4.8 References

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

Anisimova, M. and Gascuel, O. (2006) Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst. Biol.* **55**, 539-552.

Balazs, Y.S., Lisitsin, E., Carmiel, O., Shoham, G., Shoham, Y. and Schmidt, A. (2013) Identifying critical unrecognized sugar-protein interactions in GH10 xylanases from *Geobacillus stearothermophilu*s using STD NMR. *FEBS J.* **280**, 4652-4665.

Bar-Peled, M. and O'Neill, M.A. (2011) Plant nucleotide sugar formation, interconversion, and salvage by sugar recycling. *Annu. Rev. Plant Biol.* **62**, 127-155.

Baxter, H.L., Mazarei, M., Labbe, N., Kline, L.M., Cheng, Q., Windham, M.T., Mann, D.G., Fu, C., Ziebell, A., Sykes, R.W., Rodriguez, M., Jr., Davis, M.F., Mielenz, J.R., Dixon, R.A., Wang, Z.Y. and Stewart, C.N., Jr. (2014) Two-year field analysis of reduced recalcitrance transgenic switchgrass. *Plant Biotechnol. J.* **12**, 914-924.

Baxter, H.L., Poovaiah, C.R., Yee, K.L., Mazarei, M., Rodriguez, M., Thompson, O.A., Shen, H., Turner, G.B., Decker, S.R., Sykes, R.W., Chen, F., Davis, M.F., Mielenz, J.R., Davison, B.H., Dixon, R.A. and Stewart, C.N. (2015) Field evaluation of transgenic switchgrass plants overexpressing PvMYB4 for reduced biomass recalcitrance. *Bioenerg Res.* **8**, 910-921.

Burris, J.N., Mann, D.G.J., Joyce, B.L. and Stewart, C.N. (2009) An improved tissue culture system for embryogenic callus production and plant regeneration in switchgrass (*Panicum virgatum* L.). *Bioenerg. Res.* **2**, 267-274.

Carpita, N.C. (1986) Incorporation of proline and aromatic amino acids into cell walls of maize coleoptiles. *Plant Physiol.* **80**, 660-666.

Casler, M.D., Tobias, C.M., Kaeppler, S.M., Buell, C.R., Wang, Z.Y., Cao, P.J., Schmutz, J., and Ronald, P. (2011). The Switchgrass genome: Tools and strategies. *Plant Genome* 4**,** 273-282. doi: 10.3835/plantgenome2011.10.0026.

Chen, X., Ma, Q., Rao, X.L., Tang, Y.H., Wang, Y., Li, G.Y., Zhang, C., Mao, X.Z., Dixon, R., and Xu, Y. (2016). Genome-scale identification of cell-wall-related genes in awitchgrass through comparative genomics and computational analyses of transcriptomic data. *Bioenergy Research* 9**,** 172-180. doi: 10.1007/s12155-015-9674-2.

Danilova, S.A. and Dolgikh, Y.I. (2004) The stimulatory effect of the antibiotic cefotaxime on plant regeneration in maize tissue culture. *Russ. J. Plant. Physl.* **51**, 559-562.

De O. Buanafina, M. (2009). Feruloylation in grasses: current and future perspectives. *Mol Plant* **2,** 861-872. doi: 10.1093/mp/ssp067.

Decker, S.R., Carlile, M., Selig, M.J., Doeppke, C., Davis, M., Sykes, R., Turner, G. and Ziebell, A. (2012) Reducing the effect of variable starch levels in biomass recalcitrance screening. *Methods Mol. Biol.* **908**, 181-195.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M. and Gascuel, O. (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465-469.

Dhugga, K.S., Ulvskov, P., Gallagher, S.R. and Ray, P.M. (1991) Plant polypeptides reversibly glycosylated by UDP-glucose. Possible components of Golgi beta-glucan synthase in pea cells. *J. Biol. Chem.* **266**, 21977-21984.

Dixon, R.A. (2013). Microbiology: Break down the walls. *Nature* 493**,** 36-37. doi: 10.1038/493036a.

Faik, A. (2010) Xylan biosynthesis: news from the grass. *Plant Physiol.* **153**, 396-402.

Freeling, M. and Walbot, V. (1994) *The Maize Handbook*. New York:Springer-Verlag.

Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M., Jr., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A. and Wang, Z.Y. (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3803-3808.

Fu, C., Sunkar, R., Zhou, C., Shen, H., Zhang, J.Y., Matts, J., Wolf, J., Mann, D.G., Stewart, C.N., Jr., Tang, Y. and Wang, Z.Y. (2012) Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotechnol. J.* **10**, 443-452.

Guillaumie, S., Goffner, D., Barbier, O., Martinant, J.P., Pichon, M. and Barriere, Y. (2008) Expression of cell wall related genes in basal and ear internodes of silking brown-midrib-3, caffeic acid O-methyltransferase (COMT) down-regulated, and normal maize plants. *BMC Plant Biol.* **8**, 71.

Hardin, C.F., Fu, C.X., Hisano, H., Xiao, X.R., Shen, H., Stewart, C.N., Parrott, W., Dixon, R.A. and Wang, Z.Y. (2013) Standardization of switchgrass sample collection for cell wall and biomass trait analysis. *Bioenerg Res.* **6**, 755-762.

Hartley, R.D. and Ford, C.W. (1989) Phenolic constituents of plant-cell walls and wall biodegradability. *ACS Symp. Ser.* **399**, 137-145.

Harper, A.D., and Bar-Peled, M. (2002). Biosynthesis of UDP-xylose. Cloning and characterization of a novel Arabidopsis gene family, UXS, encoding soluble and putative membrane-bound UDP-glucuronic acid decarboxylase isoforms. *Plant Physiol* 130**,** 2188-2198. doi: 10.1104/pp.009654.

Hatfield, R.D., Ralph, J. and Grabber, J.H. (1999) Cell wall cross-linking by ferulates and diferulates in grasses. *J. Sci. Food. Agr.* **79**, 403-407.

Iiyama, K., Lam, T.B.T. and Stone, B.A. (1994) Covalent cross-links in the cell-wall. *Plant Physiol.* **104**, 315-320.

Konishi, T., Aohara, T., Igasaki, T., Hayashi, N., Miyazaki, Y., Takahashi, A., Hirochika, H., Iwai, H., Satoh, S. and Ishii, T. (2011) Down-regulation of UDP-arabinopyranose mutase reduces the proportion of arabinofuranose present in rice cell walls. *Phytochemistry* **72**, 1962-1968.

Konishi, T., Ohnishi-Kameyama, M., Funane, K., Miyazaki, Y., Konishi, T. and Ishii, T. (2010) An arginyl residue in rice UDP-arabinopyranose mutase is required for catalytic activity and autoglycosylation. *Carbohydr. Res.* **345**, 787-791.

Konishi, T., Takeda, T., Miyazaki, Y., Ohnishi-Kameyama, M., Hayashi, T., O'Neill, M.A. and Ishii, T. (2007) A plant mutase that interconverts UDP-arabinofuranose and UDP-arabinopyranose. *Glycobiology* **17**, 345-354.

Kotani, A., Tsuji, M., Azama, Y., Ishii, T., Takeda, T., Yamashita, T., Shimojima, M. and Konishi, T. (2013) Purification and characterization of UDP-arabinopyranose mutase from *Chlamydomonas reinhardtii*. *Biosci. Biotechnol. Biochem.* **77**, 1874-1878.

Lee, H.V., Hamid, S.B. and Zain, S.K. (2014) Conversion of lignocellulosic biomass to nanocellulose: structure and chemical process. *ScientificWorldJournal* **2014**, 631013.

Li, R.Y. and Qu, R.D. (2011) High throughput *Agrobacterium*-mediated switchgrass transformation. *Biomass. Bioenerg.* **35**, 1046-1054.

Lindsey, K., Johnson, A., Kim, P., Jackson, S. and Labbe, N. (2013) Monitoring switchgrass composition to optimize harvesting periods for bioenergy and value-added products. *Biomass Bioenerg.* **56**, 29-37.

Mann, D.G., Abercrombie, L.L., Rudis, M.R., Millwood, R.J., Dunlap, J.R. and Stewart, C.N., Jr. (2012a) Very bright orange fluorescent plants: endoplasmic reticulum targeting of orange fluorescent proteins as visual reporters in transgenic plants. *BMC Biotechnol.* **12**, 17.

Mann, D.G.J., LaFayette, P.R., Abercrombie, L.L., King, Z.R., Mazarei, M., Halter, M.C., Poovaiah, C.R., Baxter, H., Shen, H., Dixon, R.A., Parrott, W.A. and Stewart, C.N. (2012b) Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum* L.) and other monocot species. *Plant Biotechnol. J.* **10**, 226-236.

Martinez, V., Ingwers, M., Smith, J., Glushka, J., Yang, T. and Bar-Peled, M. (2012) Biosynthesis of UDP-4-keto-6-deoxyglucose and UDP-rhamnose in pathogenic fungi *Magnaporthe grisea* and *Botryotinia fuckeliana*. *J. Biol. Chem.* **287**, 879-892.

Molinari, H.B., Pellny, T.K., Freeman, J., Shewry, P.R. and Mitchell, R.A. (2013) Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in *Brachypodium distachyon*. *Front. Plant Sci.* **4**, 50.

Ragauskas, A.J., Beckham, G.T., Biddy, M.J., Chandra, R., Chen, F., Davis, M.F., Davison, B.H., Dixon, R.A., Gilna, P., Keller, M., Langan, P., Naskar, A.K., Saddler, J.N., Tschaplinski, T.J., Tuskan, G.A. and Wyman, C.E. (2014) Lignin valorization: Improving lignin processing in the biorefinery. *Science* **344**, 1246843-1-1246843-10.

Rancour, D.M., Hatfield, R.D., Marita, J.M., Rohr, N.A. and Schmitz, R.J. (2015) Cell wall composition and digestibility alterations in *Brachypodium distachyon* achieved through reduced expression of the UDP-arabinopyranose mutase. *Front. Plant Sci.* **6**, 446.

Rancour, D.M., Marita, J.M. and Hatfield, R.D. (2012) Cell wall composition throughout development for the model grass Brachypodium distachyon. *Front. Plant Sci.* **3**, 266.

Rautengarten, C., Ebert, B., Herter, T., Petzold, C.J., Ishii, T., Mukhopadhyay, A., Usadel, B. and Scheller, H.V. (2011) The interconversion of UDP-arabinopyranose and UDP-arabinofuranose is indispensable for plant development in Arabidopsis. *Plant Cell* **23**, 1373-1390.

Rennie, E.A. and Scheller, H.V. (2014) Xylan biosynthesis. *Curr. Opin. Biotechnol.* **26**, 100-107.

Scalbert, A., Monties, B., Lallemand, J.Y., Guittet, E. and Rolando, C. (1985) Ether linkage between phenolic-acids and lignin fractions from wheat straw. *Phytochemistry* **24**, 1359-1362.

Scheller, H.V. and Ulvskov, P. (2010) Hemicelluloses. *Annu. Rev. Plant Biol.* **61**, 263-289.

Schmittgen, T.D. and Livak, K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* **3**, 1101-1108.

Selig, M.J., M.P., T., Sykes, R.W., Reichel, K.L., Brunecky, R., Himmel, M.E., Davis, M.F. and Decker, S.R. (2010) Lignocellulose recalcitrance screening by integrated high-throughput hydrothermal pretreatment and enzymatic saccharification. *Ind. Biotechnol.* **6**, 104-111.

Shen, H., Fu, C.X., Xiao, X.R., Ray, T., Tang, Y.H., Wang, Z.Y. and Chen, F. (2009) Developmental control of lignification in stems of lowland switchgrass variety Alamo and the effects on saccharification efficiency. *Bioenerg Res.* **2**, 233-245.

Shen, H., He, X.Z., Poovaiah, C.R., Wuddineh, W.A., Ma, J.Y., Mann, D.G.J., Wang, H.Z., Jackson, L., Tang, Y.H., Stewart, C.N., Chen, F. and Dixon, R.A. (2012) Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytol.* **193**, 121-136.

Shen, H., Mazarei, M., Hisano, H., Escamilla-Trevino, L., Fu, C., Pu, Y., Rudis, M.R., Tang, Y., Xiao, X., Jackson, L., Li, G., Hernandez, T., Chen, F., Ragauskas, A.J., Stewart, C.N., Jr., Wang, Z.Y. and Dixon, R.A. (2013) A genomics approach to deciphering lignin biosynthesis in switchgrass. *The Plant Cell* **25**, 4342-4361.

Sternemalm, E., Hoije, A. and Gatenholm, P. (2008) Effect of arabinose substitution on the material properties of arabinoxylan films. *Carbohydr. Res.* **343**, 753-757.

Sykes, R., Yung, M., Novaes, E., Kirst, M., Peter, G. and Davis, M. (2009) High-throughput screening of plant cell-wall composition using pyrolysis molecular beam mass spectroscopy. *Methods Mol. Biol.* **581**, 169-183.

Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., Hao, Z., Zhu, X., Avci, U., Miller, J.S., Baldwin, D., Pham, C., Orlando, R., Darvill, A., Hahn, M.G., Kieliszewski, M.J. and Mohnen, D. (2013) An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* **25**, 270-287.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Updegraff, D.M. (1969) Semimicro determination of cellulose in biological materials. *Anal. Biochem.* **32**, 420-424.

Vega-Sanchez, M.E., and Ronald, P.C. (2010). Genetic and biotechnological approaches for biofuel crop improvement. *Curr Opin Biotechnol* 21**,** 218-224. doi: 10.1016/j.copbio.2010.02.002.

Whetten, R., and Sederoff, R. (1995). Lignin Biosynthesis. *Plant Cell* 7**,** 1001-1013. doi: 10.1105/tpc.7.7.1001.

Xu, B., Escamilla-Trevino, L.L., Sathitsuksanoh, N., Shen, Z., Shen, H., Zhang, Y.H., Dixon, R.A. and Zhao, B. (2011) Silencing of 4-coumarate:coenzyme A ligase in switchgrass leads to reduced lignin content and improved fermentable sugar yields for biofuel production. *New Phytol.* **192**, 611-625.

York, W.S., and O' Neill, M.A., (2008) Biochemical control of xylan biosynthesis – which end is up? *Curr. Opin. Plant. Biol.* **11(3)**, 258-265.

Zhou, X., Xu, J., Wang, Z., Cheng, J.J., Li, R. and Qu, R. (2012) Dilute sulfuric acid pretreatment of transgenic switchgrass for sugar production. *Bioresour. Technol.* **104**, 823-827.

4.9 Chapter 4 appendix

4.9.1 Tables and figures

Table 4-1 Growth of down-regulated PvUAM1 transgenic and non-transgenic (NT-ST1) switchgrass lines. Tiller number refers to the mean tally of all tillers per pot present at time of collection. Tiller height refers to the mean value of the five tallest tillers per replicate pot. Stem width of the five tallest tillers per replicate pot was measured at the potting level and averaged. Panicle number refers to the mean value of panicles present at time of collection per pot. Fresh weight refers to the mean value of fresh total biomass collected per pot. Dry weight refers to the mean value of total biomass collected and then dried for five days at 42 °C per pot. Error bars represent standard error of the mean of three replicates. Values that share the same letter are not significantly different as calculated by LSD ($p \leq 0.05$).

Line	Tiller number	LSD	Tiller height (mm)	LSD	Stem width (mm)	LSD	Panicle number	LSD	Fresh weight (g)	LSD	Dry weight (g)	LSD
270-1	32.0 ± 4.9	AB	873.67 ± 17.9	D	3.61 ± 0.12	C	3.7 ± 0.8	ABC	87.70 ± 16.24	BC	22.86 ± 4.37	BC
270-2	29.0 ± 4.9	AB	923.9 ± 17.9	D	3.68 ± 0.12	C	5.3 ± 0.8	AB	165.24 ± 16.24	A	40.21 ± 4.37	A

Table 4-1 continued

Line	Tiller number	LSD	Tiller height (mm)	LSD	Stem width (mm)	LSD	Panicle number	LSD	Fresh weight (g)	LSD	Dry weight (g)	LSD
270-4	42.3 ± 4.9	A	1089.1 ± 17.9	BC	4.12 ± 0.12	AB	5.0 ± 0.8	AB	170.35 ± 16.24	A	39.72 ± 4.37	A
270-5	33.7 ± 4.9	AB	1087.2 ± 17.9	BC	4.37 ± 0.12	A	3.3 ± 0.8	BC	135.81 ± 16.24	AB	35.64 ± 4.37	AB
270-6	40.0 ± 4.9	A	1039.9 ± 17.9	C	3.88 ± 0.12	BC	5.7 ± 0.8	A	128.15 ± 16.24	ABC	34.15 ± 4.37	AB
270-7	36.0 ± 4.9	AB	1185.9 ± 17.9	A	4.36 ± 0.12	A	4.3 ± 0.8	ABC	174.21 ± 16.24	A	41.23 ± 4.4	A
NT-ST1	22.0 ± 4.9	B	1118.3 ± 17.9	B	3.85 ± 0.12	BC	2.7 ± 0.8	C	86.34 ± 16.24	C	18.62 ± 4.37	C

Table 4-2 Glycosyl side chain analysis from stems of down-regulated PvUAM1 transgenic and non-transgenic (NT-ST1) switchgrass lines. α-xyl refers to the portion of α-xylose present. β-xylose refers to the portion of β-xylose present. 2-α-Ara refers to the internal chains of arabinose present. T-α-Araf refers to the terminal residues of arabinose present. α-4-GlcA refers to the glucuronic acid present. Values represent relative % of ^{1}H-signal of the total xylan signal. Error bars represent standard error of the mean of three replicates. Values that share the same letter are not significantly different as calculated by LSD ($p \leq 0.05$).

Line	α-Xyl (% signal)	LSD	β-Xyl (% signal)	LSD	2- α –Ara (% signal)	LSD	T-α-Araf (% signal)	LSD	α-4-GlcA (% signal)	LSD
270-1	8.43 ± 0.24	AB	91.57 ± 5.13	A	4.71 ± 0.04	A	11.2 ± 1.13	ABCD	5.19 ± 0.45	A
270-2	6.82 ± 1.44	D	93.18 ± 19.65	ABC	4.46 ± 0.89	BC	11.6 ± 3.42	D	5.78 ± 2.11	ABC
270-4	9.05 ± 1.48	BCD	90.95 ± 5.35	BC	3.76 ± 0.23	CD	13.9 ± 2.59	BCD	3.97 ± 0.73	BCD
270-5	9.28 ± 1.08	ABC	90.72 ± 7.71	ABC	3.12 ± 0.23	CD	14.5 ± 1.83	AB	2.74 ± 0.62	CD
270-6	9.46 ± 0.16	A	90.54 ± 4.44	A	2.78 ± 0.11	CD	13.6 ± 0.98	A	2.34 ± 0.36	CD
270-7	9.57 ± 0.37	A	90.43 ± 4.30	AB	3.06 ± 0.15	CD	12.5 ± 0.60	AB	2.52 ± 0.32	CD
NT-ST1	9.07 ± 0.61	CD	90.93 ± 4.16	C	3.03 ± 0.20	D	13.4 ± 1.33	CD	2.92 ± 0.25	D

Table 4-3 Glycosyl side chain analysis from leaves of down-regulated PvUAM1 transgenic and non-transgenic (NT-ST1) switchgrass lines. α-xyl refers to the portion of α-xylose present. β-xylose refers to the portion of β-xylose present. 2-α-Ara refers to the internal chains of arabinose present. T-α-Araf refers to the terminal residues of arabinose present. α-4-GlcA refers to the glucuronic acid present. Values represent relative % of ^{1}H-signal of the total xylan signal. Error bars represent standard error of the mean of three replicates. Values that share the same letter are not significantly different as calculated by LSD ($p \leq 0.05$).

Line	α-Xyl (% signal)	LSD	β-Xyl (% signal)	LSD	2- α –Ara (% signal)	LSD	T-α-Araf (% signal)	LSD	α-4-GlcA (% signal)	LSD
270-1	9.95 ± 0.44	A	90.05 ± 1.73	A	4.43 ± 0.32	A	11.0 ± 0.41	AB	3.40 ± 0.74	A
270-2	9.67 ± 0.28	ABC	90.33 ± 3.84	AB	4.74 ± 0.14	A	10.8 ± .74	BC	2.38 ± 0.33	AB
270-4	9.50 ± 0.51	C	90.50 ± 9.71	D	3.33 ± 0.29	D	11.5 ± 0.97	D	2.35 ± 0.12	B
270-5	9.57± 0.42	ABC	90.43 ± 4.51	BCD	3.73 ± 0.43	BCD	13.7 ± 1.33	CD	1.71 ± 0.49	B
270-6	10.40 ± 0.61	AB	89.60 ± 2.56	ABC	3.60 ± 0.19	BC	12.6 ± 0.70	A	2.13 ± 0.39	B
270-7	9.83 ± 0.24	BC	90.18 ± 0.49	CD	3.54 ± 0.17	CD	12.9 ± 0.04	AB	1.83 ± 0.44	B
NT-ST1	9.57 ± 0.24	ABC	90.43 ± 5.25	ABCD	3.42 ± 0.23	CD	13.0 ± 1.08	AB	2.40 ± 0.35	AB

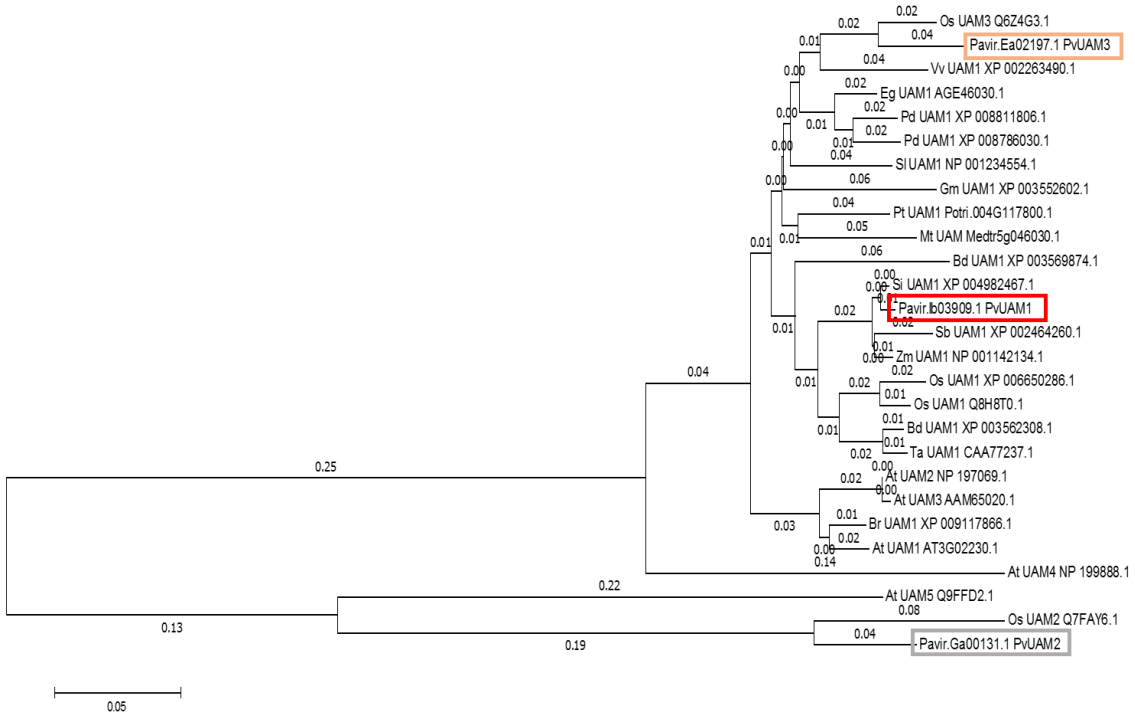

Figure 4-1 Neighbor-joining cluster analysis of UAM amino acids. The numerals on branches represent the relative differences among relevant UAM protein amino acid substitutions per site. Boxed are PvUAM1 (Pavirv000Ib03909; red), PvUAM2 (Pavir.GA00131.1; grey), PvUAM3 (Pavir.EA02197.1; peach). Tree generated using the MEGA 7.0 program (Tamura et al. 2013) of UAM amino acid sequence alignments using Gblocks at phylogeny.fr (http://phylogeny.lirmm.fr).

Boxed are PvUAM1 (Pavirv000Ib03909; red), PvUAM2 (Pavir.GA00131.1; grey), PvUAM3 (Pavir.EA02197.1; peach).

Figure 4-2 A) Representative down-regulated PvUAM1 transgenic and non-transgenic (NT-ST1) switchgrass lines. B) Relative expression of *PvUAM1* in leaf and stem tissues of transgenic and non-transgenic lines. C) Relative expression of *PvUAM1*, *PvUAM2*, and *PvUAM3* in stem tissue of transgenic and non-transgenic lines. Relative expression analysis were determined by qRT-PCR and normalized to switchgrass ubiquitin 1 (*PvUbi1*). Bars represent mean values of three replicates ± standard error. Asterisks indicate significant differences from non-transgenic control plants at $p \leq 0.05$ as tested by LSD method.

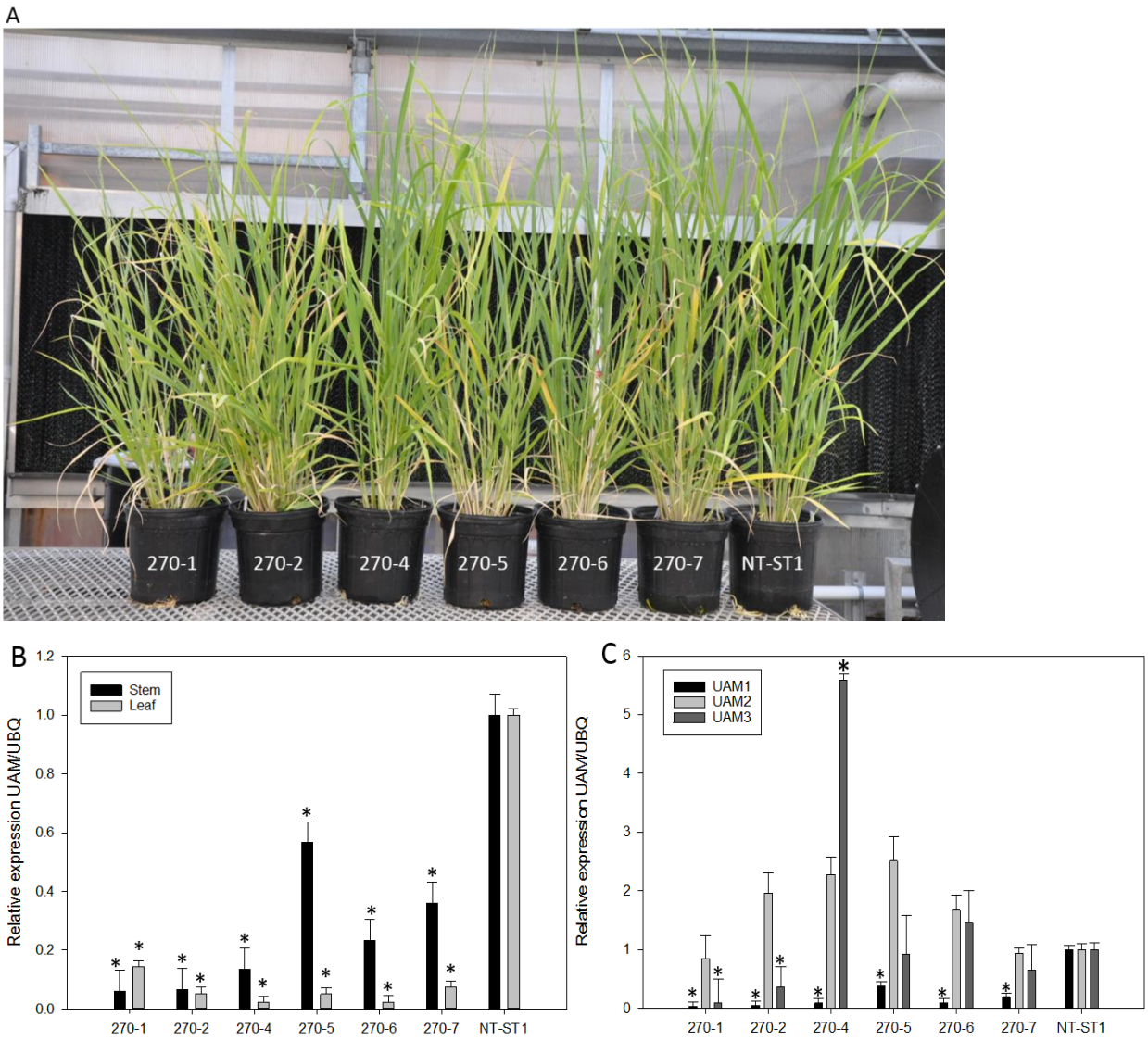

Figure 4-2 continued

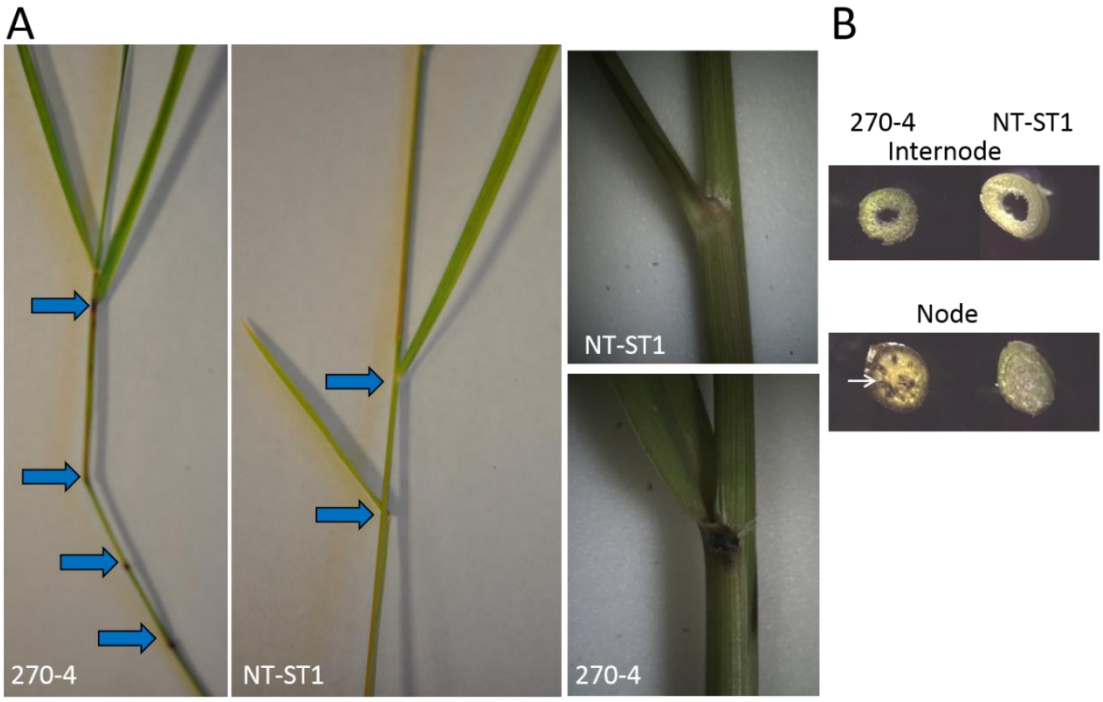

Figure 4-3 Stem node phenotype from fresh E3 (elongation growth stage) tillers in down-regulated PvUAM1 transgenic switchgrass. A) Comparison of transgenic PvUAM1 (270-4) and non-transgenic (NT-ST1) nodes and internodes. Arrows indicate nodes. B) Cross-sections of vascular bundles at nodes and internodes of transgenic and non-transgenic plants. Arrow indicates darkened vascular bundle.

Figure 4-4 Arabinose (A), xylose (B), and glucose (C) content in stem and leaves of transgenic and non-transgenic (NT-ST1) lines as determined by gas chromatography. Samples were normalized to the non-transgenic control with mol% representing the % of total cell wall-associated sugars measured. Bars represent mean values of three replicates ± standard error. Bars represented by same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

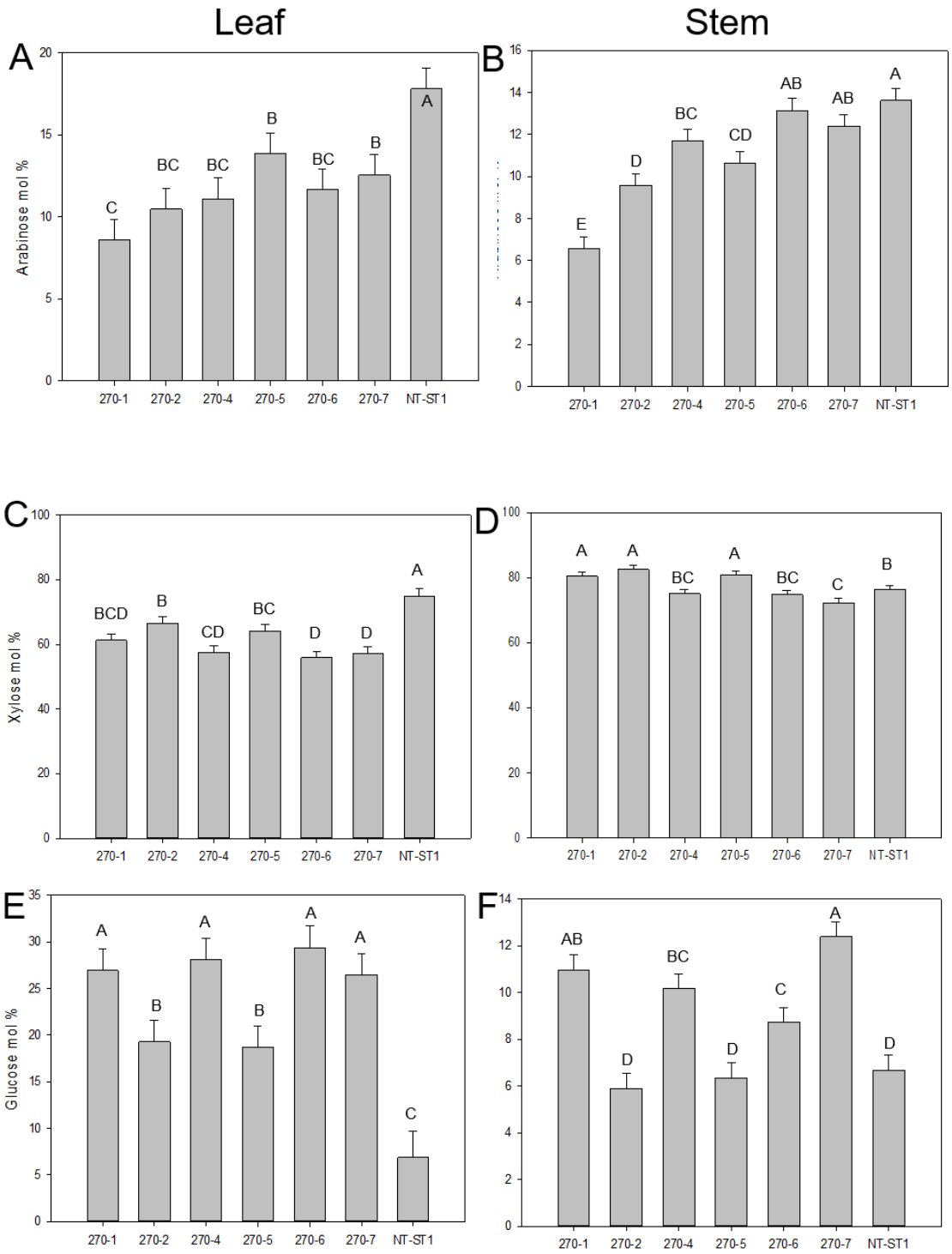

Figure 4-4 continued

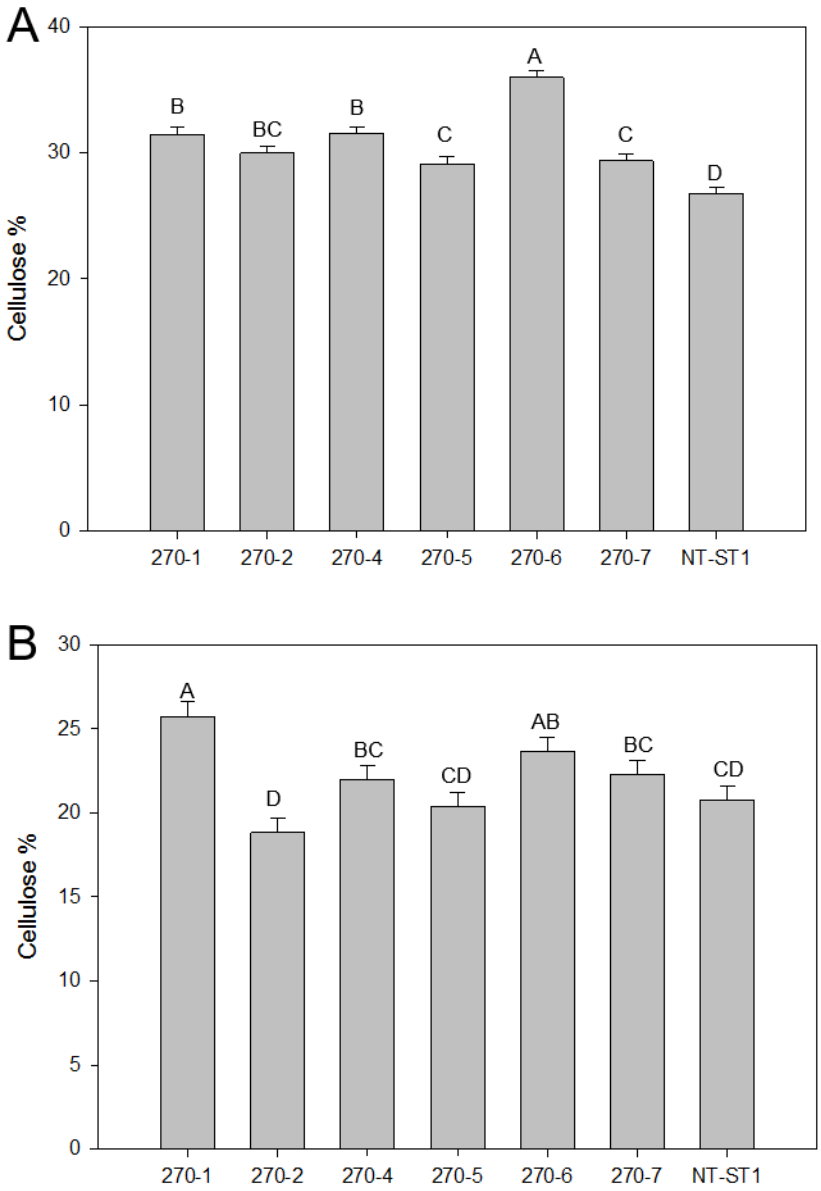

Figure 4-5 Cellulose content in stems (A) and in leaves (B) of transgenic and non-transgenic (NT-ST1) lines as determined by Updegraff reagent. Bars represent mean values of three replicates ± standard error. Bars represented by same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

Figure 4-6 Glucose (A), xylose (B), and total sugar (C) release from transgenic and non-transgenic (NT-ST1) whole tiller cell wall residues as determined by enzymatic hydrolysis. Bars represent mean values of three replicates ± standard error. Bars represented by same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

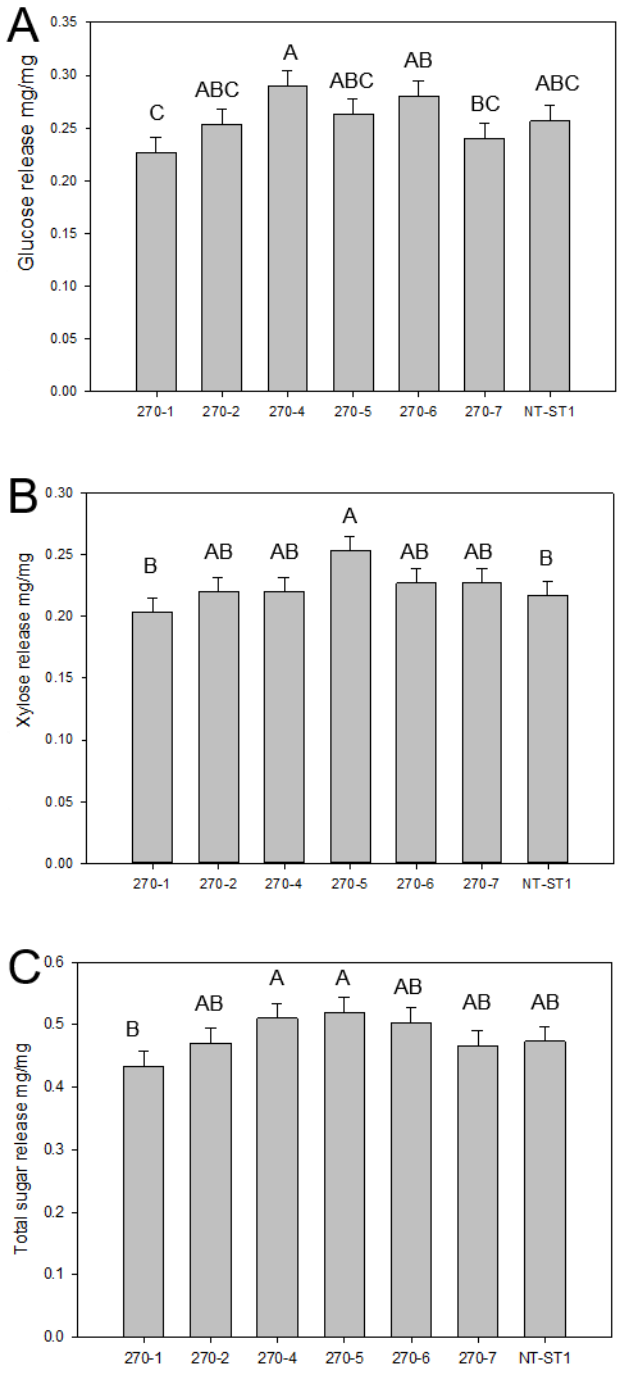

Figure 4-6 continued

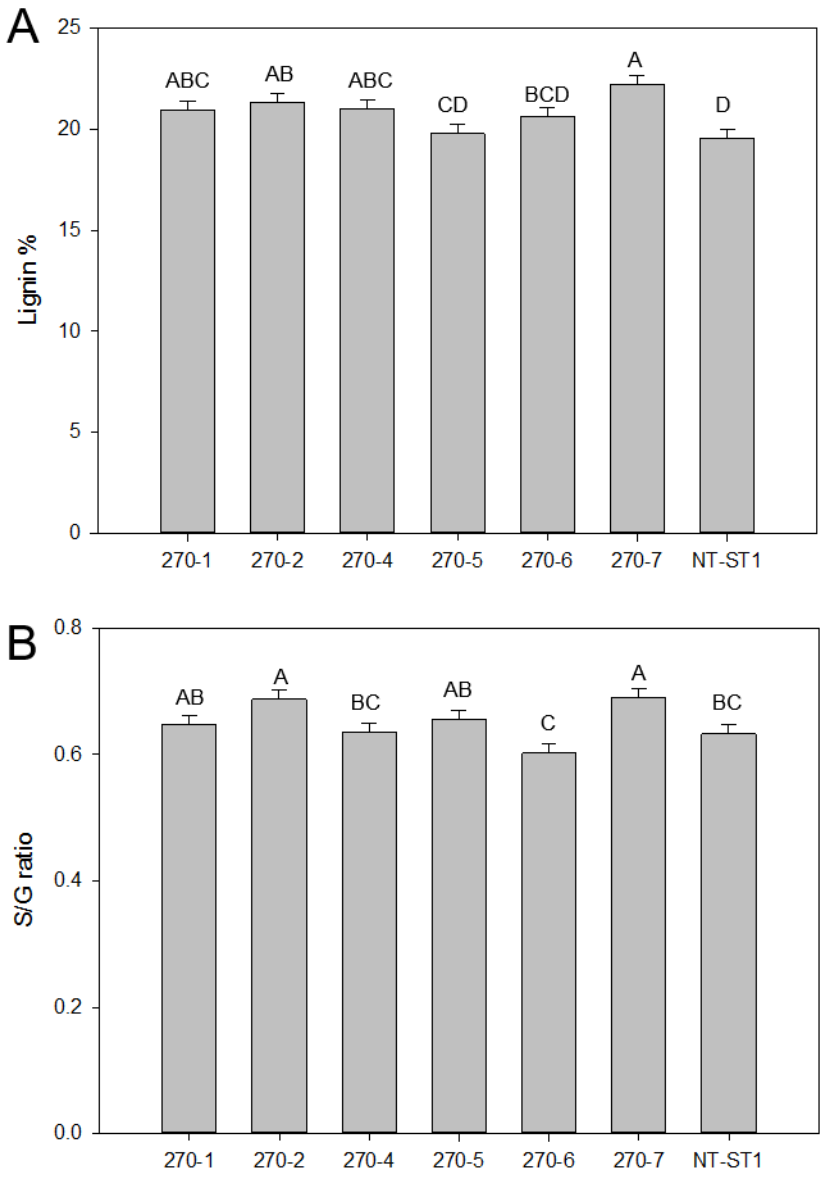

Figure 4-7 Lignin content (A) and S/G ratio (B) of down-regulated PvUAM1 transgenic and non-transgenic (NT-ST1) whole tiller cell wall residues as determined by PyMBMS. Bars represent mean values of three replicates ± standard error. Bars represented by same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

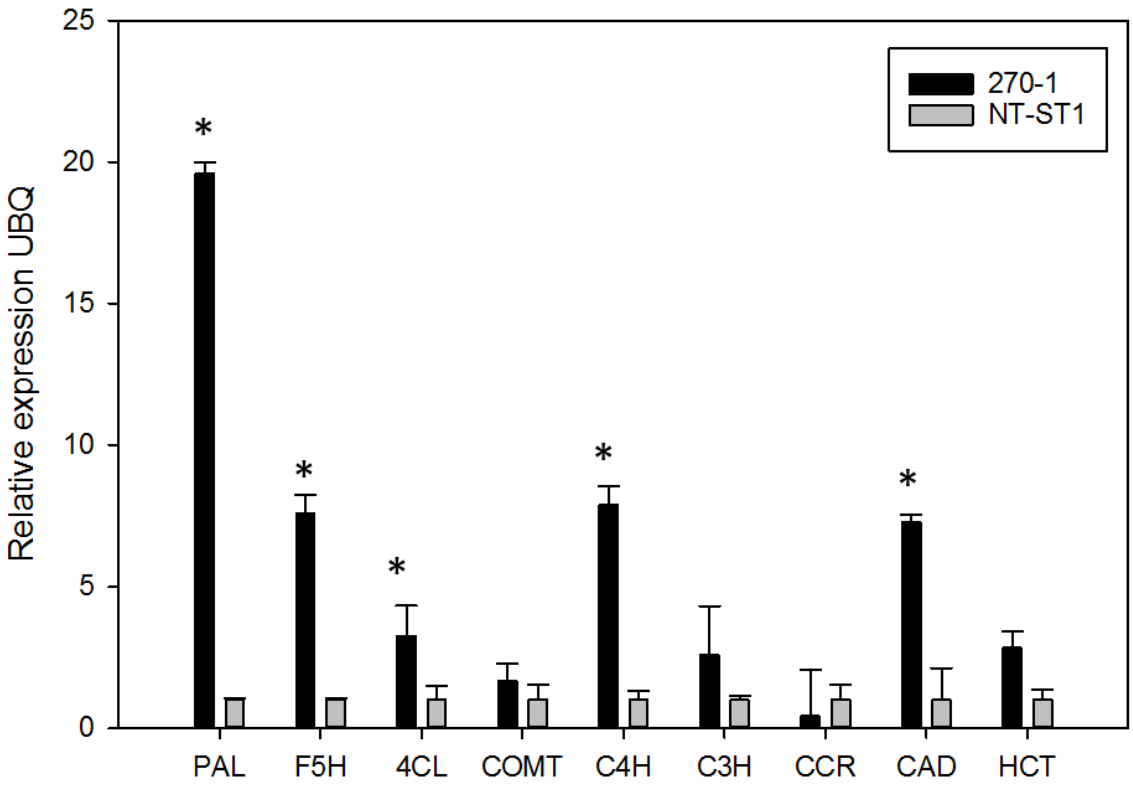

Figure 4-8 Relative expression of lignin biosynthetic genes in transgenic (270-1) and non-transgenic (NT-ST1) stem internodes as determined by qRT-PCR. The relative expressions were normalized to switchgrass ubiquitin 1 (*PvUbi1*). Bars represent mean values of three replicates ± standard error. Asterisks indicate significant differences from non-transgenic control plants at $p \leq 0.05$ as tested by LSD. PAL, phenylalanine ammonia-lyase; F5H, ferulate 5-hydroxylase; 4CL, 4-coumarate: CoA ligase; COMT, caffeic acid 3-Omethyltransferase; C4H, coumaroyl shikimate 4-hydroxylase; C3H, coumaroyl shikimate 3-hydroxylase; CCR, cinnamoyl CoA reductase; CAD, cinnamyl alcohol dehydrogenase; HCT, hydroxycinnamoyl.

Figure 4-9 Proposed model of arabinoxylan and lignin biosynthesis pathway interactions for down-regulated PvUAM1 transgenic switchgrass. Biosynthesis proteins denoted in black ovals: Ces, cellulose synthase, UGD, Uridine diphosphate (UDP)-glucose dehydrogenase; UXS, UDP-D-xylose synthase, XS, xylan synthase, XYNB, xylan synthase B, UXE, UDP-xylose esterase, UAM, UDP-arabinomutase (red lettering), PAL, phenylalanine ammonia-lyase; C4H, coumaroyl shikimate 4-hydroxylase; C3H, coumaroyl shikimate 3-hydroxylase; COMT, caffeic acid 3-O-methyltransferase; F5H, ferulate 5-hydroxylase; 4CL, 4-coumarate: CoA ligase; Ftase?, undetermined feruloyl transferase. Red arrows indicated up-regulated genes verified by qRT-PCR. Black arrow indicated down regulation of gene verified by qRT-PCR. Green arrows indicated cell wall components which have been increased. Metabolites are denoted by orange boxes: UDP-Glc, UDP-glucose; UDP-Glca, UDP-glucuronic acid; UDP-D-xyl, UDP-D-xylose; UDP-L-Ara*p*, UDP-L-arabinopyronose; UDP-L-Ara*f*, UDP-L-arabinofuranose. Blue arrows denote suspected down regulation of metabolites resulting from down regulation of *PvUAM*.

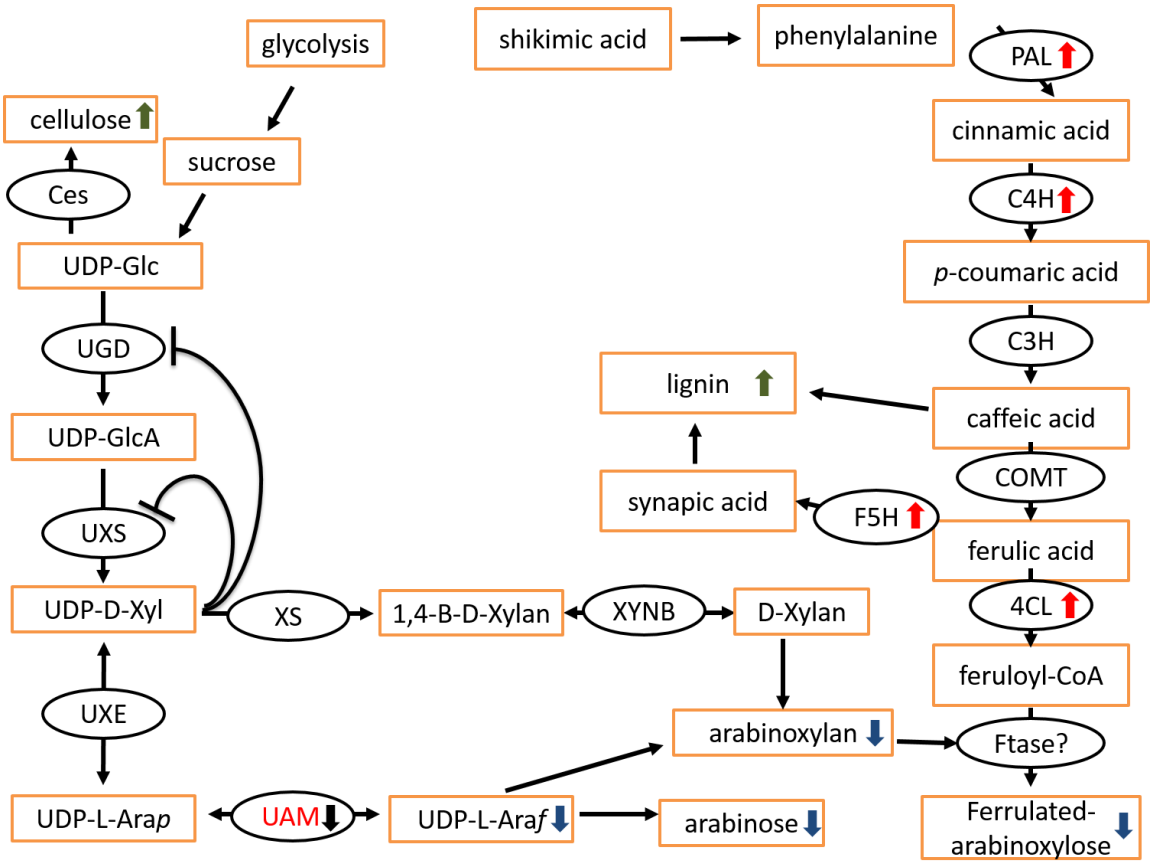

Figure 4-9 continued

4.9.2 Supplementary figures and tables

atggcgggcacggtgacggtcccgggggcgaacgtcccctccacgccgctgctcaaggacgagctcgacatcgtgatcccgacg
atccgcaacctcgacttcctggagatgtggcggcccttcttccagccgtaccacctcatcatcgtgcaggacggcgacccgaccaa
gaccatcaaggtgcccgagggcttcgactacgagctctacaaccgcaacgacatcaaccgcatcctcggccccaaggcatcctgc
atctccttcaaggactccgcctgccgctgcttcggctacatggtctccaagaagaagtacatctacaccatcgacgacgactgcttc
gttgccaaggatccatctggcaaggacatcaatgctcttgagcagcacatcaagaacctc**ctcagcccatccaccccgttcttcttt**
aacacgctgtatgacccctaccgtgagggtgctgactttgtgcgtgggtacccccttcagcctcagggagggtgcccacactgctgt
ctcccacggcctgtggctgaacatccctgactatgatgctcccacacagctggtcaagcctaaggagaggaatgaaaggtatgtt
gatgctgtcatgacaatccccaagggaaccttgttccccatgtgcggcatgaaccttgccttcgacagggatctcattggccctgct
atgtacttcggtctcatgggtgatggccagcctattggtcgctacgacgacatgtgggctgggtggtgtgtgaaggtcatctgcgac
cacttgagcctgggtgtcaagactggcctgccatacatctggcacagcaaggctagcaacccccttcgttaacttgaagaaggaata
caagggcatcttctggcaggaggacatcatccccttcttccagaacgtgaccatccccaaggagtgcgacacggtccagaagtgct
acatctacctttctgggcaggtgaaggagaagctaggcaagatcgacccctacttcgtcaagcttgccgacgccatggtcacctgg
atcgaggcctgggatgagctgaacccaaccgccgccgccgctgcgaacggcaaggccaagtag

Figure 4-S1 Full-length coding sequence of PvUAM1open reading frame. The 193 bp sequence selected for use in the RNAi cassette is bolded and underlined.

Figure 4-S2 PvUAM-RNAi vector map and Southern blot for transgene insertion analysis. A) The *PvUAM*-RNAi transformation vector used the pANIC8A RNAi plant expression vector. Key features are shown including the underlined *hph* hygromycin selectable marker gene, which was used as the probe for Southern blot analysis. *NcoI* with arrows make sites where NcoI restriction enzyme cut for Southern blot analysis. LB = Left border, OsAct1 = Rice actin promoter and intron, hph = hygromycin B phosphotransferase coding region, 35S T= 35S Terminator sequence, PvUbi1 =Switchgrass ubiquitin 1 promoter and intron, pporRFP = *Porites porites* red fluorescent protein coding region, NOS T = *Agrobacterium tumefaciens nos* terminator sequence, ZmUbi1 = Maize ubiquitin 1 promoter, R1 and R2 = *attR1* and *attR2* recombinase sites 1 and 2, *uam1* = *UAM1* sequence fragment for RNAi, RB = Right border, Kan^r = kanamycin resistance gene for bacterial selection, ColE1 = origin of replication in *E. coli*, pVS1 = origin of replication in *A. tumefaciens*, OCS T= octopine synthase terminator sequence. B) Southern blot analysis of transgenic events and non-transgenic control (NT-ST1) switchgrass genomic DNA digested by *NcoI*. The hygromycin resistance (*hph*) gene was used as probe. P1 is an uncut pANIC8A-PvUAM1 plasmid, P2 is the linearized pcr4-hph plasmid control. MW1 is the Hi-Lo DNA marker with base pair values listed. MW2 is the DIG DNA marker III with base pair values listed.

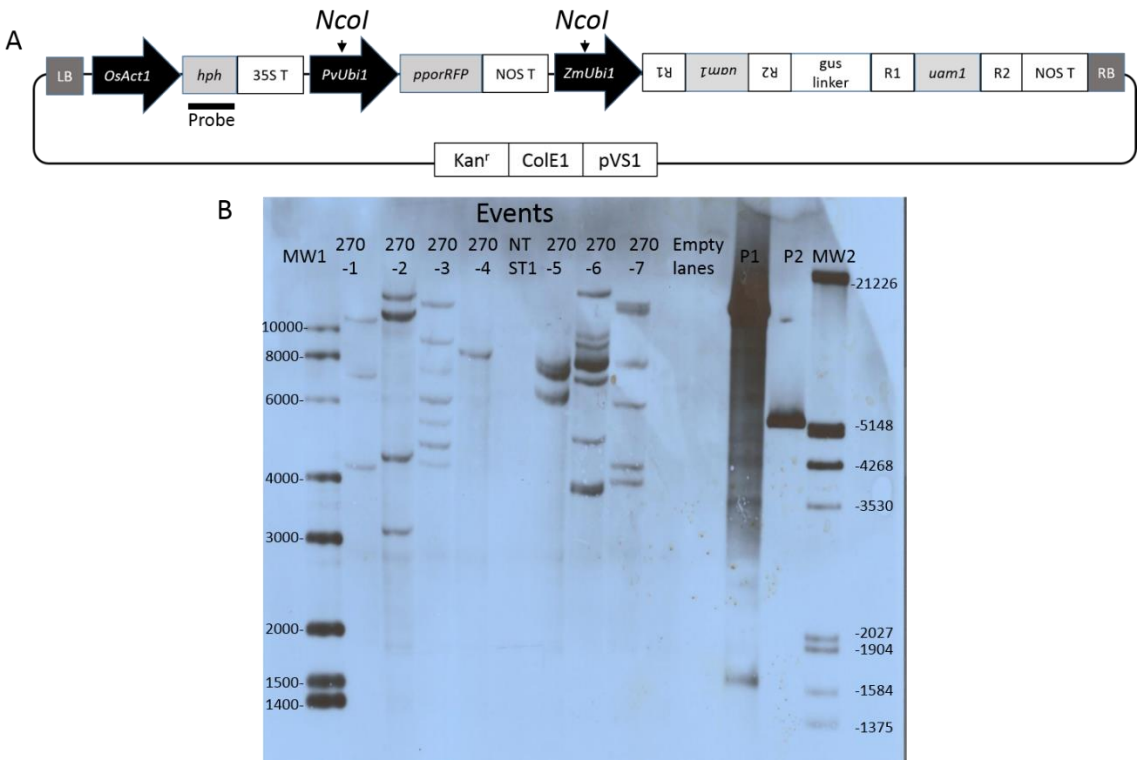

Figure 4S-2 continued

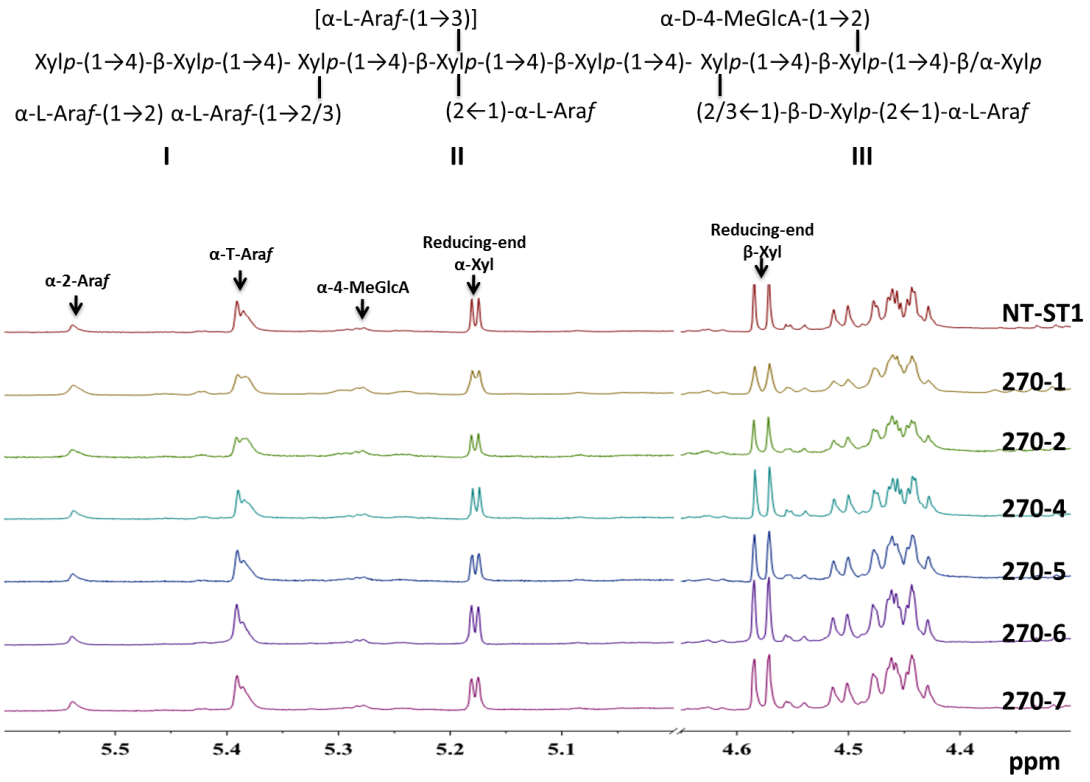

Figure 4-S3 Arabinoxylan schema with NMR peak assignments for stem samples.

Arabinoxylan sidechains are **I;** α-L-Ara*f*-(1→2)-α-L-Ara*f*, **II**; mono- or di-substituted α-L-Ara*f*, and **III**; β-D-Xyl*p*-(1→2)-α-L-Ara*f* at O-2 and/or O-3 of β-Xyl*p*-(1→4) backbone with some degree of (1→2)-linked α-D-4-MeGlcA. The α-2-Ara*f* anomeric signal represents H1 of "O-2-substituted" α-L-Ara*f* (as in sidechain **I**). The α-T-Ara*f* anomeric signal represents H1 of "terminal" α-L-Ara*f* residues (as in sidechains **I**, **II**, and **III**).

Table 4-S1 List of primers used for qRT-PCR analysis of *PvUAM1, PvUAM2,* and *PvUAM3*.

Primer name	Primer Sequence
PvUAM1-F	AACGTGACCATCCCCAAGGAGTG
PvUAM1-R	GAAGTAGGGGTCGATCTTGCCC
PvUAM2-F	GAGAAATCAGCGGAACACAACC
PvUAM2-R	GAAAATAGCAGGCCCCAACACC
PvUAM3-F	GGCATCTTCTGGCAAGAGGAGC
PvUAM3-R	CTTCGCCCTCACCTGCTTGGCC

Table 4-S2 List of primers used for qRT-PCR analysis of lignin biosynthetic genes.

Primer name	Primer sequence	Reference
C4H1_1534F	GGGCAGTTCAGCAACCAGAT	Shen, H. et al., 2012
C4H1_1611R	CGCGTTTCCGGGACTCTAG	Shen, H. et al., 2012
PvCOMT_F461	CAACCGCGTGTTCAACGA	Shen, H. et al., 2012
PvCOMT_R534	CGGTGTAGAACTCGAGCAGCTT	Shen, H. et al., 2012
4CL1_1179_F	CGAGCAGATCATGAAAGGTTACC	Shen, H. et al., 2012
4CL1_1251_R	CAGCCAGCCGTCCTTGTC	Shen, H. et al., 2012
PvCCR1. 112_F	GCGTCGTGGCTCGTCAA	Shen, H. et al., 2012
PvCCR1. 187_R	TCGGGTCATCTGGGTTCCT	Shen, H. et al., 2012
PvCAD_F116	TCACATCAAGCATCCACCATCT	Shen, H. et al., 2012
PvCAD_R184	GTTCTCGTGTCCGAGGTGTGT	Shen, H. et al., 2012
HCT_973_F	GCAGAAGGAGCAGCAGTCATC	Shen, H. et al., 2012
HCT_1035_R	CGAGCGGCAATAGTCGTTGT	Shen, H. et al., 2012
PAL_F1	CATATAGTGTGCGTGCGTGTGT	Wuddineh, W. et al., 2014
PAL_R1	CTGGCCCGCCAATCG	Wuddineh, W. et al., 2014
C3H_F1	CGTGAACAATGGGATCAGGATAG	Wuddineh, W. et al., 2014
C3H_R1	GCGGACACAACCATCTCAAATAC	Wuddineh, W. et al., 2014
F5H_F1	CCCCGTGCACTGACGATCTAT	Wuddineh, W. et al., 2014
F5H_R1	CCAAGCCAAGGGAAAACACAGTTA	Wuddineh, W. et al., 2014
F_PvUBIQUITIN	CAGCGAGGGCTCAATAATTCCA	Xu, B. et al., 2011
R_PvUBIQUITIN	TCTGGCGGACTACAATATCCA	Xu, B. et al., 2011

Table 4-S3 Galactose, rhamnose, and mannose content in stems of transgenic and non-transgenic (NT-ST1) events as determined by gas chromatography. Samples were normalized to the non-transgenic control with mol% representing the % of total cell wall-associated sugars measured. Values represent mean of three replicates ± standard error. Values with the same letters are not significantly different as calculated by LSD ($p \leq 0.05$).

Event	Galactose (mol%)	LSD	Rhamnose (mol%)	LSD	Mannose (mol%)	LSD
270-1	1.61 ± 0.29	BC	0.22 ± 0.05	B	0.17 ± 0.15	A
270-2	1.41 ± 0.29	C	0.26 ± 0.05	AB	0.35 ± 0.15	A
270-4	2.45 ± 0.29	AB	0.21 ± 0.05	B	0.37 ± 0.15	A
270-5	1.59 ± 0.29	BC	0.24 ± 0.05	B	0.27 ± 0.15	A
270-6	2.66 ± 0.29	A	0.30 ± 0.05	AB	0.35 ± 0.15	A
270-7	2.31 ± 0.29	AB	0.30 ± 0.05	AB	0.30 ± 0.15	A
NT-ST1	2.73 ± 0.29	A	0.40 ± 0.05	A	0.23 ± 0.15	A

Table 4-S4 Galactose, rhamnose, and mannose content in leaves of transgenic and non-transgenic (NT-ST1) lines as determined by gas chromatography. Samples were normalized to the non-transgenic control with mol% representing the % of total cell wall-associated sugars measured. Values represent mean of three replicates ± standard error. Values with the same letter are not significantly different as calculated by LSD ($p \leq 0.05$).

Event	Galactose (mol%)	LSD	Rhamnose (mol%)	LSD	Mannose (mol%)	LSD
270-1	2.11 ± 0.19	B	0.65 ± 0.08	B	0.62 ± 0.15	A
270-2	2.56 ± 0.19	AB	1.06 ± 0.08	A	0.29 ± 0.15	A
270-4	2.18 ± 0.19	B	0.64 ±0.08	B	0.42 ± 0.15	A
270-5	2.38 ± 0.19	AB	0.78 ± 0.08	B	0.18 ± 0.15	A
270-6	2.26 ± 0.19	AB	0.69 ± 0.08	B	0.18 ± 0.15	A
270-7	2.76 ± 0.19	A	0.76 ± 0.08	B	0.26 ± 0.15	A
NT-ST1	2.38 ± 0.19	AB	0.72 ± 0.08	B	0.25 ± 0.15	A

Chapter 5: Dissertation conclusion

Lignocellulosic feedstocks are renewable alternatives to fossil fuel feedstocks, but their cell walls are recalcitrant to degradation into simpler sugars. When the simpler sugars are released from the cell wall they can be fermented into ethanol, which can be used as fuel additive (example, in E85 gasoline). Genetically engineering the feedstocks for reduced cell wall recalcitrance would reduce the economic costs associated with cellulosic ethanol production. At the inception of the DOE bioenergy research centers, lignin was the main the target for manipulation of biofuel feedstocks. Since their inception, research has elucidated new cell wall dynamics of recalcitrance beyond the lignin barrier. Over time research showed reducing lignin was not the only component of the cell wall which could be targeted for improved yields. Genetic manipulation of cell wall gens and screening of wild plant populations for different cell wall phenotypes have demonstrated a greater understanding of cell wall biosynthesis. Switchgrass and poplar germplasm with improved saccharification have been developed altering cell wall synthesis pathways. Trait combinations of transgenic lines with field assessments are underway and will demonstrate the true efficacy for commercial production.

Future goals for the bioenergy feedstock community will include environmental goals of carbon sequestration and advanced genetic manipulation of whole biosynthesis pathways. Carbon sequestration into below ground biomass will further decrease the carbon-foot print of biofuel feedstocks and help reduce greenhouse gasses. New advances in genome editing will assist with the manipulation of whole biosynthesis pathways. Combinations of these two goals could yield

high biomass with increased carbon sequestration, biotic and abiotic stress tolerance, and potential development of "drop-in-fuel" biofuel feedstocks.

Put forth in this dissertation were two potential ways to reduce cell wall recalcitrance is to incorporate cellulase enzymes into the feedstock genome for production of its own cell wall degrading enzymes or altering hemicellulose cross linkage to lignin.

There is significant cost associated with cellulase enzymes for conversion of cellulose into simpler sugars for cellulosic ethanol production. Transgenic plants, have been used as hosts for overexpression of microbial glycosyl hydrolases from bacterial and fungal origins to reduce such costs. Insect-derived glycosyl hydrolases could play a role in reducing cell wall recalcitrance in transgenic plants owing to their high diversity of function and eukaryotic origin. Incorporation of multiple cellulase classes within feedstocks could be used in concert for *in planta* biomass degradation. When coupled with appropriate subcellular targeting as well as spatial and temporal control of synthesis transgenic glycosyl hydrolase feedstocks can result in no deleterious growth effects. Field research with bacterial cellulase producing transgenic maize demonstrates that mass plant produced cellulase enzymes is feasible.

TcEG-1 switchgrass is the first study in which an active alkaline insect cellulase was introduced and produced in any plant; specifically a dedicated bioenergy crop, switchgrass. All ten independent transgenic TcEG-1 events showed endoglucanase activity. However, the enzyme activity was decreased in oven-dried biomass compared to air-dried biomass. Sugar release

increased in one transgenic event, which also had the lowest amount of lignin of any transgenic plants in the study. Histological examination demonstrated no change to cell area or perimeter, but cell walls were wider. Cellulose crystallinity was suspected to cause the widening, however there was no correlation observed between cellulose crystallinity to saccharification or enzyme activity. Transgenic plants had narrower stems than the control, but the thicker cell walls and increase in tiller number aided in maintenance of possessing equivalent dry biomass yield. The combination of dedicated biomass feedstocks and enzyme technologies is a step toward a renewable plant-based fuel system.

Genetic manipulation of the hemicellulose pathway was done by identifying a switchgrass UDP-arabinomutase (*PvUAM*1) and downregulating it in switchgrass by RNAi. Switchgrass *PvUAM1-RNAi* transformants had decreased cell wall-associated arabinose content with a concurrent increase in cell wall-associated glucose and lignin content. A hypothesized model is that due to the decrease of available arabinoxylan caused an increase in lignin content due to excess metabolites that were not incorporated for arabinoxylan-lignin cross linking. Although enzymatic sugar release was not negatively impacted by the increase in lignin content possibly because there was more available cell wall-associated glucose in transgenic leaves and stems. Some transgenic *PvUAM1* plants were taller and produced more biomass than control, which would be useful for commercial biomass and carbon sequestration platforms as well as a lignin feedstock.

In conclusion, presented here is a review and methodology for creating plant systems that can produce their own cellulase enzymes needed for cellulosic ethanol production and

characterization of two different transgenic switchgrass lines. Both sets of transgenic experiments targeted currently unexplored methods of reducing switchgrass cell wall recalcitrance. Both transgenic lines adds to the switchgrass cell wall recalcitrance literature, which could assist in future genetic engineering of biofuel feedstocks.

Vita

Jonathan Duran Willis was born August, 23rd 1984 to the parents of James and Pam Willis in Watkinsville, GA. He graduated from Oconee County High School with Honors in 2002. Taking minimal time off between schools he quickly started his education at the University of Georgia during the summer semester. He began working the following fall in Dr. Michael Scanlon's lab in the Plant Biology department as a student worker evaluating plant development genes in maize. After three years he moved to serve as a Student Services Contractor as a lab manager for Dr. Marirosa Molina at the Environmental Protection Agency examining fecal bacteria detection techniques for water systems. Jonathan completed his BS in Applied Biotechnology and moved to Knoxville, Tennessee to begin his graduate career pursuing a MS in Entomology with Dr. Juan-Luis Jurat-Fuentes studying insect digestive system particularly their cellulose digesting cellulase enzymes. After completing his MS he switched science areas to plant sciences under Dr. Neal Stewart for completion of a PhD in Plant, Soils, and Insects. During his PhD Jonathan evaluated carbohydrate active enzyme manipulation in the biofuel crop switchgrass. In May 2015 he was married to his love Kazuyo Shimizu and live happily with their dogs. Jonathan will be joining the USDA in the summer of 2016 to continue work on transgenic switchgrass and develop methods for controlling transgenic gene flow.