MODIFYING CELL WALL COMPOSITION OF LIGNOCELLULOSIC SUGARCANE BIOMASS BY RNAI SUPPRESSION OF CAFFEIC ACID O-METHYLTRANSFERASE TO ENHANCE BIOFUEL PRODUCTION

By

JE HYEONG JUNG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2012
To my family, mentors, and friends
for their love, encouragement, and guidance
ACKNOWLEDGMENTS

I would like to gratefully acknowledge my major advisor, Dr. Fredy Altpeter, for his guidance, encouragement, and financial support. His enthusiasm, kindness, and in-depth knowledge of research have inspired my career, and he has been a great role model. I would also like to thank my committee members. Special thanks to my Co-Advisor, Dr. Maria Gallo, whose continuous motivation and positive thinking helped me accomplish my research goals. I am also grateful to Dr. Wilfred Vermerris for helping me in building a solid foundation in this research. I especially acknowledge Dr. James Preston for his continual advice and commitment to discussions. I sincerely thank Dr. Gary Peter for the thought-provoking suggestions.

Drs. Wilfred Vermerris and Kevin Kenworthy provided useful advice while I was a Teaching Assistant in Genetics. This class provided great opportunities to actively engage students during the semester. Drs. John Erickson, Robert Gilbert, and Jim Boyer kindly provided the necessary field equipment to conduct my research.

Heartfelt thanks to all my current and former lab members. I especially thank Dr. Janice Zale for editing my manuscripts and dissertation. I am indebted to Drs. Jae Yoon Kim and Walid Fouad for their generous advice in research. I am appreciative for the friendships with Elizabeth, Marco, Yang, Yuan, Yogesh, Hao, and Baskaran.

I sincerely thank Dr. Seo, Yong Weon for his continuous guidance and motivation during my research career. I am grateful to Drs. Lee, Ho Joung and Kim, Wook for helping me realize how enjoyable research can be.

At last, my most sincere thanks go to my lovely family and wife for their love and support.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS .......................................................... 4</td>
</tr>
<tr>
<td>LIST OF TABLES ................................................................... 8</td>
</tr>
<tr>
<td>LIST OF FIGURES ............................................................... 9</td>
</tr>
<tr>
<td>ABSTRACT ........................................................................ 10</td>
</tr>
<tr>
<td>CHAPTER 1 LITERATURE REVIEW ........................................... 12</td>
</tr>
</tbody>
</table>

**Sugarcane** ........................................................................ 12
- Sugarcane as Superior Feedstock for Biofuel Production ........ 12
- Origin of Sugarcane Hybrids and Genome Structure .......... 13
- Genetic Engineering in Sugarcane .................................. 15

**Plant Cell Wall** .................................................................. 17
- Cell Wall Polysaccharides .............................................. 18
- Aromatic Substances .................................................... 19
- Structural Proteins ....................................................... 20

**Lignin** ............................................................................. 21
- Monolignol Biosynthesis ............................................... 21
- Lignin Polymerization and Deposition in the Secondary Cell Walls ............... 23
- Forward and Reverse Genetics in Lignin Biosynthesis and Effects of Manipulating Lignin Biosynthetic Genes on the Formation of Lignin and Plant Development/Growth ............................................. 26
  - Phenylalanine ammonia-lyase (PAL) .......................... 26
  - Cinnamate 4-hydroxylase (C4H) ................................. 27
  - 4-coumarate:CoA ligase (4CL) .................................. 27
  - Hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyl transferase (HCT) ..................................... 28
  - 3'-hydroxylase (C3'H) ............................................... 28
  - Caffeoyl CoA O-methyltransferase (CCoAOMT) .......... 29
  - Cinnamoyl CoA reductase (CCR) ............................... 29
  - Ferulate 5-hydroxylase (F5H) .................................... 31
  - Caffeic acid O-methyltransferase (COMT) ..................... 31
  - Cinnamyl alcohol dehydrogenase (CAD) ..................... 32

**Lignocellulosic Biofuel** .................................................... 33
- The Production Process of Lignocellulosic Ethanol ................. 35
- The Intrinsic Recalcitrance of Lignocellulosic Biomass for Ethanol Production .................................................. 38
- Improved Quality of Lignocellulosic Feedstock through Lignin Modification ........................ 39
2 ISOLATION AND EXPRESSIONAL CHARACTERIZATION OF CAFFEIC ACID O-METHYLTRANSFERASE (COMT) IN SUGARCANE ................................................................. 48

Introduction ........................................................................................................................................... 48
Materials and Methods ............................................................................................................................ 50
  Isolation of COMT Gene from CP88-1762 ....................................................................................... 50
  cDNA Library Construction and Screening ...................................................................................... 51
  Sequence Analysis ............................................................................................................................... 51
  Sampling for Expression Analysis of the COMT Gene ................................................................. 52
  RT-PCR and Quantitative Real-Time RT-PCR .............................................................................. 52
Results ...................................................................................................................................................... 53
  Isolation of COMT Genes Through cDNA Library Screening and PCR ...................................... 53
  Sequence Comparison Between the Isolated COMT and TCs .................................................... 54
  Deduced Amino Acid Sequence Comparison of the Isolated COMT with Other Monocot COMTs ........................................................................................................ 55
  Expression of the COMT in Different Tissues and Developmental Stages .................................. 55
Discussion ................................................................................................................................................ 56

3 GENERATION OF COMT SUPPRESSED TRANSGENIC SUGARCANE AND EFFECTS OF COMT SUPPRESSION ON LIGNIN, PLANT GROWTH, AND SACCHARIFICATION PERFORMANCE UNDER GREENHOUSE CONDITIONS ........................................................................................................ 66

Introduction .............................................................................................................................................. 66
Materials and Methods ............................................................................................................................ 69
  Plant Growth ...................................................................................................................................... 69
  Vector Construction ............................................................................................................................. 70
  Generation of Transgenic Sugarcane ................................................................................................. 71
  Evaluation of NPTII Expression ......................................................................................................... 72
  PCR Analysis ..................................................................................................................................... 72
  Southern Blot Analysis ......................................................................................................................... 73
  Small-RNA Northern Blot .................................................................................................................... 73
  Quantitative Real-Time RT-PCR for Quantification of COMT Expression .................. 74
  Microscopic and Histochemical Analysis ......................................................................................... 75
  Sample Preparation and Determination of Lignin Content and Composition ..................................... 76
  Dilute Acid Pretreatment and Enzymatic Hydrolysis ..................................................................... 78
  Statistical Analysis ............................................................................................................................... 78
Results ...................................................................................................................................................... 79
  Generation and Molecular Characterization of Transgenic Sugarcane Lines .................. 79
  Lignin Content, Composition, and Enzymatic Saccharification ................................................. 80
  Plant Phenotype and Growth .............................................................................................................. 81
Discussion ................................................................................................................................................ 83

4 FIELD EVALUATION OF COMT SUPPRESSED TRANSGENIC SUGARCANE ........................................... 93
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>93</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>95</td>
</tr>
<tr>
<td>Field Design</td>
<td>95</td>
</tr>
<tr>
<td>Evaluation of Plant Performance</td>
<td>96</td>
</tr>
<tr>
<td>Gene Expression Analysis</td>
<td>97</td>
</tr>
<tr>
<td>Sample Preparation for the Evaluation of Cell Wall Composition and</td>
<td>97</td>
</tr>
<tr>
<td>Saccharification Efficiency</td>
<td></td>
</tr>
<tr>
<td>Lignin Content and Composition</td>
<td>97</td>
</tr>
<tr>
<td>Analysis of Hydroxycinnamic Acids</td>
<td>98</td>
</tr>
<tr>
<td>Cell Wall Carbohydrates and Starch</td>
<td>99</td>
</tr>
<tr>
<td>Dilute-Acid Pretreatment and Enzymatic Hydrolysis</td>
<td>100</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>101</td>
</tr>
<tr>
<td>Results</td>
<td>101</td>
</tr>
<tr>
<td>5 COMCLUSIONS</td>
<td>117</td>
</tr>
<tr>
<td>APPENDIX: LABORATORY PROTOCOLS</td>
<td>120</td>
</tr>
<tr>
<td>Generation of Transgenic Sugarcane using Particle Bombardment</td>
<td>120</td>
</tr>
<tr>
<td>Small RNA Northern Blot</td>
<td>123</td>
</tr>
<tr>
<td>Analysis of Cell Wall Components</td>
<td>127</td>
</tr>
<tr>
<td>Evaluation of Saccharification Efficiency</td>
<td>130</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>133</td>
</tr>
<tr>
<td>BIOGRAPHICAL SKETCH</td>
<td>151</td>
</tr>
</tbody>
</table>
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Improvement of fermentable sugar and/or ethanol yields in lignin reduced transgenic plants</td>
<td>46</td>
</tr>
<tr>
<td>3-1</td>
<td>Lignin content, composition, and glucose yields after enzymatic saccharification</td>
<td>87</td>
</tr>
<tr>
<td>4-1</td>
<td>Lignin content and composition in transgenic sugarcane</td>
<td>110</td>
</tr>
<tr>
<td>4-2</td>
<td>Composition of structural carbohydrates of the cell wall in control and transgenic sugarcane plants</td>
<td>111</td>
</tr>
<tr>
<td>4-3</td>
<td>Recovery yields of cell wall bound $p$-coumaric acid ($p$CA) and ferulic acid (FA) after mild-alkaline hydrolysis of the control and transgenic sugarcane</td>
<td>112</td>
</tr>
<tr>
<td>4-4</td>
<td>Growth characteristics of the transgenic sugarcane lines under field conditions</td>
<td>113</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1</td>
<td>Lignin biosynthesis pathway adapted from Boerjan et al. (2003) and Bonawitz and Chapple (2010)</td>
<td>47</td>
</tr>
<tr>
<td>2-1</td>
<td>Nucleotide and amino acid comparisons among the isolated sugarcane COMTs</td>
<td>61</td>
</tr>
<tr>
<td>2-2</td>
<td>Nucleotide sequence comparison of the 3' UTR between COMTa and TCs</td>
<td>62</td>
</tr>
<tr>
<td>2-3</td>
<td>Nucleotide sequence comparison of coding regions spanning SAM binding residues between TCs and COMTa</td>
<td>63</td>
</tr>
<tr>
<td>2-4</td>
<td>Amino acid sequence comparisons between sugarcane COMTs and functionally characterized COMTs in monocot plants</td>
<td>64</td>
</tr>
<tr>
<td>2-5</td>
<td>COMT expression patterns in sugarcane</td>
<td>65</td>
</tr>
<tr>
<td>3-1</td>
<td>Generation and selection of transgenic sugarcane lines and investigation of transgene integration pattern in the transgenic sugarcane lines</td>
<td>88</td>
</tr>
<tr>
<td>3-2</td>
<td>Detection of COMT siRNA and evaluation of COMT expression levels in transgenic sugarcane</td>
<td>89</td>
</tr>
<tr>
<td>3-3</td>
<td>Phenotypic evaluation of transgenic sugarcane in comparison to wild-type</td>
<td>90</td>
</tr>
<tr>
<td>3-4</td>
<td>Microscopic and histochemical evaluation of transgenic sugarcane in comparison to wild-type</td>
<td>91</td>
</tr>
<tr>
<td>3-5</td>
<td>Plant growth characteristics under greenhouse conditions</td>
<td>92</td>
</tr>
<tr>
<td>4-1</td>
<td>Real time RT-PCR analysis of COMT expression level in transgenic sugarcane</td>
<td>114</td>
</tr>
<tr>
<td>4-2</td>
<td>Field trial of the COMT suppressed transgenic sugarcane lines</td>
<td>115</td>
</tr>
<tr>
<td>4-3</td>
<td>Enzymatic saccharification performance in wild-type (WT), transgenic control harboring nptI gene alone (TC), and transgenic sugarcane lines (T41 and T4)</td>
<td>116</td>
</tr>
</tbody>
</table>
MODIFYING CELL WALL COMPOSITION OF LIGNOCELLULOSIC SUGARCANE
BIOMASS BY RNAI SUPPRESSION OF CAFFEIC ACID O-METHYLTRANSFERASE
TO ENHANCE BIOFUEL PRODUCTION

By

Je Hyeong Jung

December 2012

Chair: Fredy Altpeter
Cochair: Maria Gallo
Major: Agronomy

Sugarcane (*Saccharum* spp. hybrids) is a prime bioethanol feedstock. Currently sugarcane ethanol is produced through the fermentation of sucrose which can easily be extracted from stem internodes. Processes for the production of biofuel from the abundant lignocellulosic sugarcane residues will boost the ethanol output from sugarcane per unit land area and unit biomass. However, unlocking the vast amount of chemical energy stored in plant cell walls remains expensive, primarily because of the intrinsic recalcitrance of lignocellulosic biomass to enzymatic hydrolysis. The presence of lignin in the plant cell wall has been recognized as a major limitation to efficient bioconversion of lignocellulosic biomass to biofuel. In this study, suppression of caffeic acid *O*-methyltransferase (*COMT*) in transgenic sugarcane greatly enhances the bioconversion efficiency of lignocellulosic biomass into fermentable sugars by reducing lignin content. *COMT* suppressed transgenic sugarcane lines have been successfully generated by deploying RNA interference. Suppression of *COMT* decreases lignin content and alters lignin composition with reduced levels of syringyl unit incorporation, while the accumulation of cell wall polysaccharides remains unaffected. More
importantly, the saccharification efficiency is improved up to 32% in transgenic sugarcane lines. Furthermore, transgenic sugarcane lines require one third of the hydrolyzing time and 3 or 4-fold less enzyme to produce an equal or greater amount of fermentable sugar compared with the control lines. Plant growth performance of transgenic sugarcane depended upon the extent of lignin reduction. A transgenic line with a 6% reduction in lignin exhibited comparable growth performance to control plants under both greenhouse and field conditions, while lignin reduction between 8% and 12% resulted in impaired growth. Reducing the recalcitrance of lignocellulosic sugarcane biomass by modifying lignin biosynthesis, without compromising plant growth performance, is expected to improve the economic feasibility of lignocellulosic biofuel production. Transferring the reduced lignin trait to other high biomass varieties of sugarcane will further increase the value of sugarcane as a superior biofuel feedstock.
CHAPTER 1
LITERATURE REVIEW

Sugarcane

Sugarcane (Saccharum spp. hybrids) is an economically important crop and is the main crop source for sugar and bioethanol production. It is a highly productive perennial C$_4$ grass grown in tropical, semi-tropical, and sub-tropical regions (Tew and Cobill, 2008). Its perennial growth habit and C$_4$ photosynthetic pathway maximize carbon assimilation, while minimizing water and nitrogen inputs (Byrt et al., 2011). In 2010, more than 1.7 billion Mg of sugarcane was produced using 2% of the total cultivated area in over 90 countries (FAOSTAT, http://faostat3.fao.org/home/index.html). Its production quantity, on a fresh weight basis, is higher than any other crop in the world (FAOSTAT, http://faostat3.fao.org/home/index.html). The net production value of sugarcane worldwide was $54 billion in 2010 (FAOSTAT, http://faostat3.fao.org/home/index.html).

Sugarcane as Superior Feedstock for Biofuel Production

Sugarcane is the most important crop for sugar production, accounting for approximately 70% of the world’s sugar supply (Tew and Cobill, 2008). It is also the second most important feedstock for bioethanol production after maize. In 2011, 21 billion liters of bioethanol were produced from sugarcane, accounting for 25% of the total bioethanol produced worldwide (RFA, 2012). Sugarcane accumulates as much as 50% of its dry weight as sucrose in the internodes (Waclawovsky et al., 2010). The production of sugarcane ethanol has been commercialized in Brazil with low production costs, which are 24% less than corn ethanol in the USA (Crago et al., 2010). Furthermore, it is considered an advanced biofuel due to its high environmental
sustainability (Goldemberg, 2007; Schnepf and Yacobucci, 2012). The energy balance of sugarcane ethanol is about seven times higher than corn ethanol, reflecting much less fossil fuel utilization during production, similar to lignocellulosic ethanol (Goldemberg, 2007). Land use efficiency, in terms of ethanol productivity per unit area, is 45% higher for sugarcane ethanol compared to corn ethanol (Crago et al., 2010).

There has been a great deal of attention paid to sugarcane as a promising feedstock for lignocellulosic ethanol production due to its exceptional biomass productivity and the availability of a large amount of post-harvest residues such as green tops, leaves and bagasse. Sugarcane’s dry biomass production per unit area (39 Mg ha\(^{-1}\); stalk, leaves and tops) is significantly higher than maize (17.6 Mg ha\(^{-1}\); grain and stover), switchgrass (10.4 Mg ha\(^{-1}\); biomass), or Miscanthus (29.6 Mg ha\(^{-1}\); biomass) (Heaton et al., 2008; Waclawovsky et al., 2010). Post-harvest residues and bagasse comprise approximately 55% of the total above-ground biomass (Somerville et al., 2010; Tew and Cobill, 2008). Together with sugar-based bioethanol, the utilization of abundant lignocellulosic biomass further increases ethanol productivity per unit land area or per unit biomass. Leite et al. (2009) suggested that assuming all of the post-harvest residues and 50% of the bagasse are converted to bioethanol, 33~38% less land is required to produce an equal amount of sugar-based bioethanol. Furthermore, Somerville et al. (2010) projected that bioconversion of all of the bagasse could increase total bioethanol yield by 59%.

**Origin of Sugarcane Hybrids and Genome Structure**

Sugarcane belongs to the genus *Saccharum* L. in the *Poaceae* family. Six species have been traditionally classified as *Saccharum* genera, two of which are *S. robustum* and *S. spontaneum* and four historically cultivated species, *S. officinarum*, *S.
Among the species, *S. barberi* and *S. sinense* are thought to be natural interspecific hybrids between *S. spontaneum* and *S. officinarum*, while the origin of *S. edule* is unclear (Piperidis et al., 2010).

Current sugarcane cultivars are artificial interspecific hybrids between *S. officinarum* as female and *S. spontaneum* as male. After hybridization, F₁ hybrids are backcrossed several times to *S. officinarum* to recover high sugar yielding traits (Piperidis et al., 2010). *S. officinarum* is thought to have been domesticated from *S. robustum* and is characterized by thick stalks and high sugar content, but less tillering/ratooning ability and tolerance to biotic and abiotic stresses. *S. spontaneum* is a vigorous wild species which is genetically more diverse than *S. officinarum*. It develops relatively thin and tall stalks and contains high fiber with low sugar content. *S. spontaneum* has relatively higher biotic and abiotic tolerance than *S. officinarum* (Tew and Cobill, 2008).

Both of the parental species are considered to have an autopolyploid origin. *S. spontaneum* has a wide range of chromosome numbers and ploidy levels ranging from 2n=5x=40 to 2n=16x=128, while *S. officinarum* has 2n=8x=80. The genome of the interspecific hybrids is aneuploidy with a high level of polyploidy (average 12x) consisting of ~120 chromosomes. Seventy-eighty percent of the current sugarcane genome is derived from *S. officinarum*, 10-20% from *S. spontaneum*, and ~10% from recombination (D’Hont, 2005; Piperidis et al., 2010).

Allopolyploidization is a process of genome duplication after interspecific hybridization, and it often induces massive genome restructuring, followed by gene silencing and the subsequent loss of duplicated genes (Adams and Wendel, 2005).
Genes retained in the duplicated genome often undergo divergence in gene function and expression (Wang et al., 2012). In general, autopolyploids tend to experience less genome reshaping than allopolyploids (Parisod et al., 2010; Wang et al., 2012). The sugarcane genome resulting from interspecific hybridization between two parental autopolyploids is thought to be conserved without extensive gene loss or genome rearrangement (Garsmeur et al., 2011; Jannoo et al., 2007; Wang et al., 2010). The sugarcane genome is expected to have a high degree of functional genetic redundancy having a high level of collinearity, structural conservation, and sequence similarity/identity among homo(eo)logous haplotypes (Garsmeur et al., 2011; Jannoo et al., 2007).

**Genetic Engineering in Sugarcane**

Conventional sugarcane breeding combined with better agronomic practices have improved yields, disease resistance, and stress tolerance (Jackson, 2005). However, conventional breeding is challenging due to the lack of diversification within the sugarcane germplasm, the long breeding cycles requiring more than ten years, its complex genome structure and limited genome information (Dal-Bianco et al., 2011; Jackson, 2005). Therefore, genetic engineering through transgenic approaches will play a critical role in overcoming the limitations of conventional breeding.

Sugarcane is an attractive target for genetic transformation (Altpeter and Oraby, 2010). *Invitro* culture of sugarcane has been advanced by techniques for propagation and the generation of disease-free clones, somaclonal variants, and transgenic plants (Chengalrayan and Gallo-Meagher, 2001; Gallo-Meagher and Irvine, 1996; Lakshmanan et al., 2006; Singh et al., 2008). Vegetative propagation of transgenic sugarcane may provide stable transgene expression over generations and prevent the
segregation of stacked transgenes (Altpeter and Oraby, 2010). In addition, poor pollen viability of sugarcane along with its amenability to vegetative propagation will ensure a high level of transgene containment, which is a favorable characteristic in terms of biosafety (Bonnett et al., 2008).

Bower and Birch (1992) successfully generated the first stable transgenic sugarcane incorporating the nptII gene using microprojectile bombardment. Subsequently, a number of transgenic sugarcane lines have been generated through both particle bombardment (biolistics) and Agrobacterium-mediated gene transfer, mainly targeting herbicide resistance, tolerance to abiotic or biotic stress, increased sugar yields, and the production of value-added metabolites (Altpeter and Oraby, 2010; Arencibia et al., 1998; Gallo-Meagher and Irvine, 1996). Biolistics is applicable to a wide variety of cell types and explants, without genotype dependency, and it is favorable for integrating multiple expression cassettes for gene stacking (Altpeter et al., 2005). Utilization of minimal expression cassettes in biolistics prevents host plant contamination with vector backbones. Furthermore, although Agrobacterium-mediated gene transfer tends to result in simpler transgene integration patterns, this also can be achieved by biolistics utilizing a low concentration of minimal expression cassettes and optimizing the microprojectile size (Jackson et al., 2012; Kim et al., 2012; Taparia et al., 2012b).

Somatic embryogenesis and regeneration are essential for the generation of transgenic plants. Somatic embryos can be produced indirectly from callus induced from meristematic tissues or directly without an intervening callus phase (Arnold, 2008). Immature leaf whorls, immature inflorescences, or basal shoot apical meristems have
been used as explants for somatic embryogenesis (Altpeter and Oraby, 2010). Among these, the immature leaf whorl is ideal because it is available year round (Taparia et al., 2012a). Indirect somatic embryogenesis has been routinely used for biolistic gene transfer in sugarcane (Altpeter and Oraby, 2010). Sugarcane callus culture can be extended without compromising embryogenic potential. However, a prolonged callus phase/tissue culture period and delayed regeneration increase the rate of somaclonal variation often resulting in poor agronomic performance (Basnayake et al., 2011; Hoy et al., 2003; Vickers et al., 2005). Direct somatic embryogenesis with rapid regeneration could reduce the risk of undesirable somaclonal variants and the overall time required for the generation of transgenic plants (Lakshmanan et al., 2006). Cell fate is primarily determined by auxin to cytokinin ratio during *invitro* culture (Arnold, 2008). Direct somatic embryogenesis from immature leaf whorls of sugarcane cultivar, CP88-1762, is achieved with *p*-chlorophenoxyacetic acid:*α*-napthaleneacetic acid:6-benzylaminopurine in the ratios of 1.86:1.86:0.09 mg L⁻¹ in the tissue culture medium (Taparia et al., 2012b). The generation of transgenic sugarcane via direct somatic embryogenesis requires only 12 weeks, which is 12-24 weeks faster than those obtained through indirect somatic embryogenesis (Taparia et al., 2012b).

**Plant Cell Wall**

Plant cell walls have key roles in plant structural integrity, plant morphogenesis, water/nutrient transport, plant defense, cell differentiation, and cell communication (Cosgrove, 2005). They also are the main source for food, feed, and bio-based products, which contain abundant organic carbon in the form of cell wall polysaccharides, structural proteins, and various aromatic compounds (McCann and Rose, 2010; Vermerris, 2008). The cell wall is highly complex and heterogeneous, and its chemical
composition and structure vary among different tissues and cell types reflecting diverse and differentiated functions of the cell walls (Carpita and McCann, 2000). It is primarily classified into primary and secondary walls. The primary wall formed during cell division and expansion is mostly composed of cellulose, hemicelluloses, pectin, and a small amount of structural proteins (Carpita and McCann, 2000). The secondary cell is formed interior to the primary wall after cell growth ceases in the specialized cell types such as vessels and fibers (Donaldson, 2001). One of the distinct characteristics of these secondary cell walls is the massive impregnation of lignin into the cell wall matrix (Bonawitz and Chapple, 2010).

**Cell Wall Polysaccharides**

The most abundant polysaccharide in the cell wall is cellulose. Fifteen to thirty percent of the primary cell wall consists of cellulose on a dry weight basis, and the amount is even higher in secondary cell walls (Carpita and McCann, 2000). Cellulose is present in the form of microfibrils composed of an average of three dozen \((1\rightarrow4)-\beta-D\)-glucan chains interconnected by hydrogen bonds (Vermerris, 2008). Cellulose has a crystalline structure and its hydrophobic surface limits the accessibility of cellulolytic enzymes (Himmel et al., 2007). In growing cell walls, cellulose is embedded and interlocked in a cell wall matrix of hemicelluloses and pectins (Sørensen et al., 2010).

Hemicelluloses (cross-linking glycans) are heterogeneous polymers of polysaccharides cross-linked with one another or cellulose microfibrils (Carpita and McCann, 2000). Glucuronoarabinoxylans (GAXs) are the major hemicellulose in the cell wall of grasses and commelinoid monocots (Carpita, 1996). GAXs have xylan backbones with side chains of arabinose and glucuronic acid attached to the \(O-3\) and \(O-2\) positions of the backbone, respectively (Carpita, 1996). Less branched GAXs
confer more cell wall rigidity resulting from a relatively high degree of hydrogen bond formation with other GAXs and cellulose microfibrils (Vermerris, 2008). Xyloglucans (XyGs) are predominant hemicelluloses in dicots and non-commelinoid monocots (Carpita, 1996). XyGs have \((1\rightarrow4)\)-\(\beta\)-D-glucan chains as backbones with side chains of xylose. Some of the xylosyl units can be substituted with arabinose or galactose, and galactosyl units are sometimes substituted with fucose (Carpita and McCann, 2000).

Pectins are a heterogeneous group of polysaccharides. In the primary wall of grasses and commelinoid monocots, the amount of pectin (2-10%) is relatively less than that of dicots and non-commelinoid monocots (~35%) (Mohnen, 2008). Two main constituents of pectins are homogalacturonan (HG) and rhamnogalacturonan I (RG I) (Mohnen, 2008). HG has a backbone of \((1\rightarrow4)\)-\(\alpha\)-D-galacturonic acid (GaIA). It is often methyl esterified at the C-6 position and structurally modified to xylogalacturonan (XGA) and rhamnogalacturonan II (RG II). XGAs have side chains of xylosyl units at the O-3 position of GaIA. RG II is highly complex substituted HG with heterooligomeric side chains consisting of several sugars. RG I has a backbone of the disaccharide units, \((1\rightarrow2)\)-\(\alpha\)-L-rhamnosyl-\((1\rightarrow4)\)-\(\alpha\)-D-GaIA with side chains. Arabinans, galactans, and/or branched arabidnogalactans are attached to rhamnosyl units in RG I. These pectic polysaccharides are often cross-linked to each other forming macromolecular pectin network (Vincken et al., 2003). Non-esterified carboxyl residues of GaIA are involved in calcium ion-cross-linking of HGs. Borate diester bonding forms RGII dimer. Furthermore, transesterification could occur between pectins and other cell wall polysaccharides.

**Aromatic Substances**

The cell walls of grasses and commelinoid monocots contain considerable amounts of aromatic substances, mostly hydroxycinnamic acids, such as \(p\)-coumarate
and ferulate (Carpita, 1996). Ferulate is involved in the formation of cell wall polysaccharide cross-links (Grabber et al., 2000; Hatfield et al., 1999b). Ferulate can be esterified to the O-5 of arabinosyl units of arabinoxylans, and cell wall polysaccharides are cross-linked by radical coupling between esterified ferulates forming 8-O-4’, 8-5’, 8-8’, or 5-5’ linked diferulates. Extensive lignin-ferulate-polysaccharide complexes are also formed by copolymerization of ferulate monomers or diferulates with monolignols (Grabber et al., 2002). Furthermore, during the early stages of lignification, ferulates may play a role in providing a nucleation site for the attachment of monolignols (Ralph et al., 1995).

\( p \)-Coumarate is mainly esterified with lignin, not with cell wall polysaccharides, and it is esterified to the \( \gamma \)-position of sinapyl alcohol forming sinapyl \( p \)-coumarate conjugate (Ralph et al., 1994). The incorporation of sinapyl alcohol into the growing lignin polymer is thought to be facilitated by \( p \)-coumarate during the end-wise polymerization of lignin (Hatfield et al., 2008). Radical coupling of sinapyl alcohol alone is relatively slow, whereas radical formation is rapid with the addition of sinapyl \( p \)-coumarate conjugate.

**Structural Proteins**

There are four major classes of structural proteins: hydroxy-proline-rich glycoproteins (HRGPs, commonly extensins), proline-rich-proteins (PRPs), glycine-rich proteins (GRPs), and arabinogalactan proteins (AGPs) (Carpita and McCann, 2000). In dicots and non-commelinoid monocots, the cell wall carbohydrates are cross-linked with structural proteins. In contrast, grasses and commelinoid monocots have smaller amounts of structural proteins, but rather, aromatic substances, mostly hydroxycinnamates have functions similar to structural proteins (Carpita, 1996).
Lignin

Lignin is a major cell wall component, mainly deposited in the secondary cell wall of vascular plants. Lignin plays important roles in plant growth and development by providing mechanical strength, defense, and a path for water transportation (Boerjan et al., 2003). Lignin is a heterogeneous aromatic polymer consisting primarily of three monolignols, \( p \)-coumaryl (4-hydroxycinnamyl) alcohol, coniferyl (3-methoxy 4-hydroxycinnamyl) alcohol, and sinapyl (3,5-dimethoxy 4-hydroxycinnamyl) alcohol. After incorporation of these monolignols into a lignin polymer, they are referred to as \( p \)-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, respectively (Bonawitz and Chapple, 2010).

**Monolignol Biosynthesis**

Monolignol biosynthesis has diverged from the general phenylpropanoid pathway involved in the biosynthesis of many metabolites such as (iso)flavonoids, anthocyanins, stilbenes, tannins, phenylpropens, and coumarins (Vogt, 2010). The biosynthesis of monolignols commences with the deamination of phenylalanine, and follows a series of enzymatic reactions on the monolignol precursors: hydroxylation at the C3, C4, or C5 positions of the aromatic ring; \( O \)-methylation of the hydroxyl group at the C3 or C5 position of aromatic ring; two successive reductions at the side chain of the precursors from carboxylic acid to aldehyde and finally to alcohol (Boerjan et al., 2003; Bonawitz and Chapple, 2010).

As shown in Figure 1-1, and reviewed in Bonawitz and Chapple (2010) and Boerjan et al. (2003), monolignol biosynthesis shares three initial enzymatic reactions with the general phenylpropanoid pathway catalyzed by phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), and 4-coumarate: CoA ligase (4CL). The
Deamination of phenylalanine to cinnamic acid catalyzed by PAL is the first step, followed by the hydroxylation at the C4 position of cinnamic acid which is catalyzed by C4H to form p-coumaric acid. In grass species, tyrosine can be utilized by tyrosine ammonia-lyase (TAL) to yield p-coumaric acid (Rosler et al., 1997). p-Coumaric acid is further converted to p-coumaroyl CoA by 4CL, which is a branch-point metabolite between lignin and other phenylpropanoid metabolites (Vogt, 2010).

In the traditional view of the monolignol biosynthesis pathway, hydroxylation and O-methylation in monolignol biosynthesis are thought to occur at the level of free hydroxycinnamic acids. However, *in vivo* biochemical studies and transgenic research demonstrate that this pathway is highly unlikely (Dixon and Srinivas Reddy, 2003). Instead, hydroxylation at the C3 position of monolignol precursors occurs at shikimate or quinate ester of p-coumaroyl CoA. Firstly, hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyl transferase (HCT) converts p-coumaroyl CoA to p-coumaroyl shikimate or quinate, while p-coumaroyl CoA serves as a direct precursor for p-coumaryl alcohol and other phenylpropanoids biosynthesis. Hydroxylation at the C3 position of p-coumaroyl shikimate or quinate is then catalyzed by coumarate 3-hydroxylase (C3'H) to form caffeoyl shikimate or quinate (Schoch et al., 2001). Although C3'H is active on both shikimate and quinate ester, the shikimate derivative is considered to be preferred substrate for C3'H (Hoffmann et al., 2003). HCT is a reversible enzyme that converts caffeoyl shikimate to caffeoyl CoA. Subsequently, O-methylation at the C3 position of caffeoyl CoA is catalyzed by caffeoyl CoA O-methyltransferase (CCoAOMT) to produce feruloyl CoA.
The first reduction occurs at the level of hydroxycinnamoyl CoA thioester. Cinnamoyl CoA reductase (CCR) catalyzes the reduction of p-coumaroyl CoA to p-coumaraldehyde. However, considering the majority of G and S units in lignin, CCR is thought to preferentially convert feruloyl CoA to coniferaldehyde (Bonawitz and Chapple, 2010). The reduction of p-coumaraldehyde and coniferaldehyde to the corresponding hydroxycinnamyl alcohols, p-coumaryl alcohol and coniferyl alcohol, respectively, is catalyzed by cinnamyl alcohol dehydrogenase (CAD).

During the biosynthesis of sinapyl alcohol, hydroxylation at the C5 position of coniferaldehyde is carried out by ferulate 5-hydroxylase (F5H) yielding 5-hydroxyconiferaldehyde. F5H also can convert coniferyl alcohol to 5-hydroxyconiferyl alcohol (Humphreys et al., 1999; Osakabe et al., 1999). Subsequently, caffeic acid O-methyltransferase (COMT) catalyzes O-methylation at the C5 position of 5-hydroxyconiferaldehyde and 5-hydroxyconiferyl alcohol, yielding sinapaldehyde and sinapyl alcohol, respectively. Finally, enzymatic reduction occurs from sinapaldehyde to sinapyl alcohol by CAD.

**Lignin Polymerization and Deposition in the Secondary Cell Walls**

Following the monolignol biosynthesis, monolignols are exported from cytoplasm to the apoplast, although the mechanism of export and transport has not been fully elucidated (Bonawitz and Chapple, 2010). Several transportation models have been proposed including 4-O-glucosylated monolignol translocation, Golgi-derived vehicles, transporter mediated, and free diffusion models (Vanholme et al., 2010).

The lignin polymer is formed through radical coupling of activated monolignols that are oxidized by wall-bound peroxidases and/or laccases (Ralph et al., 2004). It has been challenging to elucidate which peroxidases and/or laccases catalyze enzymatic
oxidation because these enzymes belong to a large multigene family with overlapping activities (Vanholme et al., 2010). After the enzymatic oxidation, radical electrons on the oxygen atom at the C4 position of activated monolignols can be delocalized reversely to the C1, C3, C5, and side chain β carbon (Freudenberg, 1968). Monolignol radicals with a delocalized electron at the β carbon of side-chain are the most reactive with the highest electron density at this position (Vermerris, 2008). As reviewed in Ralph et al. (2004), during the end-wise polymerization, radical coupling occurs between the β carbon of the incoming monolignol and the 4-O- or C5 position of monolignol end unit, resulting in β-O-4 or β-5 linkages. β-1 linkages seem to be only formed between a monolignol radical and β-O-4 linked end unit. β-β linkage is only possible by dimerization of monolignols. 4-O-5 and 5-5 linkages are formed by cross-coupling of oligonols, and the frequency of these linkages is relatively minor. 4-O-5 and 5-5 linkages generate branched lignin structure, and since C5 position is not available for cross-coupling in the oligonols ended with S unit, S rich lignin could be more linear than G rich lignin (Bonawitz and Chapple, 2010). β-O-4 ether linkage is relatively liable to chemical degradation compared to biphenyl and carbon-carbon linkages (Boerjan et al., 2003). Physicochemical characteristics of lignin are primarily determined by monolignol distribution incorporated into the polymer and following inter-unit linkage patterns.

The mechanism of lignin polymerization is not fully understood. The traditional and most accepted model is that lignin polymerization is processed through random coupling of monolignol radicals. The relative abundance of the different monolignols and inter-unit linkages in lignin polymer are determined by the spatiotemporal regulation of monolignol biosynthetic genes and the availability of monolignol radicals with variable
oxidized sites at the site of polymerization (Boerjan et al., 2003). On the other hand, a directed lignin polymerization model postulates that a lignin polymer is built through selective and regiospecific coupling under rigid enzymatic reactions mediated by dirigent proteins (Davin and Lewis, 2000). However, this biochemically controlled model has limitations on explaining the racemic nature of lignin polymer, random sequence of units in lignin polymer, and the flexibility of lignifications which is the incorporation of unusual monolignols into lignin in transgenic and mutant plants (Boerjan et al., 2003; Ralph et al., 2004). Furthermore, genetic evidence is not enough to support dirigent proteins directing polymerization even in most of the model plants (Bonawitz and Chapple, 2010; Vermerris, 2008).

Lignin is deposited mainly in the secondary cell walls of specialized cells such as vessels and fibers (Donaldson, 2001). After cellulose and hemicelluloses are deposited in the secondary cell wall, lignification takes place in that region following the orientation of the cellulose microfibrils, and lignin impregnates the cell wall matrix (Boerjan et al., 2003; Bonawitz and Chapple, 2010). Lignin composition and structure are variable among and within species depending on developmental stages and cell types (Boerjan et al., 2003). The lignin of angiosperms mainly consists of G and S units with trace amounts of H units in dicots, and relatively more H units in monocots. The lignin of gymnosperms is mainly composed of G units with a lack of S units, and there is an intermediate amount of H units between dicots and monocots. Within a species, G units predominate in the earlier stage of development, and S units tend to be deposited later. The secondary cell walls of vessels are enriched in G units, while structural fibers contain more S units.
Forward and Reverse Genetics in Lignin Biosynthesis and Effects of Manipulating Lignin Biosynthetic Genes on the Formation of Lignin and Plant Development/Growth

Phenylalanine ammonia-lyase (PAL)

PAL is encoded by a multigene family. There is functional redundancy among PAL family members, although each member has a distinctive role in the general phenylpropanoid pathway (Huang et al., 2010; Rohde et al., 2004). In agreement with the role of PAL in the general phenylpropanoid pathway, T-DNA insertional mutations in *PALs: pal1, pal2, pal1/pal2 double, or pal1/pal2/pal3/pal4 quadruple mutation*, in *Arabidopsis* affects the biosynthesis of lignin and other metabolites including flavonol glucosides, tannin, and anthocyanin pigments, as well as SA (Huang et al., 2010; Rohde et al., 2004). Furthermore, transcript profiling in the mutants indicates that knock-down of PAL not only changes expression patterns of other lignin biosynthetic genes but also affects the expression of genes related to amino acid and carbohydrate metabolism (Rohde et al., 2004).

Single mutant *pal1* and *pal2* have no obvious phenotypic differences compared to wild-type. However, a *pal1/pal2* double mutation affects fertility, depending on the genetic background of *Arabidopsis* (ecotype C24 vs. Col-0), and reduces tolerance to UV-B irradiation (Huang et al., 2010; Rohde et al., 2004). The quadruple mutant shows dwarfism, male sterility, and reduced levels of SA and tolerance to pathogens (Huang et al., 2010). Down-regulation of PAL induces pleiotropic effects, and the accumulation of other metabolites important for defense mechanisms. For example, PAL suppressed tobacco is susceptible to fungal, bacterial, and viral pathogens (Dixon and Srinivasa Reddy, 2003).
Cinnamate 4-hydroxylase (C4H)

*Arabidopsis reduced epidermal fluorescence 3 (ref3)* harboring a missense mutation in *C4H* shows reductions in total lignin content and in several phenylpropanoid derivatives (Ruegger and Chapple, 2001; Schilmiller et al., 2009). Reduced C4H activity in the ref3 mutant primarily affects G unit biosynthesis, with no changes in S units. However, C4H down-regulated tobacco and alfalfa show reduced S units (Sewalt et al., 1997; Srinivasa Reddy et al., 2005).

A series of ref3 mutants with different levels of reduced C4H activity indicate that severe loss of C4H activity results in dwarfism, collapsed xylem vessels, and male sterility (Schilmiller et al., 2009).

4-coumarate:CoA ligase (4CL)

The *brown midrib (bmr2)* mutant of sorghum carries a missense mutation of the *4CL1* gene. Bmr2 with reduced 4CL activity shows 20% reductions in total lignin content and in both G and S units (Saballos et al., 2012). The down-regulation of 4CL in switchgrass and poplar results in a slight increase in H units and preferential reductions in G over S units with reduced total lignin content (Voelker et al., 2010; Xu et al., 2011). Several 4CL isoforms encoded by a multigene family are found *in planta*, and 4CL family genes are generally clustered into two groups in both monocots and dicots. Class I and II 4CLs are thought to be involved in lignin and flavonoid biosynthesis, respectively, with different substrate preferences, relative transcript abundance, and spatiotemporal expression patterns (Ehlting et al., 1999; Gui et al., 2011; Saballos et al., 2012). At4CL1, Pv4CL1 and Os4CL3 are identified as key 4CL enzymes involved in lignin biosynthesis in *Arabidopsis*, switchgrass and rice, respectively (Ehlting et al., 1999; Gui et al., 2011; Xu et al., 2011).
Brown coloration in leaf midrib, basal stem, and/or roots is observed in 4CL suppressed mutant and transgenic plants (Gui et al., 2011; Saballos et al., 2012; Voelker et al., 2010). The bmr2 mutation has no adverse effect on plant phenotype, and down regulation of Pv4CL1 in switchgrass does not affect biomass accumulation (Saballos et al., 2012; Xu et al., 2011). However, delayed plant growth and abnormal anther development are observed in transgenic rice and reduced biomass production is observed in transgenic poplar (Gui et al., 2011; Voelker et al., 2010).

**Hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyl transferase (HCT)**

Down-regulation of HCT in alfalfa and Arabidopsis reduces total lignin content with a strong reduction in G and S units and an increase in H units (Hoffmann et al., 2004; Shadle et al., 2007). Therefore, the enzymatic reaction of HCT is a crucial initial step to drive the biosynthesis of coniferyl and sinapyl alcohols.

HCT-suppressed alfalfa and Arabidopsis exhibit dwarf phenotypes with reduced biomass (Gallego-Giraldo et al., 2011; Shadle et al., 2007). Impaired plant growth in HCT down-regulated Arabidopsis appears to be associated with an increased level of SA derived from the shikimate pathway (Gallego-Giraldo et al., 2011). Involvement of shikimate during monolignol biosynthesis is required not only for supplying substrate for the 3-hydroxylation step, but also for regulating the flow of monolignol precursors (Bonawitz and Chapple, 2010).

**p-Coumarate 3’-hydroxylase (C3’H)**

The Arabidopsis ref8 mutant, carrying a missense mutation in C3’H, has increased levels of p-coumarate esters and strongly reduced total lignin content (Franke et al., 2002a; Franke et al., 2002b). Hydroxylation by C3’H is considered a committed
control step along with HCT for the biosynthesis of coniferyl and sinapyl alcohol. The ref8 mutant has small amounts of G and S lignin, and the lignin almost entirely consists of H lignin (Franke et al., 2002a). C3′H down-regulated alfalfa has strongly reduced G and S units and increased H units similar to the ref8 mutant (Ralph et al., 2006). Both the Arabidopsis mutant and the transgenic alfalfa show dwarf phenotypes.

**Caffeoyl CoA O-methyltransferase (CCoAOMT)**

Down-regulation of CCoAOMT results in reduced total lignin content with a preferential reduction in G units, without a decrease in S units, in transgenic tobacco and alfalfa plants (Guo et al., 2001; Zhong et al., 1998). These results indicate that CCoAOMT is specifically involved in the 3-Ọ-methylation of caffeoyl CoA for the biosynthesis of coniferyl alcohol. CCoAOMT-suppressed alfalfa accumulates soluble caffeic acid β-D-glucoside, possibly indicating that excess caffeoyl CoA in the metabolite pool of monolignol biosynthesis is converted into caffeic acid and then modified with glucosylation (Guo et al., 2001). Similar to transgenic alfalfa, soluble phenolic acids (vanillic, caffeic, and sinapic acids) glucosides are accumulate in CCoAOMT down-regulated poplar (Meyermans et al., 2000).

Transgenic poplar displays pink-red color on its debarked stem, and its overall morphology is similar to wild-type (Meyermans et al., 2000). There are contrasting results regarding the effects of CCoAOMT suppression on transgenic tobacco plants. Significantly decreased stem length is shown in one study (Pinçon et al., 2001b), whereas normal plant growth is shown in another (Zhong et al., 1998).

**Cinnamoyl CoA reductase (CCR)**

Down-regulation of CCR in Arabidopsis, tobacco, and poplar results in ~50% reduction of total lignin content, while the effects of CCR suppression on lignin
composition are variable (Goujon et al., 2003; Leplé et al., 2007; Piquemal et al., 1998; Zhou et al., 2010b). Lignin compositional variation in transgenic plants might depend on the species, developmental stages, culture conditions, and the suppression of different CCR isoforms. Zhou et al. (2010b) suggested that two CCR isoforms in Medicago truncatula, CCR1 and CCR2, provide non-redundant, independent routes for the biosynthesis of coniferyl and sinapyl alcohol. CCR1 preferentially catalyzes reduction of feruloyl CoA, which is the product of the CCoAOMT reaction, to coniferaldehyde. CCR2 drives an alternative route via reduction of caffeoyl CoA to caffeyl aldehyde, and its subsequent conversion to coniferaldehyde by COMT (Zhou et al., 2010b). It is unclear whether the metabolite pool of coniferaldehyde resulting from CCoAOMT/CCR1 and CCR2/COMT routes is identical or separate. Based on the observation that CCoAOMT suppression results in reduced G units without a change in S units, the CCR2/COMT route is thought to be involved in S unit formation, at least in the absence of CCoAOMT (Guo et al., 2001; Zhou et al., 2010b). In addition, the Arabidopsis CCR1 knock-out mutant deposits increased levels of feruloyl malate in the stem, and CCR down-regulated poplar accumulates high levels of ferulic acid incorporated into lignin (Leplé et al., 2007; Mir Derikvand et al., 2008). These results show a possible redirection of feruloyl CoA toward ferulic acid upon blocking CCR, and incorporation of intermediates into lignin polymer. The CCR-suppressed plants exhibit orange-brown coloration around the xylem, decreased plant height, and collapsed vessels with a less degree of secondary cell wall formation (Goujon et al., 2003; Leplé et al., 2007; Piquemal et al., 1998).
**Ferulate 5-hydroxylase (F5H)**

A F5H-deficient T-DNA tagging mutant in *Arabidopsis* (*fah1*) shows dramatic changes in lignin composition and structure. Consistent with the role of F5H in sinapyl alcohol biosynthesis, the lignin of *fah1* contains only a trace amount of S units with increased occurrences of β-5 and 5-5 linkages (Humphreys et al., 1999; Marita et al., 1999). In contrast, over-expression of F5H in *Arabidopsis* and poplar driven by the *C4H* promoter results in lignin comprised of over 90% S units (Marita et al., 1999; Stewart et al., 2009). The amount of β-O-4 linkages are similar in the F5H over-expressed poplar compared to control plants, while β-β linkages are increased (Stewart et al., 2009). In addition, the lignin with the elevated level of S unit contains no β-5 linkages and less 5-5 linkage structures. Despite dramatic alterations in relative abundance of G and S units, the F5H over-expressed transgenic poplar exhibits no differences in vessel development and cell wall thickness (Huntley et al., 2003).

**Caffeic acid O-methyltransferase (COMT)**

COMT deficiency in mutants (maize *bm3* and sorghum *bmr12*) and down-regulation of COMT in a wide variety of species, including alfalfa, tobacco, tall fescue, perennial ryegrass, switchgrass, and poplar consistently result in a strong reduction in S units, indicating a primary role of COMT in sinapyl alcohol biosynthesis (Chen et al., 2004; Fu et al., 2011a; Guo et al., 2001; Marita et al., 2003; Palmer et al., 2008; Pinçon et al., 2001a; Tu et al., 2010; Van Doorsselaere et al., 1995). Furthermore, COMT deficiency results in the accumulation of 5-hydroxyconiferaldehyde which is generally the most preferential substrate for COMT. 5-hydroxyconiferaldehyde is then reduced to 5-hydroxyconiferyl alcohol by CAD, and this atypical monolignol is incorporated into lignin.
forming a novel benzodioxane structure referred to as the 5-hydroxyguaiacyl unit (Palmer et al., 2008; Ralph et al., 2000; Van Doorsselaere et al., 1995).

COMT is encoded by a multigene family belonging to the type-1 plant O-methyltransferases (OMT), and COMT generally has broad substrate permissiveness with different substrate preferences (Louie et al., 2010; Parvathi et al., 2001; Zubieta et al., 2002). Although COMT has the highest catalytic affinity toward 5-hydroxyconiferaldehyde, 3-0 or 5-0-methylation could also occur at caffeoyl aldehyde, caffeoyl alcohol, and/or 5-hydroxyconiferyl alcohol. Incomplete depletion of S units in COMT knock-out mutants may further support the idea that sinapyl alcohol can, at least in part, originate from alternative routes driven by other OMT family members (Barrière et al., 2004; Dixon et al., 2001; Parvathi et al., 2001).

COMT defective mutants and transgenic plants show brown coloration in vascular tissues. COMT suppression generally has negligible effects on the plant growth of various transgenic plants, including alfalfa, maize, switchgrass, perennial ryegrass, and poplar (Chen and Dixon, 2007; Fu et al., 2011a; Pilate et al., 2002; Piquemal et al., 2002; Tu et al., 2010).

**Cinnamyl alcohol dehydrogenase (CAD)**

Deficiency of CAD in mutants such as the maize *bm1* and sorghum *bmr6*, and transgenic tobacco and poplar characteristically result in the incorporation of coniferaldehyde and/or sinapaldehyde into lignin polymers, which are typical substrates for CAD (Chen et al., 2012; Lapierre et al., 2004; Marita et al., 2003; Ralph et al., 1998; Saballos et al., 2008). Consistent with the lack of S units in gymnosperms, only one *CAD* gene or a relatively small gene family is present in gymnosperms (MacKay et al., 1997). On the other hand, CAD is encoded by a multigene family in angiosperms. In
aspen, one of the CAD family members is reported to have a high substrate affinity toward sinapaldehyde and is designated as sinapyl alcohol dehydrogenase (SAD) (Li et al., 2001). While the presence of SAD in herbaceous angiosperms has yet to be investigated, the enzymatic reduction of each hydroxycinnamaldehyde into corresponding alcohol seems to be governed rather by the combined action of each CAD family member in different developmental stages and tissues than by a highly specified CAD (Sibout et al., 2005; Vermerris, 2008).

CAD deficient mutants and down-regulated transgenic plants show brown coloration in lignified tissues due to the accumulation of conjugated aldehydes in these tissues (Baucher et al., 1996; Sattler et al., 2010). CAD down-regulated poplar trees, which have 15-47% of wild-type CAD activity show normal growth performance under field conditions, while severe levels of CAD suppression result in impaired growth under greenhouse conditions (Pilate et al., 2002).

**Lignocellulosic Biofuel**

Given the environmental and socio-economic problems caused by the heavy dependence on fossil energy, it is expected that renewable biofuels will replace a certain amount of fossil fuel, particularly in the transportation sector. According to the U.S. Renewable Fuel Standard (RFS) with the enactment of the Energy Policy Act of 2007, the U.S. has the goal of producing 136.3 billion liters (36 billion U.S. gallons) of biofuel by 2022 (Schnepf and Yacobucci, 2012). This amount of biofuel has the potential to replace about 27% of the gasoline by quantity in the current U.S. market (134 billion U.S. gallons consumed in 2011, U.S. Energy Information Administration, http://www.eia.gov/). Current biofuels, especially bioethanol production, mostly relies on starch and sugar feedstocks. However, there have been concerns about the negative
impacts on the food and feed supply and relatively poor environmental sustainability, particularly regarding starch based ethanol from corn. In this respect, the production of corn ethanol cannot exceed more than 56.8 billion liters (15 billion U.S. gallons) from 2015 under the RFS (Schnepf and Yacobucci, 2012).

Lignocellulosic biomass has been considered an abundant, inexpensive, and sustainable energy resource that can provide bioethanol on a sufficient scale to replace a significant amount of gasoline. The price of biomass is equivalent to 13 U.S. dollars per barrel of crude oil on the basis of energy value (Yang and Wyman, 2008). In the USA, it is projected that 1.3 billion dry Mg of lignocellulosic biomass could be available annually by 2030, and it has the potential to replace 87 billion U.S. gallons of gasoline (Carroll and Somerville, 2009), which is nearly 65% of the gasoline currently consumed in the USA (134 billion U.S. gallons consumed in 2011, U.S. Energy Information Administration, http://www.eia.gov/). Furthermore, the energy balance of lignocellulosic ethanol has higher net benefits in terms of reductions in greenhouse gas emissions and energy input, compared with fossil fuels, and also conventional ethanol from corn, wheat, and sugar beet (Goldemberg, 2007; Sims et al., 2010). Considering its abundance and sustainability, at least 60.6 billion liters (16 billion U.S. gallons) of ethanol are expected to be produced from lignocellulosic biomass by 2022 under the RFS (Schnepf & Yacobucci, 2012).

Despite the exceptionally beneficial characteristics of lignocellulosic ethanol as a renewable alternative, full commercialization has been delayed due to economic and technical limitations. The production costs of lignocellulosic ethanol, which range from 0.6-1.3 U.S. dollars per liter, are higher than gasoline and other conventional bioethanol
(Goldemberg, 2007; Sims et al., 2010). In this respect, there have been tremendous efforts in developing efficient lignocellulosic feedstock plants and bioconversion procedures. 

**The Production Process of Lignocellulosic Ethanol**

The process of ethanol production from lignocellulosic biomass includes the following steps: pretreatments of the feedstock, hydrolysis of cellulose and hemicelluloses into simple sugars, fermentation of sugars, and recovery and purification of ethanol (Lu and Mosier, 2008).

Pretreatment is a process to disrupt the naturally resistant cell wall matrix, and thereby increase the accessibility of hydrolyzing enzymes to cellulose and hemicelluloses (Mosier et al., 2005). Although pretreatment is the essential step for efficient bioconversion of lignocellulosic biomass into biofuel and other bio-based products, it is considered the most expensive cost element because of intense chemical and energy requirements (Yang and Wyman, 2008). Over the years, various pretreatment technologies have been developed and evaluated targeting high recovery yields of fermentable sugars, limiting the release of undesirable compounds that inhibit fermentation, and minimizing the chemical and energy demand (Mosier et al., 2005).

The effects of pretreatment on cell wall components vary among pretreatment methods. In general, pretreatments under alkaline conditions, such as ammonia fiber/freeze explosion (AFEX), ammonia recycle percolation (ARP), and lime pretreatment have major effects on removing or depolymerizing lignin with minor effects on solubilizing hemicelluloses. On the other hand, acidic or neutral pretreatments, such as diluted acid, steam explosion, and liquid hot water pretreatments effectively remove hemicelluloses rather than lignin (Alvira et al., 2010). There are advantages and
disadvantages to all of the current pretreatment technologies that significantly increase the accessibility of enzymes and saccharification efficiencies. For example, acid based pretreatments are applicable to a wide variety of lignocellulosic materials compared with AFEX, ARP, or lime pretreatment (Mosier et al., 2005). There is a greater chance of generating toxic compounds that inhibit fermentation, especially in steam explosion and acid pretreatment, whereas, liquid hot water, alkaline, and AFEX pretreatments minimize the release of fermentation inhibitors (Alvira et al., 2010).

After pretreatment, the enzyme hydrolysis step aims to more efficiently depolymerize polysaccharides in lignocellulosic biomass into fermentable sugars, which then can be easily converted to ethanol by fermentation microorganisms. Hydrolysis of cellulose into glucose requires the synergistic action of three different types of cellulolytic enzymes, endo-β-1,4 glucanase (endoglucanase, EG), exo-β-1,4 glucanase (cellobiohydrolase, CBH), and β-glucosidase (cellobiase) (Mansfield et al., 1999). EGs randomly cleave β-1,4 glycosidic bonds of cellulose. CBHs remove cellobiose residues from the cellulose chain hydrolyzing not only non-reducing ends, but also reducing ends of the chain. β-glucosidases finally hydrolyze cellobiose residues into glucose.

The enzyme system for hydrolyzing hemicelluloses is more complex due to the heterogeneity of hemicelluloses (Lu and Mosier, 2008; Shallom and Shoham, 2003). For hydrolyzing GAXs, which are the major hemicellulosic component in grasses, endo-β-1,4 xylanase and exo-β-1,4 xylosidase are required to liberate xylose from the xylan backbone. The side chains on GAXs can be removed by α-L-arabinofuranosidase and α-D-glucuronidases. Hemicellulolytic esterases including acetyl xylan esterase and feruloyl esterase improve xylose hydrolysis efficiency.
In the conventional process of lignocellulosic ethanol production, enzyme hydrolysis and ethanol fermentation are conducted as separate steps or as one single step, referred to as separate hydrolysis and fermentation (SHF) or simultaneous saccharification and fermentation (SSF), respectively (Lu and Mosier, 2008). If pentose sugars are fermented with hexose sugars, the SSF configuration is often classified as simultaneous saccharification and co-fermentation (SSCF) (Margeot et al., 2009). The integrated process configuration of SSF and SSCF reduce the capital cost of lignocellulosic ethanol production compared to SHF. A further integrated process, consolidated bio-processing (CBP) has been suggested as a desired process configuration for more economical lignocellulosic ethanol production (Galbe et al., 2007). In CBP, the simultaneous hydrolysis and fermentation are mediated by a single organism or a microbial consortium, while cellulolytic enzymes are produced in a separate system and added to saccharification/fermentation step in SSF and SSCF (Galbe et al., 2007). However, the integrated process is often operated under compromised operational conditions due to the different optimal conditions for enzymatic hydrolysis and fermentation steps. In addition, the microorganism mediating both cellulose hydrolysis and fermentation simultaneously in a satisfactory manner has not yet been developed (Margeot et al., 2009).

The bioconversion of fermentable sugars remains suboptimal, especially for pentose sugars in which fermentation is less efficient than hexose sugar fermentation. Pentose sugars comprise a significant portion in plant cell walls, but the native strain of *Saccharomyces cerevisiae*, a common choice of microorganism for ethanol production, is unable to convert pentose sugars into ethanol due to the lack of pentose sugar
metabolism (Alper and Stephanopoulos, 2009). Therefore, efficient pentose sugar fermentation would improve the cost-effectiveness of lignocellulosic ethanol production. Advances in metabolic engineering have significantly increased the efficiency of pentose sugar fermentation by introducing heterologous pentose sugar metabolic pathways in S. cerevisiae and by improving ethanol productivity and ethanol tolerance of pentose utilizing organisms (Van Vleet and Jeffries, 2009).

The Intrinsic Recalcitrance of Lignocellulosic Biomass for Ethanol Production

The complexity and heterogeneity of the plant cell wall provides physical and chemical resistance to hydrolysis into fermentable simple sugars (Himmel et al., 2007). The chain length of cellulose (the degree of polymerization) and the ratio of crystalline to amorphous cellulose (crystallinity) are thought to affect the rate of enzymatic hydrolysis (Mansfield et al., 1999). However, these characteristics alone do not directly explain the biomass recalcitrance to enzymatic hydrolysis (Zhao et al., 2012). Crystalline cellulose itself can be hydrolyzed by the synergistic action of the fungal cellulase complexes (Mansfield et al., 1999). The limited enzymatic hydrolysis of cellulose is associated with other factors, such as the surface area availability of the substrate for cellulolytic enzymes, pore size/volume, hemicellulose cross-linking, and lignin impregnation (Alvira et al., 2010; Zhao et al., 2012).

The degree of substitution in hemicellulose determines pore size and volume in the cell wall matrix. In grass species, the less branched GAXs result in a higher degree of self cross-linking and with cellulose microfibrils, thus the cell walls containing less branched GAXs become more rigid and dense (Vermerris, 2008). In addition, the level of feruloylation of arabinoxylans affect enzymatic hydrolysis because ferulate cross-links with arabinoyxylan and lignin interconnects the cell wall matrix (Grabber et al., 1998).
Association of hemicelluloses in biomass recalcitrance is considered less important than that of lignin because hemicelluloses are relatively liable to pretreatment and enzymatic hydrolysis (Zhao et al., 2012).

The presence of lignin in the plant cell wall is a major obstacle to enzymatic hydrolysis (Jørgensen et al., 2007; Mansfield et al., 1999; Zhao et al., 2012). Lignin primarily limits the accessibility of cellulose to cellulolytic enzymes by shielding the cellulose chain. Lignin also corresponds to lower porosity of the cell wall, thus the surface area of substrate for the enzyme attack is highly limited in the presence of lignin. Furthermore, up to 60-70% of cellulolytic enzymes added will irreversibly adsorb to lignin, thereby overall enzyme activity is reduced (Chernoglazov et al., 1988; Jørgensen et al., 2007).

**Improved Quality of Lignocellulosic Feedstock through Lignin Modification**

The presence of lignin in the lignocellulosic biomass has long been recognized as a major obstacle in the manufacturing of paper, digestibility of forage, and biochemical production, including biofuels. For the efficient bioconversion of lignocellulosic biomass into ethanol, pretreatment is an essential operating process, disrupting the cell wall matrix, removing and relocating lignin (Mosier et al., 2005). Although pretreatment is beneficial to increase ethanol production, it is an expensive and energy intensive process, accounting for about 20% of the total production costs (Yang and Wyman, 2008). Furthermore, depending on the pretreatment method, pretreatment could negatively impact subsequent steps causing degradation of sugars and the generation of undesirable compounds, such as acetic acid, furan, and phenolic derivatives, which are toxic to fermenting microorganisms (Alvira et al., 2010; Jørgensen et al., 2007). Therefore, along with the efforts in developing advanced
pretreatment methods, there has been tremendous interest in modifying lignin content and/or composition of lignocellulosic feedstock to improve the efficiency of ethanol production.

Chen and Dixon (2007) demonstrated that reducing lignin in lignocellulosic biomass significantly improved fermentable sugar yields up to 2.7-fold without pretreatment and up to 1.9-fold with dilute acid pretreatment. In that study, six different lignin biosynthetic genes (C4H, HCT, C3′H, CCaOMT, F5H, and COMT) were independently down-regulated in transgenic alfalfa. Each transgenic line displayed different levels of lignin reduction and values of S to G unit molar ratio, depending on the targeted lignin biosynthetic genes. In the set of transgenic lines, it was clearly shown that total lignin content was negatively correlated with fermentable sugar release after enzymatic hydrolysis with or without pretreatment (adjusted \( R^2 \)=0.91 or 0.79 with pretreatment or without pretreatment, respectively). Although, the chemical and structural characteristics of the lignin polymer are mainly determined by the relative abundance of each monolignol, it was inconclusive whether there was a relationship between S/G ratio and saccharification efficiency.

Improvement of saccharification efficiencies and ethanol yields have been consistently demonstrated in lignin reduced transgenic plants and mutants, demonstrating the negative impact of lignin on the bioconversion of lignocellulosic materials (Table 1-1). In the sorghum brown mid-rib mutants carrying mutations in 4CL, COMT, or CAD genes (Dien et al., 2009; Saballos et al., 2008), fermentable sugar yields were improved by 7-27% and ethanol yields were enhanced 22%, compared with the wild-type counterpart. Furthermore, a double mutation in lignin biosynthetic genes
had an additive effect on sugar and ethanol yields (Dien et al., 2009). Lignin modified transgenic switchgrasses were generated by suppressing CAD, COMT, or 4CL (Fu et al., 2011a; Fu et al., 2011b; Saathoff et al., 2011; Xu et al., 2011). Particularly, COMT-suppressed transgenic switchgrass exhibited a 22 and 38% increase in saccharification efficiency and ethanol yields, respectively (Fu et al., 2011a). Moreover, it has been demonstrated that there is great potential to reduce pretreatment severity, i.e. pretreatment reaction time, temperature and enzyme dosage, for producing a comparable amount of ethanol from lignin modified transgenic feedstocks (Fu et al., 2011a). Manipulating lignin biosynthetic genes or identifying lignin modified mutants could be a straightforward and viable strategy to increase the economic feasibility of lignocellulosic biofuel production with diminished biomass recalcitrance and pretreatment requirements.

**RNA Interference (RNAi)**

RNAi is one of the cellular regulation processes of gene expression mediated by small silencing RNA (sRNA). sRNA can be primarily categorized into two classes on the basis of biogenesis, namely microRNA (miRNA) and short interfering RNA (siRNA) (Carthew and Sontheimer, 2009). miRNA originated from an endogenous non-coding miRNA gene, and its expression is highly regulated in the genome. miRNA is processed from a miRNA gene transcript (pre-miRNA) which forms a stem-loop structured double stranded RNA (dsRNA) with imperfect complementarity. On the other hand, siRNA is generated from perfectly complementary double stranded RNA (dsRNA) precursors derived from diverse sources, such as transgenes, viral RNAs, transposons, and gene/pseudogene duplexes.
The mechanism of RNAi is reviewed in Carthew and Sontheimer (2009). In general, RNAi results in post transcriptional gene silencing (PTGS) and transcriptional gene silencing (TGS). The process begins with the formation of dsRNA of the transcripts derived from exogenous sources or endogenous sources. dsRNA can be formed by either sequence complementarity or by the action of an RNA-dependent RNA polymerase. dsRNA is then recognized and cleaved by Dicer into sRNAs. In plants, Dicers are dsRNA specific RNAase III family ribonucleases encoded by the small dicer-like (DCL) gene family (Hamilton et al., 2002; Xie et al., 2004). Each has distinctive and redundant roles in generating various types of sRNA with different lengths. In Arabidopsis, DCL1/DCL4, DCL2, and DLC3 are known to produce ~21, ~22, and 24~26 nt long sRNA, respectively. 21~22 nt long sRNAs are mostly involved in PTGS, while 24~26 nt long sRNAs mainly trigger TGS (Hamilton et al., 2002; Xie et al., 2004). Newly produced double stranded sRNAs are then loaded into one of the Argonaute (Ago) proteins in RNA induced silencing complex (RISC). During the assembly process, one of the strands of sRNA is removed or cleaved by Ago. The remaining guide strand of sRNA directs RISC and binds to the target RNA via Watson–Crick base pairing. Finally, PTGS is induced from the endonucleolytic cleavage of the target mRNA by the action of Ago, and then cellular exonucleases complete the degradation of the target mRNA. In few cases, during the cycle of PTGS mediated by siRNA or miRNA, cleaved target transcripts can be converted into dsRNAs by RNA dependent RNA polymerase and other cofactors. This secondary dsRNAs can be further processed into secondary sRNA by Dicer in plants, so that the silencing can be amplified and spread throughout the cell.
The molecular mechanism of TGS is not fully understood compared to PTGS (Carthew and Sontheimer, 2009). Specialized RNA induced transcriptional silencing complex (RITC) is assembled with sRNA which is homologous to the target chromosomal loci. sRNA within the complex appears to interact with RNA polymerase, transcribing the target region, and providing binding sites for the protein complex including DNA methyltransferase which mediates heterochromatin formation.

RNAi has been utilized as an excellent tool for the functional analysis of genes and in crop improvement. Sequence directed suppression of endogenous gene(s) is achieved by introducing the transgene which can be processed into siRNA in the host genome. RNAi has several advantages over other undirected methods such as physicochemical and insertional mutagenesis. These conventional mutagenesis approaches require generating a large number of mutant lines to identify the desirable mutation and the gene responsible for generating the phenotype (Krysan et al., 1999; Wesley et al., 2001). Moreover, particularly in a polyploid genome, these methods would be ineffective, resulting in null phenotypes due to genetic redundancy (Waterhouse and Hellwell, 2003). RNAi allows simultaneous suppression of homo(eo)logous gene copies, which may have functional redundancy and complementation (Lawrence and Pikaard, 2003; Travella et al., 2006). Silencing a specific gene or all members of a multigene family is also possible by targeting unique and conserved regions, respectively (Miki et al., 2005). RNAi is also favored to identify and characterize lethal alleles (Waterhouse and Hellwell, 2003), with different levels of target gene suppression due to the variation of silencing efficiency in different
transgenic events. Furthermore, RNAi induction can be moderated using spatiotemporal specific or inducible promoters (Ossowski et al., 2008).

Off-target silencing is considered a drawback to employing transgene induced RNAi. In general, contiguous sequence identity over at least 21 bp and/or overall homology over 88% between the transgene and endogenous sequences could induce RNAi (Travella et al., 2006; Xu et al., 2006). Furthermore, a cleavage site recognized by Dicer is unknown in dsRNAs derived from the transgene, thus the chance of off-target silencing can be increased by the generation of various species of siRNAs with diverse sequences processed from different regions of dsRNAs (Ossowski et al., 2008). It is challenging to select targeting sequences with precise specificity given the limited sequence information for most of the crop species. Instead, the transgene construct can be designed among closely related genes, with short target sequences similar to miRNA precursors, in order to reduce the number of siRNA species (Ossowski et al., 2008).

The efficiency of transgene induction mediated PTGS is affected by the transgene construct and the resulting dsRNA structure (Smith et al., 2000; Wesley et al., 2001). The inverted repeat construct with the splicing intron generally shows 90~100% induction of PTGS in transgenic events, while the occurrence of PTGS was less than 70% with a linker, spacer, or non-splicing intron. The classical sense or antisense silencing construct triggers PTGS in less than 15% of transformants.

The suppression level of the target gene differs among independent transgenic events. The effectiveness of target gene silencing is primarily affected by the productivity of siRNA from the transgene and the accessibility of siRNA to the target (Ossowski et al., 2008). Single or low copy number(s) of the transgene induce higher
levels of target gene suppression than multiple copies (Kerschen et al., 2004). Transgene integration site in the genome would be also important in the generation of sufficient siRNA, since transgene expression can be repressively influenced by flanking host DNA sequences and/or unfavorable chromosome regions (Matzke and Matzke, 1998).
Table 1-1. Improvement of fermentable sugar and/or ethanol yields in lignin reduced transgenic plants and mutants

<table>
<thead>
<tr>
<th>Plants</th>
<th>Genes affected</th>
<th>Methods</th>
<th>Lignin content a)</th>
<th>Fermentable sugar yields b)</th>
<th>Ethanol yields c)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alfalfa</td>
<td>C4H</td>
<td>Antisense</td>
<td>29%</td>
<td>41%*</td>
<td>n.d.</td>
<td>Chen and Dixon (2007)</td>
</tr>
<tr>
<td>Alfalfa</td>
<td>HCT</td>
<td>Antisense</td>
<td>52%</td>
<td>88%*</td>
<td>n.d.</td>
<td>Chen and Dixon (2007)</td>
</tr>
<tr>
<td>Alfalfa</td>
<td>C3'H</td>
<td>Antisense</td>
<td>36%</td>
<td>59%*</td>
<td>n.d.</td>
<td>Chen and Dixon (2007)</td>
</tr>
<tr>
<td>Alfalfa</td>
<td>F5H</td>
<td>Antisense</td>
<td>0%</td>
<td>0%</td>
<td>n.d.</td>
<td>Chen and Dixon (2007)</td>
</tr>
<tr>
<td>Alfalfa</td>
<td>CCR</td>
<td>Antisense</td>
<td>31%*</td>
<td>58%</td>
<td>n.d.</td>
<td>Jackson et al. (2008)</td>
</tr>
<tr>
<td>Alfalfa</td>
<td>CAD</td>
<td>Antisense</td>
<td>13%*</td>
<td>19%</td>
<td>n.d.</td>
<td>Jackson et al. (2008)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>4CL</td>
<td>Mutation 1)</td>
<td>16%</td>
<td>17%</td>
<td>n.d.</td>
<td>Saballos et al. (2008)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>unknown</td>
<td>Mutation 2)</td>
<td>17%</td>
<td>20%</td>
<td>n.d.</td>
<td>Saballos et al. (2008)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>CAD</td>
<td>Mutation 3)</td>
<td>22%</td>
<td>7%</td>
<td>n.d.</td>
<td>Saballos et al. (2008)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>COMT</td>
<td>Mutation 4)</td>
<td>13%</td>
<td>21%</td>
<td>n.d.</td>
<td>Saballos et al. (2008)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>CAD</td>
<td>Mutation 5)</td>
<td>13%</td>
<td>27%</td>
<td>22%</td>
<td>Dien et al. (2009)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>COMT</td>
<td>Mutation 6)</td>
<td>15%</td>
<td>23%</td>
<td>21%</td>
<td>Dien et al. (2009)</td>
</tr>
<tr>
<td>Sorghum</td>
<td>COMT+CAD</td>
<td>Mutation 7)</td>
<td>27%</td>
<td>34%</td>
<td>43%</td>
<td>Dien et al. (2009)</td>
</tr>
<tr>
<td>Switchgrass</td>
<td>CAD</td>
<td>RNAi</td>
<td>24%</td>
<td>44%</td>
<td>n.d.</td>
<td>Fu et al. (2011b)</td>
</tr>
<tr>
<td>Switchgrass</td>
<td>COMT</td>
<td>RNAi</td>
<td>13%</td>
<td>22%</td>
<td>38%</td>
<td>Fu et al. (2011a)</td>
</tr>
<tr>
<td>Switchgrass</td>
<td>4CL</td>
<td>RNAi</td>
<td>22%</td>
<td>57%</td>
<td>n.d.</td>
<td>Xu et al. (2012)</td>
</tr>
</tbody>
</table>

a) Values are the percentage decrease of total lignin content in transgenic or mutant plants, compared with corresponding control plants.
b) Values are the percentage increase of glucose yields per unit biomass or per unit cellulose in transgenic or mutant plants, compared with corresponding control plants. Only glucose yields from lignocellulosic materials pretreated with dilute sulfuric acid were used for this comparison.
c) Values are the percentage increase of ethanol yields per unit biomass or per unit cellulose following simultaneous saccharification and fermentation with S. cerevisiae D5A in transgenic or mutant plants, compared with corresponding control plants.

1, 2, 3, and 4) Sorghum brown mid-rib mutant bmr2, bmr3, bmr6, and bmr12, respectively. bmr2 was currently characterized harboring mutation in 4CL (Saballos et al., 2012). The gene responsible for bmr3 is unknown.

5, 6, and 7) Sorghum brown mid-rib mutant bmr6, bmr12, and bmr6 and bmr12 double mutant, respectively. These mutants were near isogenic lines with the genetic background of cultivar Atlas.

* Values with an asterisk are approximately calculated based on the bar graph presented in the corresponding reference.

Figure 1-1. Lignin biosynthesis pathway adapted from Boerjan et al. (2003) and Bonawitz and Chapple (2010). Grey arrows indicate alternative pathways that are thought to function in the absence of particular enzymes, such as CCoAOMT and COMT. Abbreviations for enzymes: PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate:CoA ligase; HCT, hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyl transferase; C3’H, coumarate 3’-hydroxylase; CCoAOMT, caffeoyl CoA O-methyltransferase; CCR, cinnamoyl CoA reductase, F5H, ferulate 5-hydroxylase; COMT, caffeic acid O-methyltransferase; CAD, cinnamyl alcohol dehydrogenase.
CHAPTER 2
ISOLATION AND EXPRESSIONAL CHARACTERIZATION OF CAFFEIC ACID O-METHYLTRANSFERASE (COMT) IN SUGARCANE

Introduction

The monolignol biosynthetic pathway involves multiple O-methylations of hydroxyl groups at the C3 and/or C5 positions of the monolignol precursors (Boerjan et al., 2003). Caffeic acid O-methyltransferase (COMT) is one of two O-methyltransferases with caffeoyl CoA 3-O-methyltransferase (CCoAOMT) in the monolignol biosynthetic pathway. In plants, COMTs are encoded by a multigene family that belongs to the type-1 family of diverse S-adenosyl-L-methionine (SAM) dependent O-methyltransferases (OMTs) (Louie et al., 2010). Unlike type-2 plant OMTs to which CCoAOMT belongs, type-1 OMTs do not require divalent cations for catalytic activity. In COMT-mediated O-methylation, the 3- or 5-hydroxyl group on the substrate is deprotonated by catalytic residues, and a reactive methyl group is transferred from SAM to the phenolate yielding S-adenosyl-L-homocysteine (SAH) and the methyl ether derivatives as final products (Zubieta et al., 2002).

In monolignol biosynthesis, COMT is primarily involved in sinapyl alcohol biosynthesis, catalyzing 5-O-methylation at 5-hydroxyconiferyl aldehyde yielding sinapyl aldehyde. COMT was historically known to catalyze 3- or 5-O-methylations at the level of free hydroxycinnamic acids including caffeic and 5-hydroxyferulic acid in monolignol biosynthesis. However, COMT exhibits a strong kinetic preference for aldehydes (5-hydroxyconiferaldehyde and caffeoyl aldehyde) and slightly less for alcohols (5-hydroxiconiferyl alcohol and caffeoyl alcohol) favoring 5-O-methylation over 3-O-methylation (Dixon et al., 2001; Li et al., 2000; Louie et al., 2010; Osakabe et al., 1999; Parvathi et al., 2001; Zubieta et al., 2002). Furthermore, COMT deficient mutants and
down-regulated transgenic plants showed a significant reduction in S units and the incorporation of 5-hydroxyconiferyl alcohol into lignin, which is derived from 5-hydroxyconiferyl aldehyde (Guo et al., 2001; Palmer et al., 2008; Ralph et al., 2000).

The major route for sinapyl alcohol biosynthesis passes through 5-O-methylation at 5-hydroxyconiferyl aldehyde. However, given the broad substrate permissiveness of COMT and presence of multiple family members, 3-O or 5-O-methylation could also occur at the level of caffeoyl aldehyde, caffeyl alcohol, and 5-hydroxyconiferyl alcohol. Residual S unit content and COMT activity in COMT knock-out mutants such as maize bm3 and sorghum bmr12 suggest that multiple COMTs could be partially involved in monolignol biosynthesis, depending on different developmental stages and cell types (Barrière et al., 2004; Dixon et al., 2001; Parvathi et al., 2001).

Sugarcane has a highly polyploid genome with an average 12 haplotypes. Despite the high level of ploidy and the interspecific origin of commercial sugarcane cultivars, homo(eo)logous genes are considered functional and able to translate complete proteins (Garsmeur et al., 2011; Jannoo et al., 2007). Considering the complexity of the sugarcane genome, ~31 different consensus COMT EST sequences are clustered within the genome (Ramos et al., 2001), which could be potential homo(eo)logous genes and/or gene family members derived from segmental duplications. Only one full-length COMT gene has been isolated and identified in sugarcane (GenBank accession no. AJ231133) (Selman-Housein et al., 1999). The previously identified COMT shows preferential expression in stems and lignified tissues such as the epidermis, xylem, and sclerenchyma (Ruelland et al., 2003; Selman-Housein et al., 1999).
In an attempt to modify lignin biosynthesis by suppressing \textit{COMT} through an RNAi approach, \textit{COMT} genes were isolated from a commercially important sugarcane cultivar, CP88-1762, by cDNA library screening and PCR. The expression patterns of the \textit{COMT} were investigated in different tissues and at various developmental stages. In addition, putative \textit{COMT} homo(eo)logous gene sequences were retrieved from a public database and compared with the isolated full-length \textit{COMT} gene.

\textbf{Materials and Methods}

\textbf{Isolation of \textit{COMT} Gene from CP88-1762}

Tissue samples comprised of leaves, internodes, nodes and immature leaf whorls of sugarcane (\textit{Saccharum} spp. hybrids) var. CP88-1762 were obtained from the University of Florida, Plant Science Research and Education Unit, Citra, Florida. Roots and newly emerging shoots were collected from the plants grown in the Technaflora B.C. Grow hydroponic solution (1/1000 dilution) (Technaflora Plant Products Ltd., Mission, British Columbia, Canada). Total RNA was extracted from each tissue using Trizol (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. Total RNA from each sample was mixed in equal proportion. cDNAs were synthesized from 1 \(\mu\)g of total RNA using the iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA). Forward (5'-GTCTCTCTCTTTGTATCCTCC-3') and reverse (5'-ATTCCGACATTAGAATCCAGAAT-3') primers were designed from 5' and 3' UTR sequences of \textit{S. officinarum} caffeic acid 3-O-methyltransferase (GenBank accession No. AJ231133). PCR was carried using a MyiQ cycler (Bio-Rad) with iTaq polymerase (Bio-Rad) with the following cycles: 15 min at 95°C, followed by 35 cycles of 94°C for 1 min, 52°C for 1 min, and 72°C for 1 min, and final elongation at 72°C for 10 min. PCR
products were cloned into pCRII-TOPO (Invitrogen), and a total of six clones were sequenced.

**cDNA Library Construction and Screening**

The mixture of the aforementioned total RNA was used for cDNA library construction. mRNA was purified from the total RNA using Oligotex mRNA Mini Kit (Qiagen, Valencia, CA, USA). The cDNA library was constructed using the ZAP-cDNA Synthesis Kit and ZAP-cDNA Gigapack III Gold Cloning Kit (Stratagene, Santa Clara, CA, USA) according to the manufacturer’s instructions. cDNAs synthesized from 5.9 μg of mRNA were fractionized using Sepharose CL-CB gel filtration column (Stratagene). cDNAs over 500 bp were ligated into Uni-ZAP XR, and recombinant lambda phage were picked and amplified. Total 4×10⁵ pfu of the cDNA containing phage were primarily screened using a 346 bp sugarcane COMT probe, spanning the SAM binding site. Secondary screening was performed using the phage with positive signals from the primary screening. Inserts-containing pBluscript phagemids were excised from Uni-ZAP XR, and cDNAs of approximately 2.0 kb were sequenced.

**Sequence Analysis**

To retrieve tentative consensus (TC) sequences for COMT, a keyword search was performed against Dana Farber Cancer Institute (DFCI) Sugarcane Gene Index (http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=s_officinarum). Nucleotide sequences and deduced amino acid sequences were analyzed using ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with the default parameters. Conserved regions of COMTs were analyzed against structurally characterized *Lolium perenne* COMT (LpCOMT, GenBank accession no. AAD10253) (Louie et al., 2010).
Sampling for Expression Analysis of the COMT Gene

Sugarcane tissues were collected from 8-month-old greenhouse grown plants. Samples were immediately frozen in liquid nitrogen and stored at -80°C until further analysis. Leaf blade, mid-rib, and sheath were collected from the top-visible dewlap (TVD) leaf (Leaf+1). After collecting a TVD leaf and leaf sheath, the remaining leaves and sheaths were removed from the stalk. Leaf whorls, shoot apical meristems, internodes, nodes, shoots, and roots were collected from the stalk. Internodes and nodes were numbered in order from shoot apical meristem to the bottom of the stalk. Shoots were newly emerging tillers and they were collected before leaf expansion. Roots were collected from below the stem-root junction, and washed with DEPC-treated water before freezing.

Epidermal, vascular, and ground tissues were collected from the fourth internode below the shoot apical meristem. After removing the leaf sheath, the internode was transversely and longitudinally sectioned and immediately frozen in liquid nitrogen. The sections were slowly thawed at 4°C. Vascular tissues were detached from ground tissues using forceps, and epidermal tissues were cut under a dissecting microscope. The tissues collected were immediately frozen in liquid nitrogen until RNA extractions could be performed.

RT-PCR and Quantitative Real-Time RT-PCR

Total RNA from each tissue was extracted using the Trizol reagent (Invitrogen) according to the manufacturer’s instructions. To prevent genomic DNA contamination, total RNA was treated with RNase-Free RQ1 DNase (Promega, Madison, WI, USA). cDNA was synthesized from 500 ng of DNase-treated total RNA using the iScript cDNA Synthesis Kit (Bio-Rad). The primers (forward: 5'-TAAATACGACACCTGCTGCT-3'
and reverse: 5'-ATTCGACAATTAGAATCCAGAACAT-3') were designed for amplification of a fragment of the 3' UTR region. Sugarcane GAPDH primers (forward: 5'-CACGGCCACTGGAAGCA-3' and reverse: 5'-TCCTCAGGGTTCTGATGCC-3') were used to amplify a fragment of the sugarcane GAPDH gene as a reference gene for normalization of transcripts as described by Iskandar et al. (2004). RT-PCR was performed in the MyiQ cycler (Bio-Rad) with Phire Hot Start DNA polymerase (New England Biolabs, Ipswich, MA, USA) with the following conditions: 30 s at 98°C, followed by 26 cycles of 98°C for 5 s, 60°C for 5 s, and 72°C for 20 s, and final elongation at 72°C for 1 min. PCR products were analyzed by 1.2% agarose gel electrophoresis and visualized with ethidium bromide. Transcript abundance of COMT was quantified in epidermal, vascular, and ground tissues of the internodes. Quantitative real-time RT-PCR was performed in the MyiQ cycler (Bio-Rad) with iQ SYBR Green Supermix (Bio-Rad) with the following conditions: 95°C for 3 min denaturation, 40 cycles at 95°C for 10 s and 55°C for 45 s. Ampliﬁon speciﬁcity was verified by melt-curve analysis from 55°C to 95°C and by agarose gel electrophoresis. The relative expression of COMT in each tissue was calculated using $2^{[Ct(GAPDH) − Ct(COMT)]}$.

Results

**Isolation of COMT Genes Through cDNA Library Screening and PCR**

cDNA library screening was performed to isolate putative COMT homo(eo)logs and other closely related genes within the COMT gene family of sugarcane. A total of six clones were sequenced out of 27 clones screened from the cDNA library derived from CP88-1762. Three different COMT genes were identified and designated as COMTa, b, and c. The PCR products were identical to COMTa and COMTb. COMTc
was not detected in the amplified PCR product. The length of their ORFs were 1089 bp and corresponded to 362 deduced amino acids in all of the isolated COMTs and the previously identified COMT (GenBank accession No. AJ231133), designated as COMTr.

The overall sequence identity including the UTRs was 98% among COMTs. The sequence identities of the coding regions and the UTRs were 99% and 96%, respectively. Three different forms of the full-length COMT genes had several nucleotide substitutions and insertions/deletions (indels), compared with COMTr (Figure 2-1). The coding region of COMTa showed 100% sequence identity with COMTr, while COMTb and c had 2 and 3 nucleotide substitutions, respectively (Figure 2-1B). Indels were observed in both 3' and 5' UTRs of isolated COMTs compared to COMTr (Figure 2-1, A and C). In the deduced amino acid sequences, COMTb and COMTc had only one substitution, Ile294Val, compared to COMTr and COMTa (Figure 2-1D).

**Sequence Comparison Between the Isolated COMT and TCs**

A total of 22 TC sequences for COMT were retrieved from the DFCI S. officinarum Gene Index database. Of the TCs, the overall coverage of the complete open reading frames (ORFs) ranged from 19% to 100%. TC112705 showed 100% coverage and was identical with COMTb within the coding region, but had several substitutions in the 3' UTRs.

Among the 22 TCs, those containing 3' UTR sequences were selected, and their 3' UTRs were compared with that of COMTa. The 3' UTR sequences of one group of TCs (TC112705, TC114280, TC127415, TC138868, and TC115978) showed relatively high sequence identity with COMTa ranging from 95 to 99% (Figure 2-2A), similar to the sequence identity among the isolated COMT (96%). The 3' UTR sequences of the other
groups (TC 120283 and TC121675) were relatively diverse compared to COMTa with 90 and 80% sequence identities, respectively (Figure 2-2B).

To investigate whether the sequences were conserved in the coding regions among TCs and COMTs, six TCs containing the sequence encoding the SAM binding residues were selected and analyzed. Sequence identity in this region from +410 to +755 of COMTa among the selected TCs (TC112705, TC116269, TC121675, TC149781, TC120283 and TC1337730) and COMTa ranged from 95% to 100% (Figure 2-3). In particular, TC121675 and TC120283, which had relatively diverse 3' UTR sequences compared with other TCs and COMTa, showed 100 and 98% sequence identity with COMTa in this region.

**Deduced Amino Acid Sequence Comparison of the Isolated COMTa with Other Monocot COMTs**

The deduced amino acid sequence from sugarcane COMTa showed 94, 91, and 79% similarity with sorghum, maize, and Lolium perenne COMTs, respectively. A pairwise comparison between the sugarcane COMT and the previously characterized additional monocot COMTs showed that the substrate, the SAM binding site, and the catalytic residues were perfectly conserved without any amino acid substitutions (Figure 2-4).

**Expression of the COMT in Different Tissues and Developmental Stages**

COMT transcripts were detected in all of the collected tissues, but the expression level was higher in the internodes, nodes, and shoots compared with the shoot apical meristem, leaves, or roots (Figure 2-5A). In the internodes and the nodes, lower transcript abundance was apparent in relatively mature tissues, the ninth internode and node (Figure 2-5A). Examining internodes, COMT was highly expressed in the vascular
bundle and in the epidermal tissues, and significantly less in ground tissue (Figure 2-5, B and C). Quantitative real-time RT-PCR results showed that expression levels of COMT in vascular and epidermal tissues were 2.5 and 2.2 fold higher than in ground tissue, respectively (Figure 2-5C).

Discussion

Three different forms of COMT were isolated from sugarcane cultivar, CP88-1762, through cDNA library screening and PCR. Structural characterization of COMTs in alfalfa and Lolium perenne elucidated the catalytic residues responsible for methyl transfer and the binding domains for recognition of SAM and the monolignol precursors (Louie et al., 2010; Zubieta et al., 2002). Maize and sorghum COMTs have been functionally characterized as the genes responsible in lignin defective brown-midrib mutants, bm3 and bmr12, respectively (Bout and Vermerris, 2003; Vignols et al., 1995). The isolated sugarcane COMT has perfectly conserved substrate and SAM binding domains and catalytic residues. Therefore, it is expected that the sugarcane COMT would have nearly identical catalytic properties toward monolignol precursors, compared with the previously characterized lignin biosynthetic COMTs.

The sugarcane COMT was preferentially expressed in developing and lignifying tissues, such as relatively immature internodes and nodes. Higher expression levels were evident from stem sections compared to leaves or roots which were similar to the previously reported COMT expression patterns in sugarcane and alfalfa (Inoue et al., 1998; Selman-Housein et al., 1999). In the internodes, COMT is highly expressed in the epidermal and vascular tissues where lignification takes place. Consistent with this result, the previously identified sugarcane COMT protein accumulated in the vascular bundles, and maize COMT transcripts were preferentially expressed in differentiated
xylem cells and sclerenchyma cells under the epidermis (Ruelland et al., 2003; Vignols et al., 1995). Thus, gene expression of the sugarcane COMT in lignifying tissues would support that the COMT is involved in monolignol biosynthesis.

Although assignment of homo(eo)logs remains to be clarified, a high level of sequence identity (> 95%) in both the coding and UTRs among the isolated COMTs and most of the TC sequences might indicate these genes are homo(eo)logs. In sugarcane, not only homologous but also homoeologous alleles display a high level of collinearity and sequence conservation (Garsmeur et al., 2011; Jannoo et al., 2007). Sequence identity among homo(eo)logous genes was 95.9 and 87.5% in the exon and intron regions, respectively (Garsmeur et al., 2011). Low copy numbers in the Southern blots hybridized with the sugarcane COMT probe also suggests a high level of sequence and structural conservation among COMT genes located in homo(eo)logous chromosomal regions (Selman-Housein et al., 1999). Moreover, other COMT family members originated by single gene duplications may have limited sequence identity and be below the limit of detection in Southern hybridizations.

Compared to gene copies derived from single duplications, such as dispersed and transposon based duplications, duplicated genes that have emerged through polyploidization, especially autopolyploidization, tend to undergo relatively low rates of sequence substitutions, gene loss, and diversification in expression and function (Parisod et al., 2010; Wang et al., 2012). Thus, it could be anticipated that COMT homo(eo)logous genes are retained in the sugarcane genome with a smaller degree of divergence in expression and function.
Given the high degree of polyploidy and potential functional redundancy among homo(eo)logous genes in the sugarcane genome, it would be challenging to isolate and identify lignin defective mutants in sugarcane. One advantage in deploying RNAi for metabolic engineering is that homo(eo)logous genes or gene family members can be simultaneously down-regulated in the host genome, targeting conserved regions among genes of interest (Lawrence and Pikaard, 2003; Travella et al., 2006). Among the isolated COMTs and TCs, coding regions spanning SAM binding domains show over 95% sequence identity. In general, contiguous sequence identity greater than 21 bp and/or overall homology greater than 88% between transgene and endogenous sequences could induce RNAi (Travella et al., 2006; Xu et al., 2006). Therefore, this coding region would be a good target for the simultaneous suppression of functionally redundant COMT genes in sugarcane.
| COMTr | MGSTAEDVAAVADEEACNYAMQLASILPMTLKNALELGLLEVLQAEAPAGKALAPEEV 60 |
| COMTa | MGSTAEDVAAVADEEACNYAMQLASILPMTLKNALELGLLEVLQAEAPAGKALAPEEV 60 |
| COMTb | MGSTAEDVAAVADEEACNYAMQLASILPMTLKNALELGLLEVLQAEAPAGKALAPEEV 60 |
| COMTc | MGSTAEDVAAVADEEACNYAMQLASILPMTLKNALELGLLEVLQAEAPAGKALAPEEV 60 |
| COMTr | VARLPVAPTNPDAADMVDRMLRLLASYDVVKCMQEDKGYERRYSAAPVGKWLTPNEDG 120 |
| COMTa | VARLPVAPTNPDAADMVDRMLRLLASYDVVKCMQEDKGYERRYSAAPVGKWLTPNEDG 120 |
| COMTb | VARLPVAPTNPDAADMVDRMLRLLASYDVVKCMQEDKGYERRYSAAPVGKWLTPNEDG 120 |
| COMTc | VARLPVAPTNPDAADMVDRMLRLLASYDVVKCMQEDKGYERRYSAAPVGKWLTPNEDG 120 |
| COMTr | VSMAALTLMNQDKVLMESNYYLKDAVLDDGIPFNKAYGMTAFEYHGDPRFNPVFNEGMK 180 |
| COMTa | VSMAALTLMNQDKVLMESNYYLKDAVLDDGIPFNKAYGMTAFEYHGDPRFNPVFNEGMK 180 |
| COMTb | VSMAALTLMNQDKVLMESNYYLKDAVLDDGIPFNKAYGMTAFEYHGDPRFNPVFNEGMK 180 |
| COMTc | VSMAALTLMNQDKVLMESNYYLKDAVLDDGIPFNKAYGMTAFEYHGDPRFNPVFNEGMK 180 |
| COMTr | NHSVIITKKLLEFYTFEGVSTLVDVGGGIGATLHAITSHHPQIKGINFDLPHVISEAPP 240 |
| COMTa | NHSVIITKKLLEFYTFEGVSTLVDVGGGIGATLHAITSHHPQIKGINFDLPHVISEAPP 240 |
| COMTb | NHSVIITKKLLEFYTFEGVSTLVDVGGGIGATLHAITSHHPQIKGINFDLPHVISEAPP 240 |
| COMTc | NHSVIITKKLLEFYTFEGVSTLVDVGGGIGATLHAITSHHPQIKGINFDLPHVISEAPP 240 |
| COMTr | FPGVQHVGDMSFKVPAIGDAILMKWILHDNDSADCAITLKLNCYDAFLENGKVIIVECVLP 300 |
| COMTa | FPGVQHVGDMSFKVPAIGDAILMKWILHDNDSADCAITLKLNCYDAFLENGKVIIVECVLP 300 |
| COMTb | FPGVQHVGDMSFKVPAIGDAILMKWILHDNDSADCAITLKLNCYDAFLENGKVIIVECVLP 300 |
| COMTc | FPGVQHVGDMSFKVPAIGDAILMKWILHDNDSADCAITLKLNCYDAFLENGKVIIVECVLP 300 |
| COMTr | VNTEAVPKAQGVFHVDMILHMLAHNFGREREFHDLKAGGFSGKFATIYANAWAIIF 360 |
| COMTa | VNTEAVPKAQGVFHVDMILHMLAHNFGREREFHDLKAGGFSGKFATIYANAWAIIF 360 |
| COMTb | VNTEAVPKAQGVFHVDMILHMLAHNFGREREFHDLKAGGFSGKFATIYANAWAIIF 360 |
| COMTc | VNTEAVPKAQGVFHVDMILHMLAHNFGREREFHDLKAGGFSGKFATIYANAWAIIF 360 |

| COMTr | IK 362 |
| COMTa | IK 362 |
| COMTb | IK 362 |
| COMTc | IK 362 |

Figure 2-1. Nucleotide and amino acid comparisons among the isolated sugarcane COMTs. A) Nucleotide sequences of the 5' UTR. B) Nucleotide sequences in the coding region. Coding sequences for start and stop codons are underlined. C) Nucleotide sequences of the 3’ UTR. D) Deduced amino acid sequence comparisons among the isolated sugarcane COMTs. Underlined red letters indicate sequence substitutions compared with COMTr.
**Figure 2-2.** Nucleotide sequence comparison of the 3′ UTR between *COMTa* and TCs.

**A)** 3′ UTR sequence of TCs showing over 95% identity with that of *COMTa*.

**B)** 3′ UTR sequence of TCs showing less than 90% identity with that of *COMTa*. Coding sequences for the stop codon are underlined. Asterisk indicates identical sequences among aligned sequences.
Figure 2-3. Nucleotide sequence comparison of coding regions spanning SAM binding residues between TCs and COMTa. Underlined red letters indicate sequence substitutions compared with COMTa.
Figure 2-4. Amino acid sequence comparisons between sugarcane COMTs and functionally characterized COMTs in monocot plants. SbCOMT (Sorghum bicolor, GenBank accession no. AAO43609). ZmCOMT (Zea Maize, GenBank accession no. Q06509). LpCOMT (Lolium perenne, GenBank accession no. AAD10253). Conserved regions among plant COMTs are presented according to the structurally characterized LpCOMT as described in Louie et al. (2010). Conserved residues involved in substrate and SAM binding are shaded yellow and red, respectively. Catalytic residues are underlined and marked with bold characters in His266, Asp267, and Glu326 (following the numbering in LpCOMT). Red letters indicate a sequence substitution between sugarcane COMTs.
Figure 2-5. **COMT** expression patterns in sugarcane. A) RT-PCR of **COMT** expression in different tissues and at different developmental stages. SAM, shoot apical meristemaitc tissue; Leaf whorl, immature leaf whorl above SAM; Leaf +1, leaf with top visible dewlap; Internode +1 ~ +9 and Node +3 ~ +9, the number of internodes or nodes counted from SAM; Shoots, emerging tiller; Roots, roots collected from stem-root junction region. B) and C) RT-PCR and quantitative real-time RT-PCR of **COMT** expression in vascular, ground, and epidermal tissues collected from the internode. Sugarcane **GAPDH** gene was used as an internal control.
CHAPTER 3
GENERATION OF COMT SUPPRESSED TRANSGENIC SUGARCANE AND EFFECTS OF COMT SUPPRESSION ON LIGNIN, PLANT GROWTH, AND SACCHARIFICATION PERFORMANCE UNDER GREENHOUSE CONDITIONS

(Re-printed from Jung et al. 2012 with permission from Plant Biotechnology Journal)

Introduction
The environmental, political and economic challenges associated with our dependence on fossil fuels, have sparked a tremendous interest in developing efficient biomass feedstocks to satisfy the increasing demand for sustainable energy sources (Byrt et al., 2011; Hisano et al., 2009; Tew and Cobill, 2008).

Sugarcane (Saccharum spp. hybrids) is a prime herbaceous biofuel feedstock. Sugarcane’s dry biomass production per unit area (39 Mg ha\(^{-1}\); stalk, leaves and tops) is significantly higher than maize (17.6 Mg ha\(^{-1}\); grain and stover), switchgrass (10.4 Mg ha\(^{-1}\); biomass), or Miscanthus (29.6 Mg ha\(^{-1}\); biomass) (Heaton et al., 2008; Waclawovsky et al., 2010). Its perennial growth habit and C\(_{4}\) photosynthetic pathway maximize carbon sequestration while minimizing light requirements, water, and nitrogen inputs (Byrt et al., 2011; Somerville et al., 2010). Sucrose accumulates in the stalk internodes of sugarcane and is either utilized for sugar production or readily fermented to the transportation fuel ethanol. The abundant lignocellulosic sugarcane residues are currently underutilized for bioenergy production (Leite et al., 2009; Somerville et al., 2010). Including these residues for second-generation biofuel production has the potential to boost biofuel yields per unit land area compared to current sucrose-based conversion technologies.

Diminished recalcitrance to enzymatic hydrolysis is a desirable trait for lignocellulosic feedstocks. The presence of lignin in the cell wall exacerbates biomass recalcitrance and limits bioconversion of lignocellulosic biomass into fermentable sugars (Jørgensen et al., 2007; Mansfield et al., 1999; Weng et al., 2008). Energy-intensive thermo-chemical pretreatments are required to degrade the cell wall matrix and to disrupt the crystalline structure of cellulose, thereby increasing binding sites for cellulolytic enzymes (Mosier et al., 2005). Pretreatments are followed by saccharification, during which the cell wall polymers, primarily cellulose and hemicellulose, are enzymatically hydrolyzed into monomeric sugars for fermentation (Lu and Mosier, 2008).

Lignin is an aromatic, hydrophobic polymer primarily consisting of \( p \)-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, which are polymerized by radical coupling of three different monolignols: \( p \)-coumaryl, coniferyl, and sinapyl alcohol, respectively (Boerjan et al., 2003). During saccharification, cellulases can adsorb irreversibly to lignin, thereby reducing the overall enzymatic activity (Chernoglazov et al., 1988; Jørgensen et al., 2007). The manipulation of lignin biosynthesis in feedstocks is a prime strategy to reduce biomass recalcitrance and improve fermentable sugar yields.

Ten different enzymes catalyze a series of hydroxylation, methylation, and side chain reduction reactions of monolignol precursors (Bonawitz and Chapple, 2010). Down-regulation of monolignol biosynthetic genes in transgenic alfalfa reduced total lignin content and altered lignin subunit composition. Fermentable sugar yields, with or without pretreatment, were shown to be negatively correlated with total lignin content (Chen and Dixon, 2007). Down-regulation of enzymes involved in the terminal steps of
the monolignol biosynthetic pathway including COMT or CAD generally had little or no negative effects on plant growth (Chen and Dixon, 2007; Fu et al., 2011a; Jackson et al., 2008; Pilate et al., 2002). In grasses, biomass conversion is also influenced by lignin subunit composition based on studies in maize and sorghum brown midrib mutants (Dien et al., 2009; Saballos et al., 2008; Vermerris et al., 2007). Recently, transgenic approaches for manipulating lignin biosynthesis have been successfully applied to switchgrass (Fu et al., 2011a; Saathoff et al., 2011; Xu et al., 2011).

Caffeic acid O-methyltransferase (COMT) functions late in the monolignol biosynthetic pathway and, despite its name, methylates 5-hydroxyconiferyl aldehyde and 5-hydroxyconiferyl alcohol to form S unit precursors, sinapyl aldehyde and sinapyl alcohol, respectively (Bout and Vermerris, 2003; Guo et al., 2001; Humphreys et al., 1999; Osakabe et al., 1999). In sugarcane, ~31 different consensus EST sequences are clustered to COMT, which could be potential allelic variants or homo(eo)logous genes (Ramos et al., 2001). The previously identified full-length COMT (GenBank accession no. AJ231133) shows 91% amino acid similarity with maize lignin-specific COMT, and is preferentially expressed in stems and roots, and in lignifying tissues such as epidermis, xylem, and sclerenchyma (Ruelland et al., 2003; Selman-Housein et al., 1999).

Conventional breeding and genetic engineering of sugarcane are challenging because of its highly polyploid genome derived from interspecific hybridization (Dal-Bianco et al., 2011; Lakshmanan et al., 2005). Although most plants examined to date share many of the enzymatic reactions leading to the synthesis of lignin (Xu et al., 2009), limited information is available for sugarcane and none of the lignin biosynthetic genes
have been functionally characterized through forward or reverse genetics (Ramos et al., 2001; Ruellan et al., 2003; Selman-Housein et al., 1999).

Biolistic or *Agrobacterium*-mediated gene transfer into sugarcane has typically been achieved by using embryogenic callus as target tissue (Arencibia et al., 1998; Bower and Birch, 1992; Gallo-Meagher and Irvine, 1996). Recent improvements for biolistic gene transfer focused on the reduction of the tissue culture period by using direct embryogenesis and by delivering minimal expression cassettes instead of plasmids (Kim et al., 2012; Taparia et al., 2012a; Taparia et al., 2012b). A number of transgenes have been introduced into sugarcane to incorporate traits including herbicide resistance, tolerance to abiotic or biotic stress, and production of sugar and value-added metabolites (Altpeter and Oraby, 2010). However, lignin modification or the application of RNAi for crop improvement has not been reported in sugarcane.

Here we describe the generation of transgenic sugarcane with RNAi suppression of *COMT* and demonstrate that the resulting transgenics are more amenable to biomass conversion for the production of fuels and chemicals.

**Materials and Methods**

**Plant Growth**

Commercially important sugarcane cv. CP88-1762 stalks were collected from the Everglades Research and Education Center, University of Florida, Belle Glade, Florida, USA. Single node segments from these wild-type plants were transplanted to 15 liter pots containing Fafard No. 2 mix (Conrad Fafard, Agawam, MA, USA) and grown to maturity under natural photoperiod in an air-conditioned greenhouse set at 28°C/22°C (day/night). Plants were irrigated once a day and fertilized biweekly with Miracle-Gro Plant and Lawn Food (Scotts Miracle-Gro, Marysville, OH, USA).
The primary transgenic lines and wild-type plants were clonally propagated by single node segments of mature plants and in 15 liter pots containing Fafard No. 2 mix. Plants were arranged in a randomized block design, with eight replications, and grown under the conditions as described above. A total of four stalks were maintained in each pot by removing juvenile tillers weekly. Above-ground fresh weight was measured from the most mature tiller of the 7-month-old plants. After removing leaves and leaf sheath, stalk diameter was measured in the middle of the stalk, and length was measured from soil surface to the apical meristem of the stalk. The basal internode of the transgenic, control, and wild-type plant were transversely or longitudinally cut, and photographed immediately without fixation and staining to display the brown coloration.

**Vector Construction**

The COMT RNAi vector was constructed using the pWFOsC4H::Bg4CLi vector (Fouad et al., 2010). To amplify a 346 bp COMT fragment from the sugarcane cDNA, a pair of primers (forward: 5ʹ-AGAGCTGGTACTACCTCAAGGACG-3ʹ, reverse: 5ʹ-GTTTAAACATGTCCCCGCCGACGTG-3ʹ) was designed to the sugarcane COMT sequence (GenBank accession no. AJ231133). This fragment was located on the COMT coding region between nucleotide +410 and +755, spanning part of the highly conserved SAM binding pocket among plant COMTs (Louie et al., 2010). PCR products were cloned into the pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA, USA), and confirmed by sequencing. XbaI and EcoRV restriction enzyme sites were introduced to the 5' end of the COMT fragment and cloned into the pCR2.1-TOPO vector. The inverted repeats of Bg4CL fragment in the pWFOsC4H::Bg4CLi (Fouad et al., 2010) were replaced by the COMT fragment with two subsequent cloning steps. The resulting COMT RNAi cassette consisted of the 1994 bp *Oryza sativa* cinnamate 4-hydroxylase
(C4H) gene promoter (GenBank accession no. AC136224), inverted repeats of the 346 bp sugarcane COMT fragment separated by the 94 bp of Paspalum notatum 4-coumarate : CoA ligase (4CL) intron (Fouad et al., 2010), and the CaMV 35S 3’ UTR (Figure 3-1, A). The pJFNPTII vector (Altpeter et al., 2000) provided the selectable marker and contained the neomycin phosphotransferase II (nptII) gene under the transcriptional control of the Zea mays ubiquitin promoter with first intron and CaMV 35S 3’ UTR.

**Generation of Transgenic Sugarcane**

Embryogenic sugarcane calli were induced as described by Chengalrayan and Gallo-Meagher (2001). Immature leaf whorls were cut transversely into 2~5 mm pieces and cultured on modified MS basal medium (CI-3), supplemented with 20 g L⁻¹ sucrose and 13.6 μM 2,4-D, with the pH adjusted to 5.8. Tissues were sub-cultured on CI-3 media every 2 weeks. Ten weeks after callus induction, particle bombardment was performed using the PDS-1000/He biolistic particle delivery system (Bio-Rad, Hercules, CA, USA) as previously described (Altpeter et al., 1996). For delivery of the minimum linear transgene cassette (MC), the COMT RNAi cassette and nptII expression cassettes were released by restriction enzyme digestion with XmnI and I-SceI, separated by gel electrophoresis, and extracted using the QIAquick Gel Extraction kit (Qiagen, Valencia, CA, USA). COMT and nptII – MC DNAs were mixed in a 2:1 molar ratio, co-precipitated onto 1 μm diameter gold micro-carriers (Bio-Rad) and accelerated into target tissues as described earlier (Altpeter et al., 2010). 24 shots of bombardment were performed with the following conditions: 1100 psi rupture pressure, 27.5 Hg vacuum pressure, 6 cm distance from the rupture disk to calli. Calli were transferred to CI-3 media containing geneticin (G-418, 30 mg L⁻¹) 6 days after bombardment, for 3
biweekly subcultures (Kim et al., 2012). Calli that survived were regenerated into plants on media containing 50 mg L\(^{-1}\) thidiazuron (TDZ) for 2 weeks. For further selection, elongation, and rooting of shoots, regenerated plantlets were transferred to hormone free media containing 30 mg L\(^{-1}\) paromomycin for four biweekly subcultures. Regenerated plants with roots were transplanted in 0.65 liter pots with Fafard No. 2 mix and placed in a growth room with 80% relative humidity with 16 h photoperiod and 500 µmol m\(^{-2}\) s\(^{-1}\) light intensity.

**Evaluation of NPTII Expression**

NPTII expression was evaluated by NPTII ELISA using a commercial kit (Agdia, Elkhart, IN, USA). Total protein was extracted from leaves and quantified with the Bradford assay (Bradford, 1976). The ELISA assay was performed using 20 µg of total protein, according to the manufacturer’s instruction. NPTII expression of putative transgenic plants was qualitatively evaluated by color development in comparison with the supplied NPTII standard and wild-type protein extracts.

**PCR Analysis**

The presence of the COMT RNAi cassette in genomic DNA extracts of transgenic lines was confirmed by PCR. Genomic DNA was extracted from leaves using the CTAB method (Murray and Thompson, 1980), and 75 ng was used per reaction as a template for amplification. Primers (forward: 5’-CCTGCTAGTCTTCTCTCATTGTT-3’ and reverse: 5’-GTGATGATGACCGAGTTCTTCT-3’) designed to the C4H promoter region and annealing to the sense fragment of COMT (Figure 3-1A) were used with an expected amplification product of 550 bp. PCR was performed in the MyiQ cycler (Bio-Rad) with iTaq DNA Polymerase (Bio-Rad) under the following conditions: 95\(^\circ\)C for 3
min denaturation, 35 cycles at 95°C for 30 s, 60°C for 30 s, 72°C for 1 min, and final extension at 72°C for 7 min.

**Southern Blot Analysis**

High molecular weight genomic DNA was extracted from leaves using the CTAB method (Murray and Thompson, 1980). Twenty microgram of genomic DNA were digested to completion with *EcoRI*, separated by electrophoresis on 1.0% agarose gel, and transferred onto the Hybond-N+ membrane using the manufacturer’s instructions (GE Healthcare Biosciences, Pittsburgh, PA, USA). Probes were generated by PCR to the *C4H* promoter region of the *COMT* RNAi cassette and labeled with ³²P-dCTP (Perkin Elmer, Waltham, MA, USA) using the Prime-a-Gene Labeling System (Promega, Mannheim, Germany). Hybridization and washing were performed according to the manufacturer’s instructions, and membranes were exposed to Kodak X-ray film (Fisher Scientific, Atlanta, GA, USA) at -80°C for 2 d.

**Small-RNA Northern Blot**

The third internode below the apical shoot meristem was collected from primary transgenic or wild-type plants, and total RNA was extracted from 3 g of internode tissue using the modified hot SDS/phenol method (Shirzadegan et al., 1991). Briefly, 3 g of internode was ground under liquid nitrogen, and homogenized by the addition of pre-warmed (65°C) 10 mL acidic phenol (pH 4.5) and 10 mL of extraction buffer containing 0.1 M LiCl, 0.1 M tris-HCl (pH 8.0), 0.01M EDTA (pH 8.0), 1% SDS (w/v), and 0.1% PVP (w/v). Following 5 min incubation in a 65°C water bath, 10 mL of chloroform was added. The homogenate was thoroughly mixed, and centrifuged at 13,000 g for 15 min at 4°C. The upper aqueous phase was extracted once more with the equal volume of chloroform. Total RNA was precipitated by adding the equal volume of isopropanol, and
washed twice with 70% ethanol. Small RNA was separated from the total RNA using the method described in Lu et al. (2007). Briefly, high molecular weight RNA was precipitated by adding 1/10 volume of 50% (w/v) PEG (M.W. 8000) and 5 M NaCl. The mixture was incubated on ice for 1 h, and centrifuged at 16,000 g for 15 min at 4°C. The supernatant containing small RNA was transferred to a new tube and small RNA was precipitated with the equal volume of isopropanol for 16 h at -20°C. Small RNA was recovered by centrifugation at 16,000 g for 30 min at 4°C and pellet was washed twice with 80% ethanol. Thirty µg of small RNA was separated by electrophoresis in 15% polyacrylamide TBE/urea gel (Bio-Rad) and transferred to the Hybond-N+ membrane (GE Healthcare Biosciences) using a semi-dry transfer cell (Bio-Rad). The 346 bp COMT fragment, used for COMT RNAi vector construction, was labeled with 32P-dCTP (Perkin Elmer) using the Prime-a-Gene Labeling System (Promega) for use as a probe. Hybridization was carried out overnight at 38°C in Church’s hybridization buffer (Brown et al., 2001). Following hybridization, the membrane was briefly rinsed 2X SSC - 0.2% SDS and then washed twice with 2X SSC - 0.2% SDS at 50°C for 20 min each. The membrane was exposed to Kodak X-ray film at -80°C for 2 d.

**Quantitative Real-Time RT-PCR for Quantification of COMT Expression**

The third internode below the apical shoot meristem was collected from the clonally produced progenies of transgenic and wild-type plants. Total RNA was extracted using Trizol reagent (Invitrogen) according to the manufacturer’s instruction. To prevent genomic DNA contamination, total RNA was treated with RNase-Free RQ1 DNase (Promega). cDNA was synthesized from 500 ng of DNase treated RNA using iScript cDNA Synthesis Kit (Bio-Rad). Primers (forward: 5’- TAAATACGCACACCTGCTGCT-3’ and reverse: 5’-
ATTCGACAATTAGAATCCAGAACAT-3') were designed for amplification of the 3' UTR region of the targeted COMT gene. Sugarcane GAPDH primers (forward: 5'-CACGGCCACTGGAAGCA-3' and reverse: 5'-TCCTCAGGGTTCTGATGCC-3') were used to amplify a fragment of the sugarcane GAPDH gene as a reference gene for normalization of transcripts as described by Iskandar et al. (2004). Quantitative real-time PCR of the transcripts was performed in the MyiQ cycler (Bio-Rad) with iQ SYBR Green Supermix (Bio-Rad) under the following conditions: 95°C for 3 min denaturation, 40 cycles at 95°C for 10 s and 55°C for 45 s. Amplicon specificity was verified by melt-curve analysis from 55°C to 95°C and by agarose gel electrophoresis. COMT expression levels in transgenic plants relative to wild-type plants were calculated using the 2-ΔΔCt method (Livak and Schmittgen, 2001).

**Microscopic and Histochemical Analysis**

The transverse stem sections were hand-cut with a razor blade from the sixth internode below the shoot apical meristem in both wild-type and transgenic line T4. Mäule and Wiesner staining were performed according to Vermerris and Nicholson (2006). For Mäule staining, sections were immersed in freshly prepared 1% (w/v) potassium permanganate for 30 min at room temperature, and then washed with water for 2 min. The stained sections were then treated with 6M HCl for 1 min and excessive HCl was removed using tissue paper. For Wiesner staining, sections were immersed with 2% (w/v) phloroglucinol in 2:1 mixture of absolute ethanol and 12M HCl for 3 min at room temperature. Sections were observed using an Olympus BH-2 light microscope (Olympus, Tokyo, Japan). Images were taken and recorded in the Infinity 1 camera and Infinity analyzer (Lumenera Corporation, Ottawa, Ontario, Canada).
Sample Preparation and Determination of Lignin Content and Composition

Stalks with 12 internodes were collected from clonally propagated transgenic or wild-type plants, and internodes 1-3 below the shoot apical meristem, all leaves and leaf sheaths were removed. The remaining mature portion of the stalks was dried at 45°C, and ground using a Wiley mill (Thomas Scientific, Swedesboro, NJ, USA) with a 1.0 mm sieve. The ground samples were passed through 0.42 mm sieve to remove irregular particles. Following three successive extractions with 50% ethanol (v/v), under sonication at 45°C for 30 min, samples were dried at 45°C until constant weight.

The modified acetyl bromide method was used to determine lignin content (Foster et al., 2010; Hatfield et al., 1999a). Two milligram of extract-free dried sample was placed in a 2 mL polypropylene tube, and 1 mL of freshly prepared 25% (w/w) acetyl bromide/glacial acetic acid solution was added. The tubes were incubated in a water bath at 50°C for 4 h, and during the last hour, samples were thoroughly mixed at 15 min intervals, and placed on ice for 30 min. One hundred microliter of each reaction were transferred into a 2 mL polypropylene tube containing 200 µL of 2 M NaOH and 1.7 mL glacial acetic acid. The absorbance of the solution was determined at 280 nm using an Evolution 300 spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). The lignin content was calculated by employing the molar extinction coefficient of 21.5 L g⁻¹ cm⁻¹ for milled sugarcane vascular bundle lignin (He and Terashima, 1990).

Lignin composition was determined by thioacidolysis method following Robinson and Mansfield (2009). One milliliter of thioacidolysis reagent (2.5% boron trifluoride etherate and 10% ethanethiol, in distilled dioxane, v/v) was added to each reaction vial containing 10 mg of extract-free sample. Each vial was flushed with N₂, and tightly sealed with Teflon-lined screw-cap. The sample was incubated at 100°C for 4 h with
manual agitation every hour. After the reaction was halted at -20°C for 5 min, 0.2 mL of tetracosane (5 mg mL⁻¹ methylene chloride) was added to each vial as an internal standard, then the sample was acidified to pH between 3 and 4 with 0.3 mL of 0.4 M sodium bicarbonate. For the extraction of reaction products, 2 mL of distilled water and 1 mL of methylene chloride were added and vortexed. After phase separation, organic phase was recovered, and simultaneously passed though the Pasteur pipette packed with 50 mg of anhydrous sodium sulfate. Extracted samples were dried at 45°C under a stream of N₂, and then resuspended in 1 mL of methylene chloride. Twenty microliter of the sample was silylated by adding 20 μL of pyridine and 100 μL of BSTFA [N,O-bis (trimethylsilyl) acetamide] at room temperature for at least 2 h. Thioacidolysis derivatives for lignin monomers were analyzed using the Varian 3800 GC gas chromatograph coupled to the Varian 1200 MS mass spectrometer (Varian, Walnut Creek, CA, USA). One microliter of the sample was injected onto a Factor-4 VF-5ht column (35 m, 0.25 mm i.d.) with helium (1.2 mL min⁻¹) as carrier gas. GC conditions were as follows: Initial column temperature, 130°C, held for 3 min, ramped at 3°C min⁻¹ to 250°C, and held for 5 min; injector temperature, 250°C; and 1:2 split ratio. The mass spectrometer was operated in electron impact mode at 70 eV. The detector was operated at 1.0 kV. Mass spectra were recorded every 0.2 s at a scanning range of 50-550 m/z, and the date was registered and analyzed using Varian MS Workstation software. Thioacidolysis monomers, G and S were quantified by employing the response factors from internal control (IS) as follows: G versus IS, 0.47; and S versus IS, 0.53 (Yue et al., 2012).
**Dilute Acid Pretreatment and Enzymatic Hydrolysis**

Dilute acid pretreatment and enzymatic hydrolysis was performed according to published protocols (Chen and Dixon, 2007; Selig et al., 2008). Extractive-free samples (0.15 g) were soaked in 1.35 mL of dilute sulfuric acid [final concentration 1.3% (w/w)], and autoclaved at 121°C for 40 min. After autoclaving, the pretreated samples were washed twice with 25 mL distilled water. The pretreated or native biomass samples were placed in 50 mL polypropylene tubes, and suspended in 5 mL of 0.1 M sodium citrate buffer (pH 4.8). Autoclaved distilled water containing 100 µL of 2% (w/w) sodium azide, 6 FPU of Kerry Biocellulase W (Kerry Bioscience, Cork, Ireland), and 6.4 pNPGU of Novozyme 188 β-glucosidase (Sigma, Saint Louis, MO, USA) were added to bring the total volume to 10 mL. The hydrolysis was performed for 120 h in a shaking incubator at 50°C and 250 rpm. After enzymatic hydrolysis, 1 mL of the each sample was collected, filtered through a 0.2 µm syringe filter, and glucose yields were analyzed using an YSI glucose analyzer (YSI Life Science, Yellow Springs, OH, USA).

**Statistical Analysis**

ANOVA was performed using Proc GLM in SAS™ Version 9.3 (SAS Institute Inc., Cary, NC, USA). Statistical significance among means for biomass, stalk diameter, and length was determined using Tukey’s multiple comparisons at $P < 0.05$. T-tests were performed to determine whether the means of total lignin content, S/G ratios, and glucose yields were statistically significant between the transgenic plants and the wild-type, tissue culture or nptII-only transgenic controls ($P < 0.05$).
Results

Generation and Molecular Characterization of Transgenic Sugarcane Lines

The RNAi inducing transgene cassette contained a highly conserved region of COMT. This may suppress COMT expression from both targeted COMT and homo(eo)logous COMT’s in the highly polyploidy sugarcane. The targeted region shared 95-100% nucleotide sequence identity with five tentative consensus (TC) sequences, TC116269, TC120283, TC121675, TC133773, and TC149781, putative homo(eo)logs of sugarcane COMT (DFCI S. officinarum Gene Index; http://compbio.dfci.harvard.edu/cgi-bin/tgi/geneprod_search.pl). The promoter of Oryza sativa C4H gene was used to induce expression of RNAi inducing transgene in lignifying tissues.

Biolistic gene transfer was used to produce sugarcane transformants. Following 24 bombardments, regeneration of plants, and selection of putative transformants on geneticin and paromomycin, 43 lines tested positive for the NPTII enzyme as determined by ELISA (Figure 3-1B). This resulted in a transformation efficiency of 1.8 transgenic plants produced per bombardment. Thirty-eight of these 43 transgenic lines tested positive for the presence of the COMT RNAi cassette as determined by PCR analysis (Figure 3-1C), resulting in a co-transformation efficiency of 88% for the unlinked nptII and COMT minimal cassettes.

To determine copy number of the COMT RNAi cassette in the transformants, genomic DNA was analyzed by Southern blot analysis (Figure 3-1D). A unique transgene integration pattern was detected for each transgenic line, confirming these were independent transformation events. Among the 12 transgenic lines analyzed, two lines (T4 and T18) displayed a relatively simple transgene integration pattern with two
copies. Four lines had less than eight hybridization signals, and six lines showed a complex integration pattern with more than eight copies.

Small-RNA northern blot analysis was performed for 17 transgenic plants to examine whether siRNA was produced from the dsRNA precursor generated by the inverted repeats of the COMT fragment. Seven of the 17 transgenic plants (41%) produced COMT siRNA, while siRNA was not detected in the wild-type plant (Figure 3-2A). The size marker indicated that siRNA from the transgenic plants was ~21 nt long, whereas another siRNA class of ~24 nt was not detected in the transgenic plants (Figure 3-2A).

Clones of primary transformants were vegetatively propagated and grown in the greenhouse, and quantitative real-time RT-PCR of the transcripts was performed in order to investigate siRNA induced suppression of COMT. The expression level was significantly reduced by 67, 97, and 97% in lines T41, T23, and T4, respectively, compared to the wild-type controls (Figure 3-2B).

**Lignin Content, Composition, and Enzymatic Saccharification**

Total lignin content and subunit composition were determined to examine the effect of COMT suppression on lignin biosynthesis in the vegetative progenies of the transgenic plants. Compared to wild-type, total lignin was reduced by 3.9, 8.4, and 13.7% in lines T41, T23, and T4, respectively (Table 3-1). Lignin of transgenic plants was composed of significantly less S unit and similar amount of G unit compared to wild-type (Table 3-1). Transgenic lines had lower S/G ratios ranging from 1.27 to 0.79, while that of wild-type was 1.47. There was no difference in total lignin content or lignin monomer composition of the transgenic control plants transformed with nptII compared to wild-type plants (Table 3-1).
To obtain an estimate of recalcitrance to enzymatic saccharification, enzymatic hydrolysis of lignocellulosic biomass from all samples was performed, with and without dilute acid pretreatments. Transgenic plants with reduced total lignin content and reduced S/G ratios, showed improvements in enzymatic digestibility of the biomass compared to the wild-type controls (Table 3-1). Without pretreatment, the biomass generated from transgenic line T4 yielded 41% more glucose than the wild-type plants, while T41 and T23 showed similar glucose yields as the wild-type plant. With dilute acid pretreatment, significant increases were observed for all of the lines T41, T23, and T4, which yielded 25%, 26%, and 51% more glucose, respectively. Pretreatment enhanced the saccharification efficiencies of the transgenics 2.1- to 2.5-fold, compared to 1.9-fold for wild-type. No significant differences were detected in glucose yields between the nptII-only and wild-type controls.

**Plant Phenotype and Growth**

Clonally propagated progeny of transgenic sugarcane displayed phenotypes similar to wild-type under greenhouse conditions, without lodging or excessive tillering (Figure 3-3A). Microscopic evaluation suggested that vascular bundle tissues and sclerenchyma fiber cells were intact in transgenic sugarcane with reduced total lignin content (Figure 3-4). In contrast to wild-type or nptII-only control plants, transverse stem sections in the internode region of the transgenic lines revealed a deep brown color (Figure 3-3C). The intensity of brown color appeared to be correlated to the level of lignin reduction. Transgenic line T4, with the highest reduction in total lignin content, displayed the darkest brown color of all transgenic lines extending all the way from the basal node to the 4th node below the apical meristem. T41 and T23 only showed brown
color in the basal internodes. The midrib of leaves from transgenic plants did not display brown coloration (data not shown).

Microscopic evaluation revealed a reddish brown color of vascular tissues and surrounding sclerenchyma cells of stem tissues from transgenic sugarcane (Figure 3-4, A and B). Following histochemical analysis with Mäule reagent vascular bundles displayed a red color in both wild-type and transgenic sugarcane. However, the sclerenchyma fiber cells surrounding the vascular bundles displayed a yellow color in transgenic sugarcane, indicating a reduction of S units of lignin (Figure 3-4, C and D). Staining with Wiesner reagent produced similar results in both wild-type and transgenic sugarcane, indicating no changes in hydroxycinnamaldehyde end groups of lignin (Figure 3-4, E and F).

The effect of COMT suppression on plant growth was investigated using clonally propagated progenies grown in randomized, complete blocks and analyzed by ANOVA. Biomass production and stalk diameters of transgenic lines T41 and T23, with moderate reductions in lignin (3.9 and 8.4%, respectively), were similar to those of the non-transgenic tissue culture control plants (TC1) and nptII-only transgenic control plants (TC2) (Figures 3-5, A and B). Transgenic line T4, with a 13.7% reduction in lignin, produced less biomass and had thinner stalks compared with the controls. The tissue culture control (TC1) and the nptII-only transgenic control (TC2) plants produced 12% and 10% less biomass than the wild-type plants, respectively. Transgenic lines T41, T23, and T4 accumulated 17%, 13%, and 35% less biomass than the wild-type plants, respectively. Since the controls and transgenic sugarcane exhibited similar or slightly taller stalk lengths compared with the wild-type sugarcane, the reduced biomass
production was mostly due to decreased stalk diameters (Figure 3-5B). Transgenic plant T4 with the most severe lignin reduction flowered in the greenhouse under natural photoperiod and did not display lodging (Figure 3-3B).

**Discussion**

Genetic improvement of sugarcane through breeding or biotechnology is challenging due to the highly polyploid genome of this interspecific hybrid. To our knowledge this is the first report of sugarcane improvement via an RNAi approach. Transgenic sugarcane plants with RNAi suppression of *COMT* had significantly lower lignin contents, ranging from 3.9 - 13.7% reduction relative to wild-type. Yields of fermentable glucose were significantly increased up to 41% even without pretreatment and up to 51% with dilute-acid pretreatment. These observations are consistent with what was observed for transgenic switchgrass in which *COMT* had been down-regulated, leading to significantly reduced total lignin content by 11.4 - 13.4% and improved saccharification efficiencies by 16.5 - 21.5% following mild pretreatment (Fu et al., 2011a), as well as maize (Vermerris et al., 2007) and sorghum (Dien et al., 2009; Saballos et al., 2008) *brown midrib* mutants with reduced COMT activity. The reduction in lignin content combined with the lower S/G ratio in the lignin of these plants results in more efficient enzymatic saccharification, possibly due to altered physico-chemical properties of the lignin, so that a smaller proportion of the cellulases irreversible adsorb on the lignin. The brown coloration of the stems is also consistent with the phenotype observed in other plants with reduced COMT activity (Fu et al., 2011a; Piquemal et al., 2002). The lack of brown midribs in the transgenic sugarcane plants is consistent with the findings in transgenic switchgrass with RNAi suppression of COMT (Fu et al., 2011a).
Lignin plays an important role in plant growth and development and serves to
protect plants from abiotic and biotic stress (Boerjan et al., 2003; Dixon and Paiva,
1995). Therefore, reducing lignin content could compromise plant performance, stress
tolerance, or defense mechanisms. Our results indicate that in vitro generated
transgenic sugarcane lines, with a moderate reduction in total lignin ranging from 3.9%
to 8.4%, showed comparable biomass production to the tissue culture or transgenic
control plants. Similarly, partial suppression of COMT through RNAi or antisense
strategies in transgenic switchgrass and maize resulted in normal phenotypes under
controlled environment conditions (Fu et al., 2011a; Piquemal et al., 2002). Suppression
of COMT may adversely affect plant growth depending on the impact on lignin content
and the genetic background. Transgenic sugarcane line T4, with 97% reduction of
COMT transcripts and 13.7% reduction in total lignin content, displayed significantly
reduced stalk diameter and biomass production compared with the control and wild-type
plants. Reduced COMT activity in brown midrib mutants of maize (bm3) and sorghum
(bmr12) resulted in ~10% reduction in stover and dry matter yields, respectively (Miller
et al., 1983; Oliver et al., 2005a). A comparison of bmr12 near-isogenic lines (NILs) in
different inbred backgrounds showed decreases in dry matter yields ranging from 6 -
22%. However, the bmr12 NIL in the genetic background of Early Hegari-Sart showed
the greatest reduction in acid detergent lignin, but was able to produce the same
amount of dry matter compared to its wild-type counterpart (Oliver et al., 2005a). This
suggested that there is an interaction between the genetic background and tolerance to
reduced lignin content for biomass production. This was corroborated by analyses of
soluble and cell wall-bound aromatics in sorghum bmr6, bmr12 and bmr6-bmr12.
mutants in different genetic backgrounds (Palmer et al., 2008). Transgenic sugarcane line T4, with the most severe lignin reduction, was still able to mature and produce flowers. Line T4 displays a simple transgene integration pattern with two copies of the COMT RNAi cassette and CP 88-1762 is a fertile cultivar. Transferring the reduced-lignin trait into energycane, a high-fiber, high-biomass variant of sugarcane (Tew and Cobill, 2008) grown primarily for biomass production is therefore expected to be feasible.

While not affecting lignin content or composition, reduced stalk diameters and biomass production were observed in tissue culture and transgenic control plants when compared with wild-type CP 88-1762. Somaclonal variation resulting in less than optimal field performance has been reported for sugarcane (Gilbert et al., 2005; Taparia et al., 2012a). A single backcross has the potential to eliminate tissue culture derived mutations (Bregitzer et al., 2008).

Sugarcane has a high level of genetic redundancy with an average of 12 homo(eo)logous haplotypes (Le Cunff et al., 2008) and most of these homo(eo)logs are considered to be functional (Garsmeur et al., 2011). Because of this functional redundancy, identification of mutant sugarcane plants with substantial changes in lignin content and/or lignin subunit composition is highly unlikely. RNAi-mediated gene silencing allows simultaneous suppression of homo(eo)logs in a high ploidy genome within members of a gene family (Lawrence and Pikaard, 2003; Miki et al., 2005). In this study, the hairpin structure of COMT transgene successfully triggered generation of siRNA and induced suppression of the targeted endogenous COMT gene expression. The reduction in total lignin content and altered S/G ratios suggest that the conserved
sequence that was used to suppress COMT may have supported co-suppression of related COMT homo(eo)logous genes.

RNAi-mediated gene suppression is a useful tool for elucidation of gene function (Miki et al., 2005; Travella et al., 2006). This study confirms that RNAi is an effective method for suppression of target genes in sugarcane. This finding is consistent with an earlier report on RNAi suppression of the phytoene desaturase in sugarcane (Osabe et al., 2009). The sugarcane COMT gene was previously annotated and characterized based on sequence identity, transcript expression pattern, and tissue/cellular localization (Ruelland et al., 2003; Selman-Housein et al., 1999). However, COMT belongs to a large S-adenosyl-L-methionine (SAM) dependent O-methyltransferase (OMT) family including lignin-specific OMT and other phenylpropanoid-specific OMT genes. Therefore, additional evidence for confirming the correct annotation was needed (Noel et al., 2003; Zhou et al., 2010a). The transgenic evidence provided by this study verifies the function of COMT participating in lignin biosynthesis, particularly in S-unit formation. Transgenic sugarcane plants displayed a significant reduction of S units without changing quantities of G units, thereby causing a lower S/G ratio compared to wild-type. Similarly RNAi suppression of COMT in switchgrass also resulted in reduced S/G ratio (Fu et al., 2011a).

Reducing the recalcitrance of lignocellulosic sugarcane biomass to enzymatic hydrolysis is expected to enhance the value of this prime biofuel feedstock. Follow-up field evaluations will elucidate the influence of the genetic background of the transgenic sugarcane lines on biofuel yield per land area.
Table 3-1. Lignin content, composition, and glucose yields after enzymatic saccharification.

<table>
<thead>
<tr>
<th>Lines a)</th>
<th>Total lignin b) (mg g⁻¹)</th>
<th>G unit c) (µmol g⁻¹ lignin)</th>
<th>S unit d)</th>
<th>S/G molar ratio</th>
<th>Glucose yield (mg g⁻¹ biomass)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pretreatment</td>
</tr>
<tr>
<td>WT</td>
<td>181.4 ± 2.2</td>
<td>158.4 ± 2.7</td>
<td>233.4 ± 2.0</td>
<td>1.47</td>
<td>190.9 ± 4.9</td>
</tr>
<tr>
<td>TC2</td>
<td>182.0 ± 2.0</td>
<td>154.0 ± 0.9</td>
<td>230.0 ± 7.9</td>
<td>1.49</td>
<td>196.9 ± 6.6</td>
</tr>
<tr>
<td>T41</td>
<td>174.3 ± 4.6*</td>
<td>150.4 ± 1.6</td>
<td>191.6 ± 7.6*</td>
<td>1.27*</td>
<td>238.7 ± 3.5*</td>
</tr>
<tr>
<td>T23</td>
<td>166.1 ± 1.1*</td>
<td>163.6 ± 4.7</td>
<td>179.8 ± 5.0*</td>
<td>1.10*</td>
<td>241.3 ± 1.8*</td>
</tr>
<tr>
<td>T4</td>
<td>156.6 ± 1.9*</td>
<td>165.8 ± 6.4</td>
<td>131.5 ± 2.3*</td>
<td>0.79*</td>
<td>288.0 ± 1.2*</td>
</tr>
</tbody>
</table>

a) Lines included WT: Wild-type sugarcane; TC2: Transgenic control harboring nptII gene alone; T41, T23, and T4: Transgenic sugarcane.
b) Total lignin content was analyzed using the acetyl bromide (AcBr) method.
c) G: Guaiacyl subunit.
d) S: Syringyl subunit.

Values are means ± standard errors of the mean (n=3 for total lignin, n=2 for lignin composition, and n=3 for glucose yields).

*Significantly different from the wild-type plants at P < 0.05 in t-test.
Figure 3-1. Generation and selection of transgenic sugarcane lines and investigation of transgene integration pattern in the transgenic sugarcane lines. A) COMT RNAi cassette consists of inverted repeats of the 346 bp COMT gene fragment separated by 94 bp Paspalum notatum 4CL intron under control of 1994 bp Oryza sativa C4H promoter and CaMV 35S 3’UTR. PF and PR indicate the primer binding site for PCR analysis to confirm transgenic events. The probe indicates the complementary binding site for Southern blot analysis. The probe indicates the complementary binding site for Southern blot analysis. B) NPTII ELISA assay. Wells A1, B1, and C1 were negative control without protein, wild-type, and positive control with NPTII standard, respectively. The rest of wells were transgenic lines. Yellow coloration indicated NPTII expression in transgenic plants, and no coloration indicated little or no expression of NPTII. C) PCR analysis of COMT RNAi cassette integration. 550 bp PCR products indicated the presence of COMT RNAi cassette in the transgenic sugarcane. T1 ~ T21: Analyzed putative transgenic sugarcane. NC: Negative control without DNA template. PC: COMT RNAi cassette. WT: Wild-type sugarcane. M: 2 kb DNA ladder. D) Southern blot analysis of transgenic sugarcane plants. PC: 9.7 kb linearized COMT RNAi vector; WT: Wild-type sugarcane; T1-T47: Transgenic sugarcane. The displayed image was generated from a single Southern blot. After exposure to x-ray film for different periods of time lanes with best resolution were merged to minimize overexposure of individual signals.
Figure 3-2. Detection of COMT siRNA and evaluation of COMT expression levels in transgenic sugarcane. A) Small-RNA northern blot. The hybridization signal indicates the generation of ~21 nt long COMT siRNA by cleavage of the hairpin structure dsRNA. Ethidium bromide stained 5S rRNA and tRNA were used as a loading control to verify the amount and integrity of the small RNA. WT1 and WT2: Wild-type sugarcane plants; NC: buffer; SM: Size marker corresponding to 21 nt long RNA; T-: Transgenic sugarcane plants; PC: T17 transgenic plant used as a positive control on a separate membrane. B) Quantitative RT-PCR analysis of COMT expression level in the transgenic sugarcane. Error bars represent the 95% confidence intervals of the $2^{-\Delta\Delta Ct}$ value of the wild-type vs. transgenic plants. Significant differences in COMT expression between the wild-type and the transgenic sugarcane are indicated by an asterisk ($n=3$, $P<0.05$ in t-test).
Figure 3-3. Phenotypic evaluation of transgenic sugarcane in comparison to wild-type. WT: Wild-type sugarcane; TC1: non-transgenic plant co-regenerated with transgenic plants from tissue culture; TC2: Transgenic control plant harboring nptII alone; T41, T23, and T4: Transgenic sugarcane plants. A) Greenhouse-grown wild-type (WT) and transgenic sugarcane plants (T4-T41). B) Flowering of the transgenic sugarcane plant (T4) under greenhouse conditions with natural photoperiod. C) Coloration of internodes of transgenic and non-transgenic plants. The basal internode of the transgenic, control, and wild-type plant were transversely or longitudinally sectioned, and the picture was taken immediately without fixation and staining. Scale bar indicates 2 cm.
Figure 3-4. Microscopic and histochemical evaluation of transgenic sugarcane in comparison to wild-type. Microscopic observations of transverse sections of the internode, without staining (A and B), stained with Mäule reagent (C and D), or stained with Wiesner reagent (E and F). Note the differences in color of vascular tissues between wild-type (A) and T4 transgenic line (B) in the absence of staining and arrows indicating yellow coloration in the sclerenchyma fiber cells in T4 transgenic line (D) compared with that in the wild-type (C). Ph: Phloem; X: Xylem; S: Sclerenchyma; P: Parenchyma. Scale bars indicate 250 µm.
Figure 3-5. Plant growth characteristics under greenhouse conditions. A) Biomass production. Values represent means of 8 clonally propagated plants, grown in an air conditioned greenhouse following randomization. Error bars represent standard error. Means with different letter are significantly different at $P < 0.05$ according to Tukey’s multiple comparisons. Fresh weight was measured from the most mature stalk of 7 month old plants. B) Stalk diameter (grey bar) and length (black bar) from the set of plants described under (A). Stalk diameter was measured in the middle of the stalk, and length was measured from the soil surface to the apical meristem of the stalk. Means of stalk diameter (a, b, c) or means of stalk length (x, y) with different letter are significantly different at $P < 0.05$ according to Tukey’s multiple comparisons.
CHAPTER 4
FIELD EVALUATION OF COMT SUPPRESSED TRANSGENIC SUGARCANE

Introduction

Sugarcane (Saccharum spp. hybrids) is a highly productive, perennial C₄ grass and a major feedstock for global bioethanol and sugar production (Tew and Cobill, 2008). In 2010, more than 1.7 billion tonnes of sugarcane were produced on 23.9 million hectares worldwide, and its production, on a dry weight basis, is higher than any other crop in the world (FAOSTAT; http://faostat3.fao.org/home/index.html#DOWNLOAD). Sugarcane accumulates up to 50% of its dry weight as sucrose in the internodes (Waclawovsky et al., 2010). Sucrose-derived ethanol production from sugarcane has been successfully commercialized in Brazil, with high environmental sustainability and average production costs that are 24% lower than the cost of ethanol production from maize in the USA (Crago et al., 2010; Goldemberg, 2007). Sugarcane stalk harvest followed by sucrose extraction generates a large amount of lignocellulosic residue (green tops, leaf litter, and bagasse) comprising 55% of total aboveground biomass (Somerville et al., 2010; Tew and Cobill, 2008; Vermerris, 2011). Including both sucrose and lignocellulosic biomass from sugarcane as feedstocks for ethanol production will boost the ethanol yield per land area while increasing sustainability and adding environmental benefits (Goldemberg, 2007; Leite et al., 2009; Somerville et al., 2010).

The presence of lignin in the cell wall is a major problem for cellululosic ethanol production because it limits the accessibility of cellulose and hemicellulose and reduces the activity of cellulolytic enzymes (Jørgensen et al., 2007; Mansfield et al., 1999). Therefore, an energy-intensive pretreatment of the biomass is typically required for disruption of the cell wall matrix and degradation of lignin. Such pretreatments may
adversely affect down-stream ethanol production by degrading sugars and generating inhibitory molecules (Alvira et al., 2010; Yang and Wyman, 2008). Therefore, reduction of lignin content or modifications of its structure are attractive targets to enhance the production of cellulosic biofuel.

Lignin is formed from the polymerization of three monolignols, $p$-coumaryl, coniferyl, and sinapyl alcohol. After incorporation of these monolignols into the lignin polymer, they are referred to as $p$-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, respectively (Bonawitz and Chapple, 2010). Grass lignin also contains considerable amounts of the hydroxycinnamic acids $p$-coumarate and ferulate. Ferulate plays a role in interconnecting hemicellulosic cell wall polysaccharides and lignin (Grabber, 2005; Grabber et al., 1996), whereas $p$-coumarate, esterified primarily to the gamma carbon of sinapyl alcohol (Ralph et al., 1994), enhances the incorporation of this monolignol in the growing lignin polymer (Hatfield et al., 2008). Among monolignol biosynthetic genes, caffeic acid $O$-methyltransferase ($COMT$) encodes the enzyme which catalyzes $O$-methylation at the C5 position of 5-hydroxyconiferaldehyde and 5-hydroxyconiferyl alcohol, yielding sinapaldehyde and sinapyl alcohol, respectively (Bout and Vermerris, 2003; Humphreys et al., 1999; Louie et al., 2010; Osakabe et al., 1999). COMT deficiency or suppression in brown-midrib mutants (maize $bm3$ and sorghum $bmr12$) or transgenic plants reduces S units in addition to total lignin content, consistent with the role of COMT in sinapyl alcohol biosynthesis (Barrière et al., 2004; Fu et al., 2011a; Guo et al., 2001; Marita et al., 2003; Palmer et al., 2008).

Down-regulation of lignin biosynthetic gene(s) or the identification of mutants with reduced lignin content is a viable strategy to improve saccharification efficiencies and/or
cellulosic ethanol yields as demonstrated with a number of crops (Chen and Dixon, 2007; Dien et al., 2009; Saballos et al., 2008) including transgenic switchgrass (Fu et al., 2011a; Saathoff et al., 2011; Xu et al., 2011) and sugarcane (Jung et al., 2012).

*COMT*-suppressed transgenic sugarcane lines were previously generated in our laboratory by RNA interference (RNAi), and they showed significant reductions in total lignin and S subunit content (Jung et al., 2012). Fermentable glucose yields were significantly increased in the transgenic sugarcane lines, following enzymatic hydrolysis of the lignocellulosic biomass. A moderate reduction of lignin (up to 8%) did not compromise biomass production under greenhouse conditions (Jung et al., 2012).

To our knowledge, data on field performance of transgenic C₄ grass species with modified lignin content or composition have not been reported. Such data are critically important to evaluate the growth and feedstock performance following production under realistic growing conditions. In this study, cell wall characteristics, saccharification efficiencies, and plant growth performance of *COMT*-suppressed sugarcane lines were evaluated in a field trial under USDA-APHIS permit 11-040-120n.

**Materials and Methods**

**Field Design**

*COMT*-suppressed transgenic sugarcane RNAi lines were generated using the commercially important cultivar CP88-1762, as previously described (Jung et al., 2012). Rootstocks of clonal propagules of the V1 (the first generation of vegetative progeny from transgenic plants), tissue culture control, transgenic control, and the original CP88-1762 (wild-type) plants were transplanted on 30 March 2011 at the University of Florida, Plant Science Research and Education Unit, Citra, Florida, USA under USDA-APHIS permit 11-040-120n. The field design was a randomized block design with three
replications. Four plots of transgenic lines representing four independent transformation events, and three plots of wild-type, non-transgenic controls derived from tissue culture, and the transgenic control, harboring the nptII gene were included in each replicate. Each plot consisted of one row with five clonal propagules planted per plot. Spacing between plants, rows, and blocks was 90, 150, and 450 cm, respectively. Each block was surrounded by one row of wild-type sugarcane plants. Weeds were removed manually during the establishment phase, and Orthene was applied in August 2011 for insect control. Plots were fertilized with 40 kg ha$^{-1}$ N, 15 kg ha$^{-1}$ P, and 60 kg ha$^{-1}$ K at planting and with 70 kg ha$^{-1}$ N, 15 kg ha$^{-1}$ P, and 60 kg ha$^{-1}$ K on 9 May 2011. Plots were irrigated once or twice a week with a rate of 10 mm depending on rainfall and harvested on 26 October 2011 for determination of biomass and compositional analysis.

**Evaluation of Plant Performance**

The fresh weight of the harvested biomass was determined on site immediately after the plot harvest. The number of stalks was counted for each plant, and five stalks were collected for milling. The leaves and leaf sheaths were removed from the stalks, the number of internodes was counted, and stalk height, from the shoot apical meristem to the soil surface, was measured. Stalk diameter was measured at the center of the 8th internode from the base of the plant. Internodes 1-4 below the shoot apical meristem were removed and the remaining portion of the stalk was crushed to extract juice using a custom-made juice extractor. Percentage soluble solid in the extracted juice (°Brix) was measured using a PAL-1 portable refractometer (ATAGO U.S.A., Inc., Bellevue, WA, USA). Crushed stalks were dried at 45°C and stored for further analysis.

Pests and diseases [e.g. sugarcane brown leaf rust (*Puccinia melanocephala*), orange leaf rust (*P. kuehnii*), and pink sugarcane mealybug (*Saccharicoccus sacchari*)]
were monitored in monthly intervals. Leaf rust was rated using a 0-4 scale as described in (Comstock et al., 2010). For evaluating the rate of mealybug infestation, six stalks per plant were randomly selected and the number of infested internodes was recorded. The rate of infestation was calculated by the number of infested internodes over the total number of internodes. The lodging incidence was monitored every month. The lodging susceptibility was scored in every plant using a 0-7 scale with 0 being no lodging (0-20° deviation from erect), and the score increasing with every 10° lodging, until 7 which equated to complete lodging.

**Gene Expression Analysis**

The third internode below the shoot apical meristem was collected from one of the millable stalks in each plot. RNA extraction, cDNA synthesis, and quantitative real-time RT-PCR were performed as previously described (Jung et al., 2012).

**Sample Preparation for the Evaluation of Cell Wall Composition and Saccharification Efficiency**

Crushed stalks, dried at 45°C as described above, were ground using a Wiley mill (Thomas Scientific, Swedesboro, NJ, USA) with a 1.0 mm sieve, and the samples were further passed through a 0.42 mm sieve. Soluble extract was removed from the samples by three successive extractions with 50% ethanol (v/v) and sonication at 45°C for 30 min. Extract-free samples were dried at 45°C until constant weight was achieved in order to analyze lignin, hydroxycinnamic acids, and cell wall polysaccharides, and the evaluation of saccharification efficiencies.

**Lignin Content and Composition**

Total lignin content was measured using the acetyl bromide (AcBr) method as previously described (Jung et al., 2012). Lignin composition was determined by
thioacidolysis (Robinson and Mansfield, 2009). Thioacidolysis derivatives for lignin monomers were analyzed using a Varian 3800 gas chromatograph (GC) coupled to the Varian 1200 mass spectrometer (MS) (Varian, Walnut Creek, CA, USA). One microliter of the sample was injected onto a Factor-4 VF-5ht column (35 m, 0.25 mm i.d.) (Varian) with helium (1.2 ml min⁻¹) as carrier gas. GC conditions were as follows: initial column temperature 130°C, held for 3 min, ramped at 3°C min⁻¹ to 250°C, and held for 5 min; injector temperature, 250°C; and 1:2 split ratio. The mass spectrometer was operated in electron impact mode at 70 eV. The detector was operated at 1.0 kV. Mass spectra were recorded every 0.2 s at a scanning range of m/z 50-550, and the data were recorded and analyzed using Varian MS Workstation software. Thioacidolysis monomers, G and S, were quantified by integration of the respective peak areas, and employing the response factors of G and S against an internal standard, tetracosane (IS) as follows: G versus IS, 0.47; and S versus IS, 0.53 (Yue et al., 2012).

**Analysis of Hydroxycinnamic Acids**

Ester-linked cell wall bound p-coumaric acid (pCA) and ferulic acid (FA) were determined following Hatfield et al. (1999c) with modifications. Briefly, 25 mg of extract-free sample was saponified with 1.7 ml of 2 M NaOH and 20 µl 2-hydroxycinnamic acid (1 mg ml⁻¹ in 2 M NaOH) (Sigma Aldrich, Saint Louis, MO, USA) as an internal standard. The reaction was carried out at room temperature for 20 h in the dark, then acidified with 0.3 ml 12 M HCl. Samples were extracted three times with an equal volume of diethyl ether, and supernatants were combined and dried under a stream of N₂. Derivatization was performed by adding 40 µl methoxyamine hydrochloride (20 mg ml⁻¹ in pyridine) (Sigma Aldrich) and incubating at 37°C for 90 min. The sample was trimethylsilylated by adding 60 µl of MSTFA [N-methyl-N-(trimethylsilyl)
trifluoroacetamide] (Sigma Aldrich), and incubating at 37°C for 30 min. pCA and FA were identified and quantified using GC-MS with a Factor-4 VF-5ht column (35 m, 0.25 mm i.d.) (Varian). One microliter of the sample was injected and helium was used as a carrier gas (1.2 ml min⁻¹). The injector temperature was 220°C; the split ratio was 1:10. The oven temperature was held for 4 min at 70°C, and increased to 250°C at 10°C per min without holding, then ramped at 40°C per min, and held constant for 2 min. MS operating procedure was the same as that for lignin compositional analysis except for a scanning range of m/z 45-650. Peaks from derivatized pCA, FA, and internal standard were identified by characteristic mass spectrum ions obtained from the reference compounds. Concentrations of pCA, FA, and internal standard were determined using the standard curve. The concentrations of pCA and FA were normalized with the concentration of internal standard among the samples.

**Cell Wall Carbohydrates and Starch**

Cell wall carbohydrates in the stalk samples were determined according to the National Renewable Energy Laboratory (NREL) protocol by Sluiter et al. (2008). Three hundred milligram of extract-free sample were hydrolyzed with 72% H₂SO₄ at 30°C for 1 h and then treated with 4% H₂SO₄ at 121°C for 1 h. Liberated monomeric sugars were identified and quantified with an Agilent/HP 1090 HPLC equipped with an RI detector (Agilent Technologies, Santa Clara, CA, USA). The HPLC analysis was carried out using a HPX-87H column (Bio-Rad, Hercules, CA, USA), operating at 50°C with a 0.004 M H₂SO₄ mobile phase at a flow rate of 0.4 ml min⁻¹. Starch content in the sample was determined by the NREL procedure by Ehrman (1996). Glucose yield was measured using a YSI glucose analyzer (YSI Life Science, Yellow Springs, OH, USA) following
enzymatic digestion of the extracted starch, and its value was adjusted by multiplying 0.9 for the weight gained by hydration.

**Dilute-Acid Pretreatment and Enzymatic Hydrolysis**

Extract-free samples possessing 0.1 g of cellulose were pretreated as previously described (Chen and Dixon, 2007). The amount of cellulose was defined as the amount of glucose from cell wall carbohydrate corrected by subtracting the amount of glucose from starch (Selig et al., 2008). Five sets of pretreated samples were prepared per sugarcane line to evaluate saccharification performance at different hydrolysis times and enzyme dosages.

The enzymatic hydrolysis was performed according to the NREL protocol by Selig et al. (2008) with modifications. The pretreated sample was suspended in 8 ml distilled water, and placed in a pre-weighed 50 ml polypropylene tube. Ten milliliter of 0.1 M sodium citrate buffer (pH 4.8) and 200 µl of 2% (w/w) sodium azide (Sigma Aldrich) were added to each tube. The enzymatic hydrolysis was performed by adding the appropriate volume of Kerry Biocellulase W (Kerry Bioscience, Cork, Ireland) and Novozyme 188 β-glucosidase (Sigma Aldrich). One set of the pretreated samples was used as the no-enzyme blank (no addition of cellulolytic enzymes). The appropriate volume of distilled water was added to each tube to bring the total mass of each sample to 20 g, assuming all solutions and the biomass sample have 1.000 g ml⁻¹ specific gravity. The hydrolysis was carried out for 168 h in a shaking incubator at 50°C and 250 rpm. The time course of enzymatic hydrolysis with 60 FPU g⁻¹ cellulose of Kerry Biocellulase W and 64 ρNPGU g⁻¹ cellulose of Novozyme 188 β-glucosidase was monitored by measuring glucose yields at the time points of 0, 3, 6, 24, 48, and 72 h. The effect of enzyme loading on saccharification was evaluated with Kerry Biocellulase
W loadings of 5, 20, or 60 FPU g\(^{-1}\) cellulose. Novozyme 188 β-glucosidase loading was constant as 64 pNPGU g\(^{-1}\) cellulose. One milliliter of the hydrolyzed sample was collected for each treatment and time point, and filtered through a 0.2 µm syringe filter. Glucose yields were analyzed using the YSI glucose analyzer (YSI Life Science). Saccharification efficiency was calculated as the ratio of glucose released following the enzymatic hydrolysis to the amount of glucose present in the cell wall before the hydrolysis. The glucose amount derived from starch in the stalk was subtracted in order to calculate the glucose amount in the cell wall. Glucose yields were adjusted by multiplying 0.9 for the weight gained by hydration.

**Statistical Analysis**

ANOVA was performed using Proc GLM in SAS\textsuperscript{TM} Version 9.3 (SAS Institute Inc., Cary, NC, USA). Statistical significance among the means for plant growth performance was determined using Fisher’s protected LSD test at \(P < 0.05\). Significant differences (\(P < 0.05\)) of means for the content of cell wall components among the samples were determined using Tukey’s test. For the gene expression analysis, the \(t\)-test was used to determine whether the means of \(\Delta\)Ct were significantly different between wild-type and a transgenic line (\(P < 0.05\)).

**Results**

**COMT Gene Suppression in Transgenic Sugarcane Lines**

*COMT* gene suppression through RNAi was evaluated six months after initiation of the field trial. The transcript abundance of *COMT* was significantly reduced by 80, 89, 92, and 91\% in the transgenic lines, T41, T23, T31, and T4, respectively, compared to wild-type (WT) sugarcane (Figure 4-1).
Effects of COMT Suppression on Lignin Content and Composition

Total lignin content in stalk samples of transgenic and control plants grown in the field was determined using the acetyl bromide (AcBr) method. There was no significant difference in total lignin content of the WT, the non-transgenic tissue culture control (NT), or the nptII-transgenic control (TC) (Table 4-1). However, the transgenic lines had 4.5-11.1% lower total lignin content compared to WT. In comparison to the corresponding TC, the level of total lignin reduction was 5.5, 7.5, 11.2, and 12.0% in the transgenic lines T41, T23, T31, and T4, respectively (Table 4-1).

The S subunit content was significantly reduced by 16% and 49% in T41 and T4, respectively, compared to the TC control (Table 4-1). There were no significant differences in G subunit content among the transgenic lines, WT, or the TC line (Table 4-1). Due to a reduction in S subunits, T41 and T4 transgenic lines had significantly lower S/G molar ratios of 1.17 and 0.72, respectively, compared to 1.47 for the TC and 1.48 for the WT.

Effects of COMT Suppression on Cell Wall Carbohydrates and Cell Wall Bound Hydroxycinnamic Acids

The consequences of COMT suppression on other cell wall components, such as cellulose, hemicellulose, and cell wall bound hydroxycinnamic acids, were investigated in the two transgenic lines, T41 and T4, which represented the lines with the least and most reduction in lignin content, respectively, versus the WT and the TC (Table 4-2). The amounts of glucose mostly derived from cellulose did not differ significantly between the transgenic and control plants. The amount of xylose, major hemicellulose component, was significantly increased in T4 compared to T41 or control plants, while
the amount of arabinose was not different among the lines. The total amount of cell wall sugars did not differ significantly between the transgenic and control plants.

The content of cell wall esterified $p$-coumaric acid ($p$CA) and ferulic acid (FA) were evaluated following mild alkaline hydrolysis. The esterified $p$CA content was significantly decreased by 8% and 32% in T41 and T4, respectively, compared to the WT (Table 4-3). Esterified FA did not differ significantly between the transgenic and control plants.

**Effect of Lignin Reduction on Saccharification Efficiency**

Enzymatic hydrolysis following dilute-acid pretreatment was performed to evaluate the bioconversion efficiency of lignocellulosic biomass from field-grown transgenic sugarcane into directly fermentable sugars. The time course of saccharification describes elevated saccharification efficiencies of the transgenic biomass compared to WT and TC during the entire 72 h period of enzymatic hydrolysis (Figure 4-3A). A maximum saccharification rate was reached at 72 h of enzymatic hydrolysis, and saccharification efficiencies for T41 and T4 were 23.2 and 32.4% higher than for the TC control. Cellulose in the transgenic lines was converted to glucose more rapidly compared to WT or the TC control. Transgenic lines T41 and T4 had saccharification efficiencies of 49.2 and 54.8%, respectively, after only 24 h of enzymatic hydrolysis, values which exceeded the 48.4 and 46.7% conversion for WT or TC at 72 h of enzymatic hydrolysis, respectively.

To evaluate the effect of enzyme dosage on the saccharification of transgenic and control sugarcane plants, hydrolysis was performed with different cellulase loadings of 5, 20, and 60 filter paper unit (FPU) per gram of cellulose. All lines showed the highest saccharification efficiencies at the highest cellulase loading (60 FPU) (Figure 4-
3B). However, lignocellulosic biomass from transgenic sugarcane plants was more effectively converted to glucose than that of WT or the TC control, regardless of enzyme dosage. For T41 and T4, saccharification efficiencies at 5 FPU were similar to those at 20 FPU for WT or the TC control. Furthermore, at 20 FPU, saccharification efficiencies of 49.9 and 55.9% in T41 and T4 were significantly higher than those obtained with 60 FPU in the WT or TC control, respectively.

**Growth Performance of the Transgenic Sugarcane Grown under Field Conditions**

Clonally produced rootstocks derived from the vegetative progeny (V1) of transgenic events and control plants established well and were grown for seven months under field conditions in Citra, FL (Figure 4-2, A and B). The transgenic line, T41, with 6% lignin reduction, displayed no significant difference in biomass production compared with the TC and NT lines (Table 4-4). Stalk length and diameter were also not significantly different between T41, TC and NT lines. However, lines of T23, T31 and T4 with 8, 11, and 12% reduction in lignin, respectively, displayed a 21, 64, and 65% reduction in biomass, respectively, compared to TC lines. This reduction in biomass can primarily be attributed to reduced stalk diameter compared to the controls. In comparison to the original CP88-1762 WT plants, both the NT tissue culture controls and the TC control lines had an 18% reduction in biomass.

For the transgenic lines, the number of internodes per stalk was not statistically different from WT or the control lines indicating that the developmental stage was similar among the lines at the time of harvest (Table 4-4). The number of stalks per plant was not different among the lines, except in transgenic line T31, which had fewer stalks per plant (Table 4-4). The amount of soluble solids in the stalk did not differ significantly between the transgenic lines T41 and T23 and control lines. However,
transgenic lines T31 and T4 displayed significantly reduced concentrations of soluble solids in their stalks compared to the control lines (Table 4-4).

Diseases, pests and lodging were monitored monthly. A minor occurrence of orange rust (*Puccinia kuehnii*) was observed during the main growth period and was rated between 0-1 in all of the lines, with no significant differences between the lines. A pink sugarcane mealybug infestation was observed in August and the pest was eliminated by pesticide application after scoring. The rate of mealybug infestation did not significantly differ between the transgenic and control plants, ranging from 10 to 37% of the internodes being infested. The transgenic and control plants did not show any lodging throughout the growing season until an isolated thunderstorm hit the field on 10 October 2011 with a maximum wind speed of 68 km/h. After the storm, lodging was most severe in the part of the field facing the prevailing wind direction during the severe thunderstorm, with no significant difference between transgenic lines and WT or TC controls (Figure 4-2).

**Discussion**

To our knowledge, this is the first report on the field performance of a transgenic C₄ grass with modified cell wall composition. The field-grown *COMT*-suppressed transgenic sugarcane lines showed improvement in bioconversion efficiency of mature lignocellulosic biomass to fermentable glucose. Suppression of *COMT* transcripts ranging from 80% to 91% resulted in a reduction of lignin content by 6% to 12% in different transgenic lines and an improvement of saccharification efficiency by 19% to 32% compared to non-transgenic and transgenic (*nptII*-only) controls. The transgenic lines required one-third of the hydrolysis time and 3- to 4-fold less enzyme to produce an equal or higher amount of glucose than control plants. These findings are consistent
with earlier reports on greenhouse-grown COMT-suppressed transgenic switchgrass (Fu et al., 2011a) or sugarcane (Jung et al., 2012). COMT-suppressed switchgrass increased ethanol production by 38% or enabled reduced pretreatments and 3- to 4-fold less enzyme loading to produce an equal amount of ethanol compared to the control. Elevated biofuel production from lignocellulosic sugarcane biomass will boost the sucrose-derived biofuel yields for this prime biofuel crop. Accelerated rates of bioconversion, less severe pretreatment conditions and/or lower enzyme loadings to achieve yields of fermentable sugars equal to or higher than with wild-type biomass are expected to significantly reduce biofuel production costs.

The RNAi-induced level of COMT suppression in the different transgenic lines that was observed earlier under greenhouse conditions (Jung et al., 2012) was stably maintained under field conditions. Reductions of total lignin and S subunit content in the field-grown transgenic lines were consistent with those previously reported for greenhouse-grown plants. Suppression of COMT in transgenic sugarcane specifically resulted in a significant reduction of S subunit content without affecting G subunit content, similar to COMT-deficient maize brown midrib (bm3) and sorghum bmr12 mutants (Barrière et al., 2004; Palmer et al., 2008).

The transgenic sugarcane lines displayed reduced p-coumaric acid (pCA) levels. pCA is primarily esterified to S lignin subunits in grasses (Grabber et al., 1996; Ralph et al., 1994). The reduced pCA content most likely resulted from a reduction of S subunits in the COMT-suppressed transgenic sugarcane plants, similar to maize bm3 and sorghum bmr12 mutants (Marita et al., 2003; Palmer et al., 2008; Piquemal et al., 2002), and COMT-suppressed transgenic maize (Piquemal et al., 2002). Ferulic acid
(FA) cross-linking of the cell wall in grass species negatively affects enzymatic hydrolysis (Grabber et al., 1998). FA is esterified to arabinoxylans, and xylans are interconnected by radical coupling of esterified FA into FA dimers or trimers. Furthermore, esterified FA is incorporated into lignin, forming an extensive lignin-ferulate-polysaccharide complex (Grabber, 2005; Ralph, 2010). Elevated levels of FA were not observed in the sugarcane lines with COMT suppression.

In this study, there was no difference in the amount of glucose in the transgenic sugarcane lines analyzed, while xylose was significantly increased in the transgenic line T4 that displayed the greatest lignin reduction. Similarly, xylose amounts in the stem of COMT-suppressed transgenic switchgrass were increased in both T0 and T1 plants (Fu et al., 2011a). However, it is unclear whether there is a compensatory increase in cell wall carbohydrates for lignin reduction as proposed in 4-coumarate:CoA ligase (4CL)-suppressed transgenic poplar (Hu et al., 1999). The maize bm3 mutant, cinnamyl alcohol dehydrogenase (CAD)- or 4CL-suppressed transgenic switchgrass did not exhibit an increase in cell wall carbohydrates (Marita et al., 2003; Saathoff et al., 2011; Xu et al., 2011). Furthermore, no increase in the amount of cellulose or the transcription level of cellulose biosynthetic genes was observed in a set of Arabidopsis lignin mutants (Vanholme et al., 2012).

Plant biomass yield is one of the most important factors for determining the economic viability of a lignocellulosic feedstock. Despite the beneficial effects of reducing lignin for ethanol conversion, blocking the monolignol biosynthetic pathway may be associated with impaired growth and development of the transgenic plants. Particularly, suppression of hydroxycinnamoyl CoA shikimate/quinate hydroxycinnamoyl
transferase (HCT), \( \rho \)-coumarate 3'-hydroxylase (C3'H), or cinnamoyl CoA reductase (CCR) frequently leads to developmental arrest, dwarfism, collapsed xylem vessels, and/or abnormal flowering (Gallego-Giraldo et al., 2011; Goujon et al., 2003; Piquemal et al., 1998; Ralph et al., 2006; Shadle et al., 2007; Srinivasa Reddy et al., 2005). These negative effects may result from disruption of non-lignin metabolite biosynthesis, such as coniferaldehyde derivatives and/or shikimate derivatives, which influence cell growth and defense mechanisms (Bonawitz and Chapple, 2010; Gallego-Giraldo et al., 2011). Interestingly, COMT-suppressed plants were reported to display normal plant growth in a variety of species including tobacco, alfalfa, maize, and switchgrass under greenhouse conditions, and poplar under field conditions (Chen and Dixon, 2007; Fu et al., 2011a; Pilate et al., 2002; Pinçon et al., 2001b; Piquemal et al., 2002). However, complete knock-out of COMT in maize and sorghum brown midrib mutants was generally associated with lower biomass yields (Lee and Brewbaker, 1984; Miller et al., 1983; Oliver et al., 2005a; Oliver et al., 2005b; Sattler et al., 2010). In our results, 80% suppression of COMT and 6% reduction of lignin did not cause significant reduction of biomass production or related traits, such as stem diameter, height, tillering, and accumulated sugars compared to non-transgenic tissue culture-derived control plants. However, COMT suppression of 91% and a 12% reduction in lignin compromised biomass production under both greenhouse (Jung et al., 2012) and field conditions. These results indicate that adverse effects on plant growth can be avoided by targeted suppression of specific lignin biosynthetic gene(s), which do not cause pleiotropic effects, and by determining the level of lignin modification that allows high biomass yield
along with improved conversion performance. This strategy is facilitated by the range of 
gene expression/suppression that is typically found among different transgenic events.

The tolerance to reduced lignin may also differ among different species, and 
even, within a species. In brown midrib mutants, agronomic traits in the mutant cultivar or hybrid are influenced by the interaction between the specific gene and the genetic background (Pedersen et al., 2005; Sattler et al., 2010). For example, a sorghum 
COMT-deficient bmr12 near-isogenic line in the genetic background of Early Hegari-
Sart exhibited similar plant height and dry matter yield as the wild-type counterpart, unlike what was observed for the comparison in the background of Atlas, Kansas Collier, or Rox Orange (Oliver et al., 2005a).

Although line T4 produced the least amount of biomass, it displayed the highest 
saccharification efficiency and the lowest lignin content among the transgenic 
sugarcane lines. This would suggest its use as a parent in future breeding efforts to 
explore if further improvements can be achieved by crossing COMT-down-regulated 
sugarcane lines (e.g. T4) to genetically diverse accessions (e.g. high biomass type 
energycanes). Crossing would also eliminate tissue culture-derived mutations (Bregitzer 
et al., 2008). Among the transgenic sugarcane lines evaluated under both greenhouse 
and field conditions, T41, has the greatest promise for crop improvement since its field 
performance was comparable to control lines, but with a 19 to 23% increase in 
saccharification efficiency. Improving the saccharification efficiency of lignocellulosic 
sugarcane biomass by modifying lignin biosynthesis will greatly benefit the biofuels 
industry.
Table 4-1. Lignin content and composition in transgenic sugarcane.

<table>
<thead>
<tr>
<th>Lines(^a)</th>
<th>AcBr lignin(^b) (mg g(^{-1}) DW)</th>
<th>G units(^c) (µmol g(^{-1}) AcBr lignin)</th>
<th>S units(^d) (µmol g(^{-1}) AcBr lignin)</th>
<th>S/G molar ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>192.2 ± 1.9(^a)</td>
<td>97.4 ± 5.4(^a)</td>
<td>143.9 ± 3.5(^a)</td>
<td>1.48(^a)</td>
</tr>
<tr>
<td>NT</td>
<td>193.3 ± 2.1(^a)</td>
<td>na(^f)</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>TC</td>
<td>194.3 ± 1.7(^a)</td>
<td>102.4 ± 4.0(^a)</td>
<td>150.0 ± 4.8(^a)</td>
<td>1.47(^a)</td>
</tr>
<tr>
<td>T41</td>
<td>183.6 ± 0.7(^b)</td>
<td>106.8 ± 1.6(^a)</td>
<td>125.4 ± 2.6(^b)</td>
<td>1.17(^b)</td>
</tr>
<tr>
<td>T23</td>
<td>179.8 ± 1.2(^b)</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>T31</td>
<td>172.5 ± 0.8(^c)</td>
<td>na</td>
<td>na</td>
<td>na</td>
</tr>
<tr>
<td>T4</td>
<td>170.9 ± 1.3(^c)</td>
<td>107.2 ± 5.7(^a)</td>
<td>77.1 ± 1.0(^c)</td>
<td>0.72(^c)</td>
</tr>
</tbody>
</table>

\(^a\) Lines included WT: Wild-type sugarcane, NT: Non-transgenic control derived from tissue culture, TC: Transgenic control harboring nptII gene, and T41, T23, T31, and T4: Transgenic sugarcane lines.
\(^b\) Total lignin content was analyzed using the acetyl bromide (AcBr) method.
\(^c\) G: Guaiacyl subunit.
\(^d\) S: Syringyl subunit.
\(^e\) DW: Dry weight.
\(^f\) na: not analyzed.

Values are means ± standard errors of the mean; different letters within the same column indicate significant differences among means (\(n=3\), \(P < 0.05\)) as determined by Tukey’s test.
Table 4-2. Composition of structural carbohydrates of the cell wall in control and transgenic sugarcane plants.

<table>
<thead>
<tr>
<th>Linesa)</th>
<th>Glucose</th>
<th>Xylose</th>
<th>Arabinose</th>
<th>Total sugar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mg g⁻¹ DWb)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>462.0 ± 17.5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>234.9 ± 6.2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.7 ± 0.3&lt;sup&gt;a&lt;/sup&gt;</td>
<td>704.6 ± 23.6&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>TC</td>
<td>446.1 ± 2.6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>228.6 ± 1.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>7.5 ± 1.0&lt;sup&gt;a&lt;/sup&gt;</td>
<td>682.3 ± 4.5&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>T41</td>
<td>429.0 ± 2.7&lt;sup&gt;a&lt;/sup&gt;</td>
<td>223.6 ± 0.6&lt;sup&gt;b&lt;/sup&gt;</td>
<td>9.5 ± 0.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>662.1 ± 3.1&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>T4</td>
<td>441.8 ± 5.0&lt;sup&gt;a&lt;/sup&gt;</td>
<td>250.2 ± 0.8&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9.2 ± 0.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>701.2 ± 5.4&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Lines included WT: Wild-type sugarcane, TC: Transgenic control harboring nptII gene, and T41 and T4: Transgenic sugarcane lines.

<sup>b</sup> DW: Dry weight.

Values were means ± standard errors of the mean; different letters in the same column indicated significant differences among means (n=3, P < 0.05) as determined by Tukey’s test.
Table 4-3. Recovery yields of cell wall bound $p$-coumaric acid ($p$CA) and ferulic acid (FA) after mild-alkaline hydrolysis of the control and transgenic sugarcane.

<table>
<thead>
<tr>
<th>Lines$^a)$</th>
<th>$p$CA</th>
<th>FA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mg g$^{-1}$ DW$^b)$</td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>14.4 ± 0.3$^a$</td>
<td>2.2 ± 0.2$^a$</td>
</tr>
<tr>
<td>TC</td>
<td>14.5 ± 0.3$^a$</td>
<td>2.3 ± 0.1$^a$</td>
</tr>
<tr>
<td>T41</td>
<td>13.3 ± 0.3$^b$</td>
<td>2.4 ± 0.2$^a$</td>
</tr>
<tr>
<td>T4</td>
<td>9.9 ± 0.1$^c$</td>
<td>2.7 ± 0.1$^a$</td>
</tr>
</tbody>
</table>

$^a)$ Lines included WT: Wild-type sugarcane, TC: Transgenic control harboring nptII gene, and T41 and T4: Transgenic sugarcane lines.

$^b)$ DW: Dry weight.

Values were means ± standard errors of the mean; different letters in the same column indicated significant differences among means ($n=3$, $P < 0.05$) as determined by Tukey’s test.
Table 4-4. Growth characteristics of the transgenic sugarcane lines under field conditions.

<table>
<thead>
<tr>
<th>Linesa)</th>
<th>No. of Internodes</th>
<th>Biomass (kg plant(^{-1}))</th>
<th>Stalk height (cm)</th>
<th>Stalk diameter (mm)</th>
<th>No. of Stalks</th>
<th>Soluble solids (% Brix)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>18.5(^a)</td>
<td>18.5(^a)</td>
<td>204(^a)</td>
<td>28.2(^a)</td>
<td>20.5(^a)</td>
<td>20.0(^a)</td>
</tr>
<tr>
<td>NT</td>
<td>17.6(^a)</td>
<td>15.2(^b)</td>
<td>180(^bc)</td>
<td>27.5(^a)</td>
<td>18.7(^a)</td>
<td>19.9(^a)</td>
</tr>
<tr>
<td>TC</td>
<td>18.3(^a)</td>
<td>15.7(^b)</td>
<td>185(^b)</td>
<td>26.5(^ab)</td>
<td>19.8(^a)</td>
<td>20.4(^a)</td>
</tr>
<tr>
<td>T41</td>
<td>18.2(^a)</td>
<td>15.7(^b)</td>
<td>182(^b)</td>
<td>25.1(^bc)</td>
<td>22.8(^a)</td>
<td>20.3(^a)</td>
</tr>
<tr>
<td>T23</td>
<td>18.9(^a)</td>
<td>12.4(^c)</td>
<td>191(^ab)</td>
<td>23.7(^cd)</td>
<td>23.0(^a)</td>
<td>20.1(^a)</td>
</tr>
<tr>
<td>T31</td>
<td>17.9(^a)</td>
<td>5.5(^d)</td>
<td>161(^dc)</td>
<td>21.6(^de)</td>
<td>14.0(^b)</td>
<td>17.5(^b)</td>
</tr>
<tr>
<td>T4</td>
<td>18.7(^a)</td>
<td>5.3(^d)</td>
<td>155(^d)</td>
<td>19.8(^e)</td>
<td>19.0(^a)</td>
<td>17.2(^b)</td>
</tr>
</tbody>
</table>

a) Lines included WT: Wild-type sugarcane, NT: Non-transgenic control derived from tissue culture, TC: Transgenic control harboring nptII gene, and T41, T23, T31, and T4: Transgenic sugarcane lines.

Values are means. Different letters in the same column indicate significant differences among means (\(n=3\), \(P < 0.05\)) as determined by Fisher’s Least Significant Differences.
Figure 4-1. Real time RT-PCR analysis of COMT expression level in transgenic sugarcane. WT: wild-type sugarcane; T41, T23, T31, and T4: Transgenic sugarcane lines. Error bars represent the 95% confidence intervals of the $2^{-\Delta\Delta C_t}$ value ($n=3$).
Figure 4-2. Field trial of the *COMT* suppressed transgenic sugarcane lines. A) The field trial of transgenic sugarcane 6 months after establishment October 1<sup>st</sup> 2011. B) Sugarcane growth in block 1 at the time of harvest. October 26<sup>th</sup> 2011. WT: wild-type sugarcane; TC; transgenic control harboring *nptII*; T41, T23, T31, and T4: Transgenic sugarcane lines. Border plants were removed before harvest.
Figure 4-3. Enzymatic saccharification performance in wild-type (WT), transgenic control harboring nptII gene alone (TC), and transgenic sugarcane lines (T41 and T4). A) Time-course of enzymatic hydrolysis of ground extract-free stalk sample with 60 FPU g\(^{-1}\) cellulose following dilute acid pretreatment. B) Saccharification efficiencies at different cellulase dosages (5, 20, and 60 FPU g\(^{-1}\) cellulose). Error bars represent standard errors of the mean (n=3).
CHAPTER 5
CONCLUSIONS

Lignin reduced transgenic sugarcane has been successfully developed by suppressing caffeic acid O-methyltransferase (COMT). The suppression level of the gene, cell wall characteristics, enzymatic saccharification efficiency, and agronomic performance of transgenic sugarcane plants were evaluated under both greenhouse and replicated field conditions.

Despite the high level of ploidy and genetic redundancy in the sugarcane genome, transgene induced RNAi targeting the highly conserved sequences among the genetically redundant COMT genes, effectively suppressed the expression of COMT by 80-91%.

Corresponding to the level of COMT suppression, total lignin content was reduced by 6-12% in different transgenic lines. The suppression of COMT also altered lignin composition. Lignin in the transgenics contained 49% fewer S units without changing the G unit content compared to that of control plants. The reduced S unit content in the lignin polymer subsequently resulted in a reduced level of p-coumarate incorporation into lignin. No impact of COMT suppression was observed on the accumulation of total cell wall carbohydrates.

A reduction in total lignin content was associated with diminished recalcitrance of lignocellulosic biomass to enzymatic hydrolysis. The saccharification efficiency of lignin-reduced transgenic lines was 32% higher than in control lines. Cellulose in the transgenic lignocellulosic material was converted to sugars more effectively at any given hydrolyzing time and enzyme dosage.
A moderate reduction of total lignin by up to 6% had negligible impacts on biomass yield, plant height, stalk diameter, tillering ability, and soluble solid accumulation compared with control lines under both greenhouse and field conditions. However, 8-12% lignin reduction had negative effects on plant growth performance under field conditions.

Future research should focus on the transfer of these traits for reduced lignin and biomass recalcitrance into other high biomass yielding sugarcane or energycane. Given the association between the genetic background and tolerance to reduced lignin, crosses with transgenic sugarcane from diverse genetic backgrounds will further improve biomass production under optimal conditions. It is also expected that somaclonal variation observed in the transgenic sugarcane will be eliminated during the (back) crossing procedure. Finally, it will be critical to investigate the ability of these transgenic sugarcane lines with reduced lignin to maintain agronomic fitness after extensive field testing across multiple years and through successive vegetative progenies.

In conclusion, the improvement of feedstock quality has been an important task for efficient and cost-competitive production of biofuel. Sugarcane can be considered a superior feedstock compared to conventional starch feedstocks or other dedicated biofuel crops. The economic feasibility and productivity of sugarcane ethanol can be significantly increased by the co-utilization of readily extractable sugar and abundant lignocellulosic biomass. Reducing lignin content by suppressing lignin biosynthetic genes has been established as a straightforward strategy to reduce biomass
recalcitrance, thereby improving the bioconversion efficiency of lignocellulosic sugarcane biomass.
Generation of Transgenic Sugarcane using Particle Bombardment

Callus Induction from Immature Leaf Whorls

1. Harvest sugarcane tops with two or three internodes below the shoot apical meristem.
2. Wipe the outermost leaf sheath with 70% ethanol, and remove a few additional layers of outer leaves under aseptic conditions.
3. Transversely cut immature leaf whorl in 2-5 mm thick sections from the region above shoot apical meristem.
4. Place explants on a callus induction media (CI-3; see media preparation below) and subculture to the new medium of the same composition biweekly.
5. Maintain cultures under low light intensity (30 μmol m\(^{-2}\) s\(^{-1}\) light) at 28ºC and 16 h/8 h (light/dark) photoperiod in an incubator.
6. Continue the callus induction phase for 10 weeks. Calli are bombarded 10-weeks after culture initiation.

Preparation of Gold Microparticle Stock (60mg mL\(^{-1}\))

1. Weigh 60 mg of 1.0 µm gold particles and place in a sterile 1.5 mL tube.
2. Add 1mL of 70% ethanol and vortex for 3-5 min.
3. Centrifuge briefly (5 s) to pellet the microparticles.
4. Discard the supernatant followed by three washes with 1mL autoclaved ddH\(_2\)O.
5. Vortex for 1 min.
6. Centrifuge briefly (3-5 s) and remove the supernatant.
7. Add 1mL sterile 50% (v/v) glycerol.
8. Store the gold stock at -20ºC.

Preparation of DNA Coated Microparticles

1. Vortex gold stock suspension and transfer 30 µL into a sterile 1.5 mL tube.
2. Add the expression cassettes of the target gene and selectable marker gene as 2:1 molar ratio, respectively, and add sterilized ddH\(_2\)O to a final volume of 60 µL and vortex at low speed for 30 s.
3. Place 20 µL of 0.1 M freshly prepared spermidine and 50 µL of 2.5 M CaCl\(_2\) on the lid of the 1.5 mL tube.
4. Mix all components by closing the lid, and vortex for 1 min.
5. Centrifuge briefly (3-5 s) to pellet the gold.
6. Discard the supernatant without disturbing the pellet, and add 250 µL of absolute ethanol as a wash.
7. Centrifuge briefly (3-5 s) and discard the supernatant.
8. Repeat the previous wash with ethanol once more.
9. Re-suspend the pellet in 90 µL absolute ethanol by sonication for 1 s.
10. Keep the DNA coated microparticles on ice.

**Biolistic Bombardment using on PDS-1000/He® Particle Delivery System**

1. Place callus on CI-3 media supplemented with 0.4 M sorbitol 4-6 h prior to bombardment.
2. Turn on PDS-1000/He® Particle Delivery System and vacuum pump. Ensure the helium supply is at least 200 psi above the desired pressure optimum.
3. Place the rupture disk in the centre of the rupture disk holder and secure it properly inside the chamber.
4. Place macrocarriers into holders with forceps and push down with a sterile blunt object to secure it in holders.
5. Re-suspend the DNA-coated microparticles by vortexing briefly.
6. Apply 5 µl of the suspension of DNA coated microparticles into the center (inner 5 mm diameter) of the macrocarrier. Allow for complete evaporation of ethanol before bombardment.
7. Place stopping screen into the macrocarrier plate and insert the inverted macrocarrier assembly on top. Secure the lid on top of the shelf assembly.
8. Place macrocarrier plate containing the macrocarrier at the highest level of the inner chamber.
9. Place tissue culture plate on the second shelf below the macrocarrier plate (6 cm below).
10. Initiate a vacuum to 27.5 Hg, press and hold the fire button until the disc ruptures at 1100 psi.
11. Vent the vacuum and remove bombarded calli.
12. Dismantle the assembly and prepare for the next shot.

**Selection and Regeneration Protocol**

1. Sixteen hours (overnight) after bombardment, transfer the bombarded calli to CI-3 media and maintain the culture for 6 d under 30 µmol m⁻² s⁻¹ light intensity at 28ºC and 16 h/8 h (light/dark) photoperiod in an incubator.
2. Transfer the calli onto selection medium which consists of CI-3 with 30 mg L⁻¹ geneticin (G-418). Phytagel is replaced by agarose (Type I, Sigma). Track the identity of independent callus lines through the selection and regeneration procedure to recover independent events. Subculture calli to new selection medium 3X biweekly.
3. Transfer the selected calli to regeneration CI-3 medium supplement with 50 mg L⁻¹ thidiazuron (TDZ), and culture for 2 weeks under 100 µmol m⁻² s⁻¹ light intensity at 28ºC and 16 h/8 h (light / dark) photoperiod in an incubator.
4. Transfer calli with shoots to selection rooting CI-3 media without 2,4-D containing 30 mg L⁻¹ paromomycin. Subculture 4X biweekly under 100 µmol m⁻² s⁻¹ light intensity at 28ºC and 16 h/8 h (light/dark) photoperiod in an incubator. If the plantlet reaches the lid, use a deep Petri-dish.
5. Transplant elongated shoots with roots into soil and maintain in a growth chamber with 80% relative humidity under 500 µmol m⁻² s⁻¹ light intensity at 28ºC and 16 h/8 h.
(light/dark) photoperiod. Keep the regenerated plants covered during the first 4-6 d with a transparent container to maintain humidity.

6. After 14 d of acclimatization in the growth chamber, move the transgenic plants to glasshouse maintained at 28ºC day and 22ºC night, under natural photoperiod.

Stock Solution and Media Preparation

Macro-stock solution

<table>
<thead>
<tr>
<th>ddH₂O</th>
<th>800 mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonium Nitrate</td>
<td>16.5 g</td>
</tr>
<tr>
<td>Potassium Nitrate</td>
<td>19.0 g</td>
</tr>
<tr>
<td>Calcium Chloride dehydrate</td>
<td>4.4 g</td>
</tr>
<tr>
<td>Magnesium Sulfate heptahydrate</td>
<td>3.7 g</td>
</tr>
<tr>
<td>Potassium Phosphate, monobasic</td>
<td>1.7 g</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Fill up to 1000 mL</td>
</tr>
</tbody>
</table>

Mix all ingredients under constant stirring. Store in bottle at 4 ºC.

Micro-stock solution

<table>
<thead>
<tr>
<th>ddH₂O</th>
<th>400 mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potassium Iodide</td>
<td>0.04150 g</td>
</tr>
<tr>
<td>Boric Acid</td>
<td>0.31000 g</td>
</tr>
<tr>
<td>Manganese Sulfate</td>
<td>0.64000 g</td>
</tr>
<tr>
<td>Zinc Sulfate heptahydrate</td>
<td>0.43000 g</td>
</tr>
<tr>
<td>Sodium Molybdate dihydrate</td>
<td>0.01250 g</td>
</tr>
<tr>
<td>Cupric Sulfate pentahydrate</td>
<td>0.00125 g</td>
</tr>
<tr>
<td>Cobalt Chloride hexahydrate</td>
<td>0.00125 g</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Fill up to 500 mL</td>
</tr>
</tbody>
</table>

Mix all ingredients under constant stirring. Store in bottle at 4 ºC.

Fe-stock solution

<table>
<thead>
<tr>
<th>ddH₂O</th>
<th>400 mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na₂EDTA</td>
<td>0.93 g</td>
</tr>
<tr>
<td>FeSO₄ ·7H₂O</td>
<td>0.65 g</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Fill up to 500 mL</td>
</tr>
</tbody>
</table>

Heat 400 mL ddH₂O in beaker, but do not boil water. Add Na₂EDTA to hot water under constant stirring. Once it dissolves, remove from heat before adding FeSO₄ but continue stirring. Store in light protective bottle at 4 ºC.

2,4-Dichlorophenoxyacetic acid (3 mg mL⁻¹)

0.15 g powder dissolved in 500 μL 1N NaOH.

Make up to 50 mL with ddH₂O. Store in aliquots at -20ºC. Use 1 mL L⁻¹ media.

CuSO₄-stock (12.45 mg mL⁻¹)

0.6225g of CuSO₄·5H₂O dissolved in 50 mL ddH₂O.

Store at -20ºC. Use 100 μL L⁻¹ media.
B5G-Vitamin-Stock solution/filter-sterilized

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>ddH₂O</td>
<td>90 mL</td>
</tr>
<tr>
<td>Nicotinic Acid</td>
<td>0.10 g</td>
</tr>
<tr>
<td>Thiamine HCl</td>
<td>1.00 g</td>
</tr>
<tr>
<td>Pyridoxine HCl</td>
<td>0.10 g</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.20 g</td>
</tr>
<tr>
<td>Myo-Inositol</td>
<td>10.0 g</td>
</tr>
</tbody>
</table>

ddH₂O Fill up to 100 mL

Bring to clean bench and filter the solution through 0.2 μm syringe filter and place into prepared sterile 1.5 mL tubes. Store tubes at -20°C and only open them in clean bench when preparing media. Before adding to media, thaw it completely and vortex.

Geneticin sulfate for callus selection media (30 mg mL⁻¹)

Dissolve 0.3 g Geneticin G418 in 10 mL ddH₂O. Filter-sterilize and store in aliquots at -20°C. Use 1 mL L⁻¹ CI-3 media.

Thidiazuron (TDZ) for shoot regeneration media (50 mg mL⁻¹)

0.5 g TDZ dissolved in 500 μL 1N NaOH. Make up to 10 mL with ddH₂O. Store in aliquots at -20°C. Use 1 mL L⁻¹ CI-3 media.

Paromomycin sulfate for selection rooting media (30 mg mL⁻¹)

Dissolve 0.3 g paromomycin sulphate in 10 mL ddH₂O. Filter-sterilize and store in aliquots at -20°C. Use 1 mL L⁻¹ CI-3 media without 2,4-D.

Callus induction CI-3 media

Pre-Autoclave

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>ddH₂O</td>
<td>400 mL</td>
</tr>
<tr>
<td>Sucrose</td>
<td>10 g</td>
</tr>
<tr>
<td>Macro-stock</td>
<td>50 mL</td>
</tr>
<tr>
<td>Micro-stock</td>
<td>5 mL</td>
</tr>
<tr>
<td>Fe-stock³</td>
<td>10 mL</td>
</tr>
<tr>
<td>2,4-Dichlorophenoxyacetic acid-stock</td>
<td>500 μL</td>
</tr>
<tr>
<td>CuSO₄-stock</td>
<td>50 μL</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Fill up to 500 mL</td>
</tr>
<tr>
<td>pH</td>
<td>5.8 for each bottle</td>
</tr>
<tr>
<td>Phytagel (Gelrite)</td>
<td>1.5 g per bottle</td>
</tr>
</tbody>
</table>

Post-Autoclave

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>B5G-Vitamin-stock</td>
<td>500 μL per bottle</td>
</tr>
</tbody>
</table>

Small RNA Northern Blot

Total RNA Extraction

1. Prepare 10 mL total RNA extraction buffer and 10 mL phenol (pH 4.5) for each sample, and warm to 55°C.
2. Grind 3 g of internode tissue under liquid nitrogen, and place the sample in a 50 mL polypropylene tube. Do not let the sample thaw.
3. Add 10 mL extraction buffer and 10 mL phenol, and mix thoroughly by vortexing. Incubate at 55°C for 5 min.
4. Cool to room-temperature with shaking for 5 min.
5. Add 10 mL of chloroform and shake at room-temperature for 20 min.
6. Centrifuge 13,000 g at 4°C for 15 min.
7. Transfer supernatant to a fresh 50 mL polypropylene tube.
8. Add an equal volume of chloroform to the tube and shake at room-temperature for 15 min.
9. Centrifuge 13,000 g at 4°C for 10 min.
10. Transfer the supernatant to a fresh 50 mL polypropylene tube.
11. Add an equal volume of isopropanol and incubate at -20°C for 16 h.
12. Centrifuge 13,000 g at 4°C for 20 min to precipitate total RNA.
13. Discard the supernatant, add 10 mL of cold 80% ethanol, and centrifuge 13,000 g at 4°C for 5 min.
15. Remove supernatant and let RNA pellet dry at room-temperature for 5 min. Do not completely dry the pellet.
16. Add 250 µL of DEPC treated water and dissolve RNA. Take 10 µL of total RNA and check RNA quality using formaldehyde gel electrophoresis.

**Small RNA Separation from Total RNA**

1. Take 240 µL of total RNA and add 30 µL of 50% (w/v) PEG (M.W. 8,000) and 30 µL of 5 M NaCl. The final concentrations of PEG and NaCl are 5% and 0.5M, respectively.
2. Mix thoroughly and incubate on ice for at least 30 min.
3. Centrifuge at 16,000 g for 10 min at 4°C to pellet the high molecular weight RNA.
4. Transfer supernatant to a fresh 1.5 mL tube.
5. Add an equal volume of cold isopropanol. Mix well and place at -20°C for at least 2 h.
6. Centrifuge at 16,000 g for 30 min at 4°C to pellet the small RNA.
7. Carefully remove the supernatant, add 1 mL of cold 80% ethanol, and centrifuge at 16,000 g at 4°C for 5 min.
8. Air dry for 5 min at room-temperature, and dissolve small RNA with 10 µL of DEPC treated water.
9. Check RNA quantity and quality using the Nano-Drop and 1.5% agarose gel electrophoresis.

**Membrane Transfer**

1. Assemble Mini-PROTEAN system (Bio-rad) with 1×TBE buffer, 15% polyacrylamide/urea ready-gel (Cat No. 161-1189, 12 well, Bio-rad), and fill the outside of the tank with ice or set the assembly at 4°C.
2. After the system assembly, wash the slots with 1x TBE using a 1 mL syringe.
3. Pre-run a gel for 45 min at 180 V/400 mA.
4. Prepare 20 µL of RNA loading sample with 20 µg of small RNA in 10 µL DEPC treated water and 10 µL of loading buffer. Prepare 100 fmole of 23-mer DNA oligo in
20 μL of loading sample as a positive control and a size marker. For the controls, 21 and 24-mer RNA oligos can be used.
5. Denature the samples and the controls at 65°C for 10 min, and put on ice immediately.
6. Wash slots again with 1x TBE using 1 mL syringe.
7. Load samples, and run at 180 V/400 mA for about 75 min until bromphenol blue dye reaches the bottom of the gel.
8. Take a gel out of the running chamber and stain the gel with EtBr solution (1 μg mL⁻¹ 1xTBE buffer) for 5 min.
9. Verify small RNA quality with a photo, and de-stain EtBr with 1xTBE until the membrane transfer.
10. Soak Hybond-N+ membrane (GE Healthcare Biosciences) with 1xTBE buffer for 5 min.
11. Prepare Semi-Dry Transfer Cell (Bio-Rad) in the following orders: Bottom → Whatman paper → membrane → Gel → Whatman paper → Top
12. Perform membrane transfer at 10 V/400 mA for 35 min. Do not exceed 35 min.
13. Wash membrane with 1xTBE for 1 min
14. Let the membrane dry briefly at room temperature.
15. Perform UV cross-linking for 1 min.
16. Bake membrane for at least 1 h at 80°C.
17. Store membrane at 4°C until use.

**Probe Labeling, Hybridization, and Detection**

1. Check the radioactive working area for previous radioisotope contamination before commencing work.
2. Thaw Prime-a-Gene® Labeling System (Promega), and place ³²P-dCTP (Perkin Elmer) behind the plexiglass shield to thaw.
3. Warm the Church’s Hybridization buffer (see buffer preparation below) to 38°C.
4. Boil 25 ng of probe in 30 μL ddH₂O for 5 min, and place the denatured probe in ice for 5 min.
5. Set up the labeling reaction as follows according to the manufacture’s instruction. Add the following to the denatured probe
   - 5x labelling buffer 10 μL
   - Unlabeled dNTP mix 2 μL
   - BSA 2 μL
   - Klenow (5U/μL) 1 μL
   - Total volume 45 μL
6. Move the mixture behind the plexiglass shield, add 5 μL of ³²P-dCTP, and mix well by pipetting. The final volume of the reaction is 50 μL.
7. Incubate the mixture behind the plexiglass shield at room-temperature for 4 h.
8. During the probe incubation time, place the membrane inside a hybridization tube, and perform pre-hybridization in the hybridization oven for 4 h at 38°C with 1 mL of Church’s Hybridization buffer per 10 cm² of membrane. Do not add blocking DNA, i.e. sheared salmon sperm DNA.
9. Just before performing hybridization, boil the probe for 5 min behind the plexiglass shield. Do not add blocking DNA.
10. Discard the pre-hybridization buffer and add new hybridization buffer into the tube.
11. Immediately after probe boiling, add probe mixture into the tube behind the plexiglass shield.
12. Perform hybridization at 38°C for 16 h in the hybridization oven.
13. After hybridization, prepare washing solution, 2×SSC - 0.2% SDS, and warm to 50°C.
14. Working behind the plexiglass shield, dispose of the hybridization solution into the hazardous waste container using a funnel, taking care to avoid any spills.
15. Wash the membrane three times with washing solution for 20 min at 50°C.
16. Remove the tubes from the oven and place them behind the plexiglass shield. Dispose of the wash solution into the hazardous waste container, and wrap the membrane in Saran wrap.
17. Check for radioactivity on the membrane using a survey meter. In addition, check the working area for any radioactive contamination.
18. If there is radioactivity on a blank lane, wash the membrane once more with 1×SSC - 0.1% SDS for 20 min at 50°C.
19. Place the membrane with an X-ray film (Kodak) in an autoradiography cassette and allow 16-18 h for exposure. Place the cassette at -80°C during exposure.

Buffer Preparation

DEPC treated water (0.1%, v/v)
Dissolve 1 mL DEPC in 1 L ddH₂O under constant stirring for 16 h. Autoclave the solution.

Total RNA extraction buffer
0.1M LiCl
0.1M Tris-HCl, pH 8.0
0.01M EDTA, pH 8.0
1% (w/v) SDS
0.1% (w/v) PVP (FW 40,000)
Store at room-temperature

Loading buffer
98% (v/v) deionized formamide
10 mM EDTA pH 8.0
0.025% (w/v) xylene cyanol
0.025% (w/v) bromphenol blue
Store at −20 °C

Church’s hybridization buffer without BSA
0.5 M sodium phosphate buffer, pH 7.2
1 mM EDTA, pH 8.0
7% (w/v) SDS
Store up to 1 year at room-temperature

20X SSC
3 M sodium chloride
0.3 M sodium citrate

Analysis of Cell Wall Components

Determination of Dry Weight (DW) of the Extract Free Sample

1. Dry crucibles at 105°C for at least 4 h and cool-down in a desiccator. Weigh crucibles to the nearest 0.1 mg.
2. Weigh 100 ± 10 mg extract-free sample in a crucible and record the weight to the nearest 0.1 mg.
3. Dry the crucible with the sample at 105°C for at least 4 h and cool-down in a desiccator. Weigh to the nearest 0.1 mg.
4. Place the sample back into the dry oven at 105°C and dry to the constant weight, which is defined as ± 0.1% change in the weight percent solids upon 1 h of re-drying the sample.

Determination of Total Lignin Content using Acetyl Bromide

1. Weigh 2 ± 0.1 mg of extract-free sample, and place 2 mL of polypropylene tube with screw cap.
2. Add 1 mL of freshly prepared 25% (w/w) acetyl bromide in glacial acetic acid into the tube. Prepare the blank solution without the sample. Perform the reaction under the fume hood. Do not use molecular grade acetic acid.
3. Incubate the tube in a water bath at 50°C for 4 h, and during the last hour, thoroughly mix the sample at 15 min intervals.
4. Take the tube out of the water bath, and incubate the tube on ice for 30 min.
5. Prepare a new tube containing 200 µL of 2 M NaOH and 1.7 mL of glacial acetic acid.
6. Take 100 µL from the reaction mixture and transfer to the new tube with NaOH and acetic acid. Prepare three reactions per sample given inaccuracies due to pipetting errors.
7. Mix thoroughly and transfer all of the mixture to a UV quartz cuvette. Do not use a plastic disposable cuvette.
8. Measure the absorbance at 280 nm.
9. The lignin content can be calculated by employing the molar extinction coefficient of 21.5 L g⁻¹ cm⁻¹ for milled sugarcane vascular bundle lignin.

Determination of Lignin Composition using Thioacidolysis following GC/MS

1. Weigh 10 ± 0.1 mg of extract-free sample and place into a 5 mL glass reaction vial with Teflon-lined screw-cap.
2. Add 1 mL of freshly prepared reaction mixture and blank with nitrogen gas prior to sealing. Reaction mixture: 2.5% (v/v) boron trifluoride etherate and 10% (v/v) ethanethiol, in recently distilled dioxane. Prepare reaction mixture under the fume hood wearing a respiratory protection mask.

3. Incubate at 100 ºC for 4 h with manual agitation hourly.
4. Stop the reaction at -20 ºC for 5 min.
5. Add 0.2 mL of internal standard (tetracosane, 5 mg mL⁻¹ methylene chloride) to each vial.
6. Add 0.3 mL of 0.4 M sodium bicarbonate, enough to bring the reaction pH to between 3 and 4, check using pH paper.
7. Add 2 mL of water, 1 mL of methylene chloride, vortex, and let settle to separate the phases.
8. Prepare a Pasteur pipette packed with a small piece of Kimwipes and ~ 50 mg of granular anhydrous sodium sulfate.
9. Remove an aliquot (1.5 mL) of organic lower phase using an autopipette from the reaction vial, and simultaneously pass through the Pasteur pipette, and transfer directly into a 2 mL polypropylene tube.
10. Dry the sample under a stream of nitrogen at 45 ºC for 90 min, and re-suspend sample in 1 mL of methylene chloride.
11. Prepare 20 µL of pyridine and 100 µL of N,O-bis(trimethylsilyl)acetamide (BSTFA) in a GC glass vial.
12. Take 20 µL of resuspended sample and add into the vial.
13. Incubate reaction mixture for at least 2 h at room-temperature.
14. Prepare tetracosane standard ranging from 100 µg to 1000 µg mL⁻¹ methylene chloride.
15. Set up the GC (Varian 3800) conditions as follows;
   - Carrier: helium 1.2 mL min⁻¹
   - Injector temperature: 250 ºC
   - Oven temperature: hold at 130 ºC for 3 min, increase to 250 ºC at 3ºC min⁻¹, and hold constant for 5 min.
16. Set up the MS (Varian 1200) operational conditions as follows;
   - Electron impact mode at 70 eV
   - Detector operation at 1.2 kV
   - Mass range m/z 50-550, and scan every 0.2 s
17. After setting the GC/MS, inject 1 µL of sample and standard into Factor-4 VF-5ht column (35 m, 0.25 mm i.d.) with 1:10 split ratio.
18. Analyze the data with MS Workstation software. Peaks from thioacidolysis derivatives for H, G, and S units can be identified by characteristic mass spectrum ions of m/z 239, 269, and 299, respectively. The peak from IS can be identified based on the peak from tetracosane standard.
19. Retrieve peak area corresponding to H, G, and S unit, and internal standard (IS).
20. Calculate the recovery rate based on the peak area of IS and tetracosane standard curve, and normalize the lignin monomer peak area with the recovery rate.
21. Calculate the quantity of lignin monomers using the response factors of each monomer against IS as follows: H versus IS, 0.42, G versus IS, 0.47, and S versus IS, 0.53.
22. Calculate the molar concentration of each monomer. Molecular weights for H, G, and S are 388, 418, and 448, respectively.

**Determination of Cell Wall Carbohydrates, Acid Insoluble/Soluble Lignin, and Ash Content**

1. Burn filtering crucibles at 575°C for 24 h, and cool-down in a desiccator, and record the weight to the nearest 0.1 mg.
2. Weigh 300 ± 0.1 mg of extract-free sample, and place into a glass pressure tube (38 x 203 mm) with Teflon front seal plugs.
3. Add 3 mL of 72% (w/w) sulfuric acid to each tube and mix thoroughly to wet the sample.
4. Incubate the tube in a shaking water bath at 30°C for 1 h.
5. Add 84 mL of ddH2O to dilute sulfuric acid to make 4% solution, and mix samples thoroughly by inverting to eliminate the phase separation.
6. Perform autoclave at 121°C for 1 h.
7. Vacuum filter the autoclaved hydrolysis solution through previously weighed filtering crucibles.
8. Transfer 20 mL of an aliquot to a sample storage bottle for the determination of acid soluble lignin.
9. For acid soluble lignin, measure the absorbance of the aliquot at 240 nm blanked with 4% sulfuric acid. Use molar extinction coefficient of 25.0 L g⁻¹ cm⁻¹ for the calculation of acid soluble lignin.
10. Transfer and filter all remaining solids in the pressure tube using ddH2O.
11. Dry crucible and acid insoluble residues at 105°C until a constant weight is achieved.
12. Cool down the crucible in a desiccator and record the weight to the nearest 0.1 mg.
13. Burn the crucible at 575°C for 24 h for the determination of ash content. Cool down the crucible in a desiccator and record the weight to the nearest 0.1 mg.
14. Transfer 20 mL of an aliquot after step 10 into a 50 mL Erlenmeyer flask.
15. Add calcium carbonate to bring the sample pH 5-6. Manually agitate the sample frequently and monitor the pH while adding calcium carbonate using pH paper. After reaching pH to 5-6, stop adding calcium carbonate, and allow the sample to settle.
16. Decant the supernatant and filtrate through a 0.2 µm syringe filter. Collect 1 mL of filtrate into a HPLC glass vial.
17. Analyze the calibration standard (D-glucose, D-xylose, L-arabinose), internal standard, and samples by HPLC coupled with RI detector using a Bio-rad Aminex HPX-87H column and operating at 65°C with 4 mM H2SO4 mobile phase at a flow rate of 0.6 mL min⁻¹.

**Determination of Ester-Linked p-Coumarate and Ferulate Content**

1. Weigh 25 ± 0.1 mg of extract-free sample and place into a 2 mL polypropylene tube with a screw cap.
2. Add 1.7 mL of 2 M NaOH containing 20 µL of 2-hydroxycinnamic acid (1 mg mL⁻¹ 2M NaOH) as an internal control, and blank with nitrogen gas prior to sealing.
3. Incubate at room-temperature for 20 h under dark conditions with shaking.
4. Add 0.3 mL of 12 M HCl to acidify the sample.
5. Take 1 mL of reaction mixture and transfer into a new 2 mL centrifuge tube.
6. Add 1 mL of diethyl ether, vortex for 5 s, and centrifuge 16,000 g for 5 min.
7. Take the supernatant and add the equal volume of diethyl ether, and repeat step 6.
8. Repeat step 7.
9. Combine the supernatants and dry under a stream of nitrogen gas.
10. Add 40 µL of methoxyamine hydrochloride (20 mg mL\(^{-1}\) pyridine) and incubate at 37°C for 90 min.
11. Adding 60 µl of MSTFA [N-methyl-N-(trimethylsilyl) trifluoroacetamide], and incubate at 37°C for 30 min.
12. Prepare the standard reference for \(p\)-coumarate, ferulate, and 2-hydroxyccinnamic acid.
13. Set up the GC (Varian 3800) conditions as follows;
   Carrier: helium 1.2 mL min\(^{-1}\)
   Injector temperature: 220 °C
   Oven temperature: hold at 70 °C for 4 min, increase to 250 °C at 10°C min\(^{-1}\), and hold constant for 2 min.
14. Set up the MS (Varian 1200) operational conditions as follows;
   Electron impact mode at 70 eV
   Detector operation at 1.0 kV
   Mass range m/z 45-650, and scan every 0.2 s
15. After setting the GC/MS, inject 1 µL of sample and standard into Factor-4 VF-5ht column (35 m, 0.25 mm i.d.) with 1:10 split ratio.
16. Analyze data with MS Workstation software.
17. Peaks from derivatized \(p\)-coumarate, ferulate, and internal standard can be identified by their characteristic mass spectrum ions obtained from the reference compounds. The concentration of \(p\)-coumarate, ferulate, and internal standards can be determined using the standard curve. The concentration of \(p\)-coumarate and ferulate must be normalized with the concentration of internal standard among the samples.

**Evaluation of Saccharification Efficiency**

**Diluted Sulfuric Acid Pretreatment**

1. Weigh a sample equivalent to 0.1 g of cellulose and place into a 50 mL glass tube. The cellulose content is determined as “glucose minus the contribution of any starch present in the sample.” Prepare 0.1 g of No.1 Watmann filter paper for the positive control and check the saccharification efficiency.
2. Mix the sample at a solid loading of 10% (w/w) with dilute sulfuric acid (final concentration of 1.3%, w/w).
3. Incubate the sample at room-temperature for 1 h.
4. Autoclave the sample at 121°C for 40 min.
5. Wash the sample with 50 mL of ddH\(_2\)O and drain the liquid through nylon vacuum filter.
6. Collect the all remaining pre-treated residues with 50 mL of ddH₂O and repeat step 5 twice.
7. Collect the all remaining pre-treated residues with 8 mL of ddH₂O and place it into a pre-weighed 50 mL polypropylene tube. Do not let the sample dry.

**Enzymatic Saccharification and Determination of Glucose Yields**

1. Add 10 mL of 0.1 M sodium citrate buffer (pH 4.8) and 200 µL of 2% (w/w) sodium azide into the tube with pre-treated sample.
2. Add the appropriate volume of cellulases which is dependent on the activities of each enzyme and the enzyme dosage per unit biomass. Prepare the sample without adding cellulases for the enzyme blank control.
3. Add the appropriate volume of ddH₂O to bring the final mass of the sample to 20 g, assuming all solutions and the biomass sample have 1.000 g mL⁻¹ specific gravity.
4. Incubate the sample in a shacking incubator at 50°C and 250 rpm for 72 h.
5. Take 1 mL of the hydrolyzed sample while shaking the sample, and filter it through a 0.2 µm syringe filter.
6. Measure the glucose amount from the filtrate using the YSI glucose analyzer.
7. Correct the YSI reading value for hydration by multiplying with 0.9.

**Determination of Starch Content**

1. Weigh 100 ± 0.1 mg of sample and place into a 50 mL Erlenmeyer flask. Use 100 ± 0.1 mg of amylopectin as a standard reference material which is run in parallel to a batch of samples.
2. Add 5 mL of ddH₂O and mix thoroughly to wet the sample.
3. Add 2 mL of 2N NaOH and incubate at 90°C for 20 min. Swirl to mix every 5 min.
4. Add 2 mL of 2N HCl and swirl to mix. Cool the sample to below 50°C.
5. Add 2 mL of 1M sodium acetate buffer (pH 4.2) and swirl to mix.
6. Add 1 mL of amylglucosidase (60 U mL⁻¹), mix well, and incubate the sample in a shaking incubator at 40 °C and 200 rpm for 60 min.
7. Take the sample out of the incubator and add 1 mL of 25% (w/v) trichloroacetic acid to stop the enzyme reaction.
8. Cool down the sample to room-temperature and transfer it to a 20 mL of volumetric flask. Rinse out all of the remaining residues using 0.4 M sodium phosphate buffer (pH 6.2) and transfer to the flask.
9. Add additional sodium phosphate buffer to the final volume of 20 mL.
10. Mix well, take 1 mL of the sample and filter through a 0.2 µm syringe filter.
11. Measure the glucose amount from the filtrate using the YSI glucose analyzer.
12. Measure the free glucose amount in the enzyme, and subtract it from the value of samples.
13. Correct the YSI reading value for hydration by multiplying 0.9.
14. Correct the value with the digestion efficiency of standard reference material with a given enzyme.
### Buffer Preparation

#### 0.1 M sodium citrate buffer (pH 4.8)

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ddH₂O</td>
<td>400 mL</td>
</tr>
<tr>
<td>Citric acid monohydrate</td>
<td>4.4 g</td>
</tr>
<tr>
<td>Trisodium citrate dihydrate</td>
<td>8.5 g</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Fill up to 500 mL</td>
</tr>
</tbody>
</table>

#### 1 M sodium acetate buffer (pH 4.2)

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ddH₂O</td>
<td>400 mL</td>
</tr>
<tr>
<td>Acetic acid (glacial)</td>
<td>23.4 g</td>
</tr>
<tr>
<td>Sodium acetate anhydrous</td>
<td>9.1 g</td>
</tr>
<tr>
<td>ddH₂O</td>
<td>Fill up to 500 mL</td>
</tr>
</tbody>
</table>

#### 0.4 M sodium phosphate buffer (pH 6.2)

<table>
<thead>
<tr>
<th>Ingredient</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4 M sodium phosphate monobasic</td>
<td>815 mL</td>
</tr>
<tr>
<td>0.4 M sodium phosphate dibasic</td>
<td>195 mL</td>
</tr>
<tr>
<td>Total</td>
<td>1000 mL</td>
</tr>
</tbody>
</table>


BIOGRAPHICAL SKETCH

Je Hyeong Jung was born in Seoul, Korea. He went to the Korean University, Seoul, Korea, in 1998. During his undergraduate degree, he was quite active in both academic and extra-curricular activities. He was awarded the Academic Excellence Scholarship for three semesters. He was the President of the Undergraduate Student’s Association in the Department of Crop Science. He also served 26 months in the army, and after finishing military service, he travelled to Canada for 6 months.

After he graduated with a B.S. (crop science) in 2005, he joined the Graduate School at the Korean University and pursued an M.S. in life sciences and biotechnology under the guidance of Dr. Yong Weon Seo. After obtaining an M.S. degree in 2007, he got an offer from KT&G Central Research Institute, Daejeon, Korea, and participated in several research programs as an Assistant Researcher, until he came to the U.S. to pursue a Ph.D.

In 2009, he joined Dr. Fredy Altpeter’s lab at University of Florida. He was a Teaching Assistant in Genetics from 2009 to 2012. He was awarded The Paul Robin Harris Memorial Scholarship for three consecutive years from the Agronomy Department, University of Florida. He won the Wilton Earle Award at the International Plant Biotechnology Conference in 2010 and The Philip White Memorial Award at the Society for In Vitro Biology in 2011.