Investigation of lignin biosynthesis in sugarcane for improved lignocellulosic ethanol production

By

Patrick Bewg

Bachelor of Applied Science (Hons.)

A thesis submitted for the degree of Doctor of Philosophy

Centre for Tropical Crops and Biocommodities

Queensland University of Technology

2015

Keywords

Bagasse

Lignin biosynthesis

Lignocellulose

RNAi

Second generation bioethanol

Sugarcane

*Zm*MYB31

*Zm*MYB42

Abstract

Second generation bioethanol is a renewable energy resource produced from lignocellulosic biomass with the potential to reduce reliance on oil-based energy. Sugarcane bagasse is an abundant source of lignocellulosic material available for bioethanol production. The utilisation of bagasse for biofuel production would be environmentally and economically beneficial, however, the lignin polymers restrict polysaccharide degradation by hydrolytic enzymes. Pretreatment is currently required to overcome the recalcitrant nature of lignin polymers, the cost of which prevents the cost-competitive production of bioethanol from lignocellulosic biomass. There is a strong consensus in published literature that reducing lignin content can increase glucose liberation during enzymatic hydrolysis of both wild type and genetically modified plants, including genetically modified sugarcane. Whilst lignin biosynthesis has received increasing research attention in some plant species, lignin biosynthesis and its manipulation in sugarcane has been explored only in recent publications and remains far from being fully understood. This thesis focuses on contributing to the limited knowledge available concerning lignin biosynthesis and secondary cell wall deposition in wild type sugarcane and the saccharification potential of genetically modified, lignin-reduced sugarcane bagasse through the completion of three specific aims.

The first aim was to develop a lignin biosynthesis profile in wild type sugarcane by correlating gene expression data with cell wall compositional data in sugarcane internodes. The expression levels of nine genes within the lignin biosynthesis pathway (PAL, C4H, C3H, 4CL, CCoAOMT, F5H, CAD, CCR and COMT) were quantified in five sugarcane stem sections of increasing maturity (section A through section E) and in root tissue. Analysis found two distinct expression patterns in maturing stem tissue. The first pattern saw highest gene expression in section A (youngest tissue), which then decreased as tissue matured (sections B – E). This pattern was strongest in PAL and CCR, and less pronounced in 4CL, COMT and CAD. In root tissue CAD expression did not differ significantly from any stem section, whereas root expression of PAL, CCR, 4CL and COMT was equivalent to that of stem section A, but higher than more mature stem sections. The second expression pattern saw little to no change in transcription levels of C4H, C3H, CCoAOMT and F5H across the five stem

sections. Expression of C3H, CCoAOMT and F5H within root tissue did not differ from any stem section whereas the root expression of C4H was approximately 9-fold higher than its expression in any stem section. The expression pattern of the nine genes did not appear linked with their position within the lignin biosynthesis pathway, suggesting the individual role of each gene may influence its expression pattern. Cell wall compositional analysis of the five stem sections found total lignin content significantly increased between section A (the youngest stem tissue) and section B, but in more mature stem, no differences were seen between sections B, C, D and E. There were no differences in glucose, xylose or galactose across the five stem sections, although section A had significantly higher levels of arabinose than the more mature stem regions. This suggests that the deposition of structural carbohydrates occurs early during sugarcane stem formation.

The second aim of this thesis was to increase saccharification of sugarcane bagasse by reducing lignin content through the transgenic expression of two maize R2R3 MYB transcription factors, *Zm*MYB31 and *Zm*MYB42, known to control down-regulation of lignin biosynthesis in Arabidopsis. Using maize cDNA as a cloning template, the open reading frame (ORF) each MYB gene (MYB31 and MYB42) was cloned to both include and exclude adjacent sequences of the 5' and 3' untranslated region (UTR). This resulted in two MYB31 constructs (MYB31 ORF and MYB31 UTR) and two MYB42 constructs (MYB42 ORF and MYB42 UTR).

A total of 33 MYB31 (14 ORF and 19 UTR) and 23 MYB42 (10 ORF and 13 UTR) sugarcane plants were confirmed qPCR positive for transgene expression. Analysis of young and maturing internodes of glasshouse grown plants found MYB31 exerted greater down-regulatory control over a higher number of lignin biosynthesis pathway genes than MYB42. Surprisingly, only one out of 14 MYB31 plants analysed had significantly less total lignin and two additional plants had less acid insoluble lignin, with approximately half of the MYB31 plants having increased hemicellulose contents. Of the 14 MYB42 plants analysed for cell wall composition, six had less total lignin but showed little change to cell wall polysaccharide levels. The three plants with the lowest total lignin contents for each MYB line underwent enzymatic hydrolysis. Two MYB31 UTR plants and no MYB31 ORF plants released a greater amount of glucose after 72 hours of enzymatic hydrolysis, whereas all MYB42 ORF and UTR plants performed better than the transgenic controls. Phenotypically there were no differences between MYB transgenic sugarcane and transgenic controls for height and internode number. Reduced internode diameters were observed in some

MYB31 plants but not in MYB42 plants. Juice sugar concentrations were also quantified and all plants were found to be comparable with controls except a single MYB31 plant having significantly lower levels of sucrose. Improved bagasse digestibility without negatively impacting juice sucrose levels increases the economic value of these plants to the sugarcane industry.

The third aim of this research was to generate lignin altered transgenic sugarcane with improved saccharification by specific RNAi targeting of three lignin biosynthetic pathway genes; CCoAOMT, F5H and COMT. In total, 12 CCoAOMT-RNAi, 15 F5H-RNAi and 13 COMT-RNAi plants were regenerated, with each regenerated transgenic plant harbouring only one of the three RNAi cassettes generated. As well as potentially reducing lignin content, targeting these genes may also affect the H:G:S lignin monomer ratios, which has been linked with changes in biomass digestibility. Initial analysis of sugarcane harbouring RNAi cassettes suggested down-regulation of each of the target genes was occurring and analysis of glasshouse grown plants found this trend continued during plant development. Gene down-regulation was greater in more mature tissue than young tissue and stronger in the CCoAOMT and F5H lines than in the COMT line. Little correlation was seen between levels of RNAi gene down-regulation and cell wall composition. Across the three transgenic lines only one COMT-RNAi plant showed a reduction to lignin content. Unexpectedly, one CCoAOMT-RNAi and one F5H-RNAi plant had increased lignin deposition. Across the three RNAi lines there was little change to structural carbohydrate levels. After 72 hours of enzymatic hydrolysis, the lignin reduced COMT-RNAi plant released significantly higher levels of glucose. Additionally, one plant from each RNAi line with no differences to lignin content also released significantly more glucose. No phenotypic differences were detected between RNAi sugarcane and controls comparing height, internode number and internode diameter and all plants had juice sucrose levels equivalent to controls with two F5H-RNAi plants having significantly higher levels of sucrose. For plants to be beneficial to the sugarcane industry, it is important that sucrose production from lignin altered plants is not reduced.

Current knowledge of lignin biosynthesis in monocots is limited, even more so in sugarcane. Increasing interest in the production of second generation bioethanol from lignocellulosic biomass, such as sugarcane bagasse, has led to an increase in lignin-focused research as lignin polymers are a major hurdle to the production of cost-competitive biofuel. The knowledge and findings of this thesis into the biosynthesis and deposition of lignin in

sugarcane will be beneficial to the production or modification of cultivars with improved bioethanol production qualities. Additionally, there are no current reports of transcription factors being utilised to reduce the lignin content in transgenic sugarcane and the positive results in saccharification after MYB42 expression is of research interest. Furthermore, CCoAOMT and F5H have not previously been targeted by RNAi in sugarcane, and the combined results with COMT down-regulation suggest there may also be potential in further exploration of this avenue of research. The production of second generation bioethanol from sugarcane bagasse will have environmental benefits as they will reduce reliance on oil-based energy as well as economic benefits to both the Australian, and more widely, the global sugarcane industry though product diversification.

Table of Contents

List of Figures

CHAPTER 1

CHAPTER 4

CHAPTER 5

CHAPTER 6

List of Tables

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

List of Abbreviations

2,4-D	2,4-dichlorophenoxy acetic acid
4CL	4-coumarate-CoA ligase
AC element	Promoter motif containing adenosine and cytosine residues
AGO1	Argonaute protein
Am	*Antirrhinum majus*
am	Ante meridiem
ANOVA	Analysis of variance
At	*Arabidopsis thaliana*
AUS	Australia
BLAST	Basic local alignment search tool
bp	base pair
BSES	Bureau of Sugar Experiment Stations
BUX	Buckinghamshire
BW	Baden-Württemberg
C3H	Coumarate 3-hydroxylase
C4H	Cinnamate 4-hydroxylase
CA	California
$CaCl_2$	Calcium chloride
$CaCo_3$	Calcium carbonate
CAD	Cinnamyl alcohol dehydrogenase
CaMV 35S	Cauliflower Mosaic Virus 35S
CARF	Central Analytical Research Facility
CCoAOMT	Caffeoyl CoA 3-O-methyltransferase
CCR	Cinnamoyl-CoA reductase
cDNA	Complimentary deoxyribonucleic acid
CHE	Switzerland
CI	Crystallinity index
Cm	*Chrysanthemum morifolium*
COMT	Caffeic acid 3-O-methyltransferase
C-terminal	Carboxyl terminal
DE	Germany
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
DS	Dominant suppression
dsRNA	Double stranded ribonucleic acid
E	Expressing modifier
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
Eg	*Eucalyptus gunnii*
EST	Expressed sequence tag

F5H	Ferulate 5-hydroxylase
FPU	Filter paper units
FR	Fibrous roots
G monomer	Guaiacyl phenylpropanoid monomer
G4	Group number 4
gDNA	Genomic deoxyribonucleic acid
GFP	Green fluorescent protein
GM	Genetically modified
GUDSF	Griffith University DNA sequencing facility
GUS	*β*-glucuronidase reporter gene
H monomer	p-hydroxyphenyl phenylpropanoid monomer
HCT	Hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase
HPIC	High performance ion chromatography
HPLC	High performance liquid chromatography
hpRNA	Hairpin ribonucleic acid
ICUMSA	International Commission for Uniform Methods of Sugar Analysis
IL	Illinois
IN	Internode
IRE	Ireland
IPTG	Isopropyl-β-D-thiogalactopyranoside
ISR	Initial storage roots
iUbi	Maize ubiquitin 5' intron
KCl	Potassium chloride
KOH	Potassium hydroxide
KS	Kansas
LAC	Laccase
LB	Lysozyme broth
Ll	*Leucaena leucocephala*
MA	Massachusetts
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
miRNA	Micro ribonucleic acid
$MnCl_2$	Manganese(II) chloride
MO	Missouri
mRNA	Messenger ribonucleic acid
MS	Murashige and Skoog basal medium
MSC_3	MSC_3 medium semi-solid
MYB	Transcription factor first identified in avian myeloblastosis virus
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
NE	Non-expressing modifier
NE	Normalised expression
NJ	New Jersey
NLD	Netherlands

NMR	Nuclear magnetic resonance
norm	Normalised
nos	Nopaline synthase terminator
nptII	neomycin phosphotransferase II gene
NSW	New South Wales
Nt	*Nicotiana tabacum*
N-terminal	Amino terminal
OE	Over-expressed
oligo-dT	Deoxy-thymine oligonucleotide
ORF	Open reading frame
Os	*Oryza sativa*
P1	*maize pericarp color1* gene
PA	Pennsylvania
PAL	Phenylalanine ammonia lyase
PAP1	Production of anthocyanin pigment 1-Dominant
pBS	pBlueScript construct backbone
pDNA	Plasmid deoxyribonucleic acid
PDS	Phytoene desaturase gene
PER	Peroxidase
Pg	*Picea glauca*
pm	Post meridiem
Pn	*Paspalum notatum*
Pp	*Pinus pinaster*
Pt	*Pinus taeda*
PT	Pretreated bagasse
PTGS	Post transcriptional gene silencing
Pto	*Populus tomentosa*
Ptr	*Populus trichocarpa*
Ptt	*Populus tremula* L. x *tremuloides* Michx
pUC19	pUC19 plasmid backbone
Pv	*Panicum virgatum*
QLD	Queensland
qPCR	Quantitative polymerase chain reaction
QUT	Queensland University of Technology
R2	DNA binding repeat 2 motif
R3	DNA binding repeat 3 motif
RISC	Ribonucleic acid induced silencing complex
RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
RT-PCR	Reverse transcriptase polymerase chain reaction
S monomer	Syringyl phenylpropanoid monomer
SA	South Australia
Sal B	Salvianolic acid B
SEM	Standard error of the mean
siRNA	Short interfering ribonucleic acid
SOB	SOB medium liquid buffer

ssRNA	Single stranded ribonucleic acid
syntron	Synthetic intron
Ta	*Triticum aestivum*
TAE	Tris-acetate EDTA buffer
Taq	*Thermus aquaticus*
TB	Transformation buffer
TGS	Transcriptional gene silencing
Ubi	Maize ubiquitin promoter
UFO1	*Unstable factor for orange1* gene
UK	United Kingdom
UKN	Transgenic controls plants harbouring *Zm*Ubi-nptII-nos/pUC19 cassette
USA	United States of America
UT	Untreated bagasse
UTR	Untranslated region
UV-Vis	Ultra violet - visible light wavelengths
VIC	Victoria
Vv	*Vitis vinifera*
WA	Washington State
WI	Wisconsin
WT	Wild type
x-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
yo	Year old
Zm	*Zea mays*

Units and Symbols

°	Degrees	kPa	Kilopascal(s)
°C	Degrees Celsius	L	Litre
ΔCt	Delta critical threshold	mL	Millilitre(s)
Å	Angstrom	µL	Microlitre(s)
cm	Centimetre(s)	M	Molar
mm	Millimetre(s)	mM	Millimolar
µm	Micrometre(s)	µM	Micromolar
Cu/Kα_1	Copper *x* unit	pM	Picomolar
g	g-force	mA	Milliampere
g	Gram	rpm	Revolutions per minute
mg	Milligram(s)	V	Volts
µg	Microgram(s)	kV	Kilovolt
ng	Nanogram(s)	w/v	Weight/volume
hrs	Hours	w/w	Weight/weight

Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature: QUT Verified Signature

Date: 30/07/15

Acknowledgements

The completion of this research was possible due to the contributions and advice from many people. First and foremost I would like to thank Heather Coleman for being a wonderfully patient and encouraging supervisor and to acknowledge my associate supervisors Mark Kinkema and Rob Harding for project advice and thesis proof-reading. I also express my gratitude to James Dale for allowing me the opportunity to complete my PhD in the CTCB laboratories.

The molecular biology components of this research were completed with the help of the following researchers and I would like to thank (in alphabetical order) Milovan Bokan, Ben Dugdale, Jason Geijskes, Anthony James and Brett Williams. I would especially like to acknowledge Anthony Brinin for many hours of sugarcane-related help both in the laboratory and at Carseldine and Kylie Shand for patiently discussing and answering innumerable questions throughout my PhD.

This research also involved many chemistry based research techniques and I am grateful for the help of (in alphabetical order) Mark Harrison, Lalehvash Moghaddam, Danny Nguyen, Darryn Rackemann, Kylie Shand, Caroline Thai and Jan Zhang. I would especially like to thank Wanda Stolz for her help and guidance during all phases of my chemistry-based research.

I also appreciate Dani Tikel for helping many times with product ordering, paperwork and organising repairs of equipment and Carseldine glasshouse as well as Melinda McLachlan, Di O'Rourke and Joanne Spowart for administration related queries and requests. Many thanks also to Tony Raftery for granting me access to his McCrone Micronising Mills, Charleson Poovaiah for crystallinity analysis, and to Anne Morley, Jenny Davies and Amanda Meller from CSSF for autoclaving hundreds of bags of potting mix for me.

To the remaining staff and students of CTCB I would like to offer a general thank you. I doubt there are many members who have not helped me by answering questions or offering advice at one point or another. I would like to acknowledge and thank the Sugar

Research and Development Corporation and the Australian Research Council for my scholarship and project funding.

On a more personal note, I would like to thank my family for financial support and for remaining encouraging each time I decided to begin a new university course. Finally, I would like to thank Tess James for countless hours of support and takeaway coffees in the botanic gardens during many twin rivers days; and tangentially, thanks to Triple J, SoundCloud and YouTube for soundtracking many long hours over the last few years.

Chapter 1

Introduction

1.1 Second generation bioethanol

Escalating demand for oil-based energies is both unsustainable and environmentally damaging (Murray and King, 2012). Increased global awareness and concern has stimulated research into alternative and renewable energy sources (Borrion *et al.*, 2012; Limayem and Ricke, 2012). One attractive option is plant-derived biofuels. Bioethanol, one such biofuel, is an energy source produced from the fermentation of plant-derived sugars. The utilisation of bioethanol as a sustainable transport fuel will have many environmental and social benefits that traditional fossil fuels cannot offer such as lower carbon emissions, improved economic stability and reduced reliance on oil from politically unstable countries (Goldemberg, 2007; Murray and King, 2012; Yang *et al.*, 2013).

While first generation bioethanol is currently in large-scale production from sugar and corn starch, the carbohydrates utilised for fermentation are controversially sourced from the edible, high-value portion of crops, contributing to the food versus fuel debate (Yuan *et al.*, 2008). Second generation bioethanol attempts to overcome this drawback by utilising the carbohydrate content of lignocellulosic material, such as sugarcane bagasse or corn stover, the lower-valued residual waste material produced during sugar production. An overview of the complexities and considerations for economic and environmental viability of second generation biofuel production are comprehensively discussed in Borrion *et al.* (2012) and Limayem and Ricke (2012). The most significant challenges include sourcing lignocellulosic biomass, optimising different approaches and conditions for pretreatment, hydrolysis and fermentation stages and the potential application of biotechnology to these processes.

1.2 Lignocellulosic biomass for biofuels

Lignocellulosic biomass is composed of cellulose, hemicellulose and lignin (Hisano *et al.*, 2009). Hydrolysis reduces cellulose and hemicellulose to fermentable monosaccharides and

as cellulose is the most abundant polymer on earth (Delmer and Haigler, 2002) this provides a large, untapped resource for the production of bioethanol. However, utilising lignocellulosic biomass is significantly more challenging than first generation bioethanol production (Borrion *et al.*, 2012; Vanholme *et al.*, 2010). The recalcitrant nature of the cell wall imparted by the presence of lignin polymers incurs a necessary yet costly pretreatment step (Benjamin *et al.*, 2013; Chen and Dixon, 2007) which presents a major challenge to the production of cost-competitive, commercial scale second generation bioethanol (Benjamin *et al.*, 2013; Yang *et al.*, 2013; Yuan *et al.*, 2008). Although lignin is necessary for plant development, water transport and defense (Rogers and Campbell, 2004), those same properties make it a physical and chemical barrier preventing access of hydrolytic enzymes to cellulose (Benjamin *et al.*, 2013; Chen and Dixon, 2007; Zhao *et al.*, 2012). Therefore pretreatment of lignocellulosic material to remove lignin is currently a crucial yet costly step in the enzymatic conversion of polysaccharides to simple sugars (Yuan *et al.*, 2008).

1.3 Lignin biosynthesis

The evolution of the lignin biosynthesis pathway, part of the much larger phenylpropanoid biosynthesis pathway, is one of the key steps allowing plants to occupy terrestrial environments (Boerjan *et al.*, 2003; Weng and Chapple, 2010). Lignin biosynthesis and deposition provides land plants the structural rigidity to support themselves and the hydrophobic nature of xylem lignin polymers allows for efficient water transportation (Ma, 2007; Weng and Chapple, 2010). Lignin monomers are formed through the phenylpropanoid pathway by the deamination of phenylalanine (Figure 1.1; Boerjan *et al.*, 2003; Liu, 2012). A combination of successive hydroxylation and methylation reactions modify the aromatic ring, and esterification and reduction of the external carboxylic group result in p-coumaryl, coniferyl and sinapyl monolignols being produced (Boerjan *et al.*, 2003; Halpin, 2004; Liu, 2012). Respective polymerisation of these monolignols produces *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) phenylpropanoid monomers which comprise the lignin polymer (Boerjan *et al.*, 2003).

The genetic regulation of cell wall biosynthesis is complex and well-reviewed (Gray *et al.*, 2012; Wang and Dixon, 2012) as is the assembly of the lignin polymer (Buanafina, 2009; Liu, 2012). Many genes identified have multiple levels of regulation and feedback with the number of genes identified in cell wall formation and maintenance (or regulation thereof)

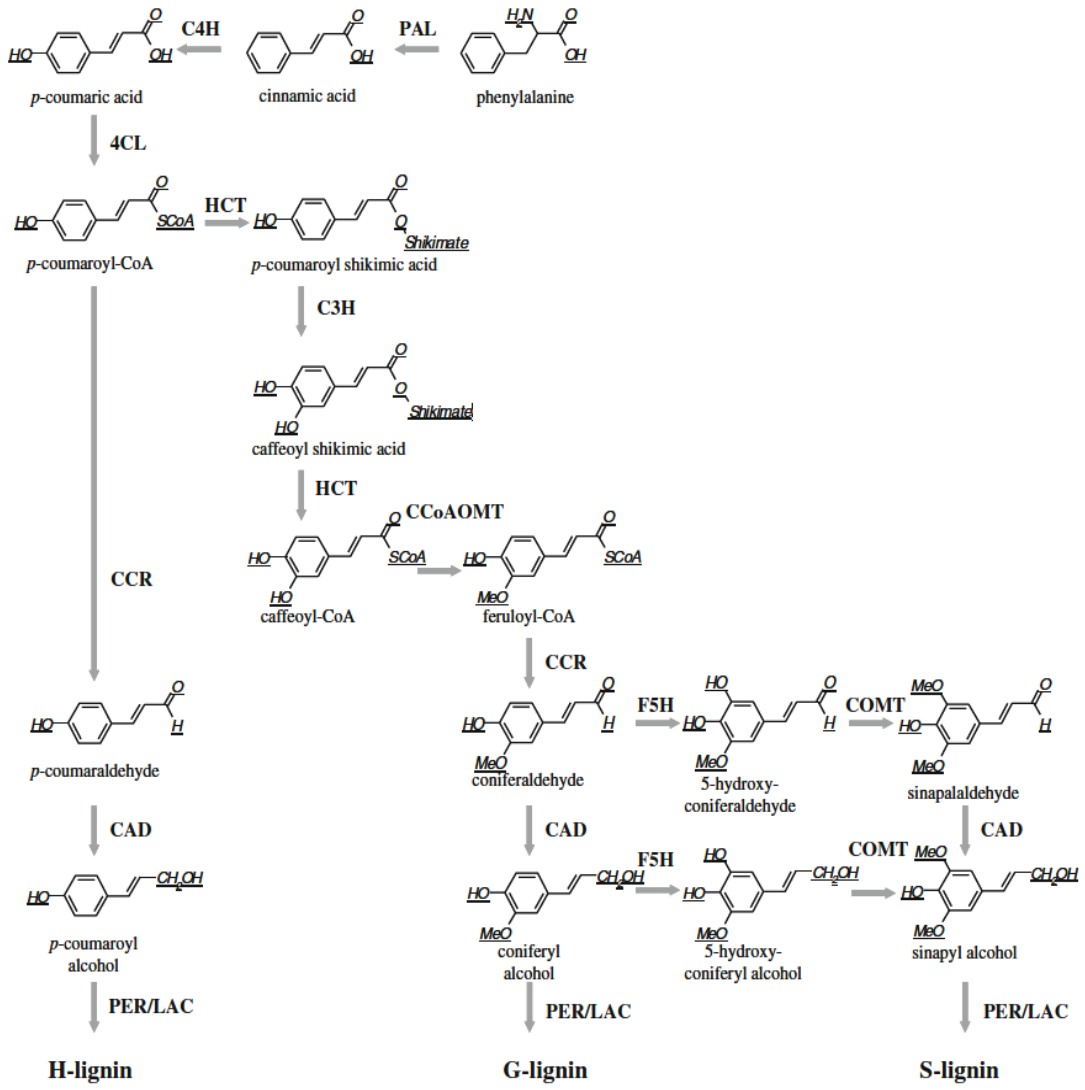


FIGURE 1.1: The biosynthesis of lignin monomers H, G and S begins with the deamination of phenylalanine and requires successive reactions by ten individual enzymes. Abbreviations: Phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate-CoA ligase (4CL), cinnamoyl-CoA reductase (CCR), hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase (HCT), coumarate 3-hydroxylase (C3H), caffeoyl CoA 3-O-methyltransferase (CCoAOMT), ferulate 5-hydroxylase (F5H), caffeic acid 3-O-methyltransferase (COMT), cinnamyl alcohol dehydrogenase (CAD), peroxidase (PER), and laccase (LAC) (Hisano *et al.*, 2009). Ten genes (PAL, C4H, 4CL, CCR, HCT, C3H, CCoAOMT, F5H, COMT and CAD) are considered necessary for the synthesis of p-coumaryl, coniferyl, and sinapyl alcohols, or monolignols (Hisano *et al.*, 2009; Weng *et al.*, 2008). Subsequent dehydrogenative polymerisation of these monolignols by PER and LAC genes form p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) lignin monomers (Hisano *et al.*, 2009; Weng *et al.*, 2008).

continuing to grow (Zhao and Dixon, 2011). New evidence suggests cell wall formation also has an additional level of miRNA regulation (Li and Lu, 2014; Ong and Wickneswari, 2012). In contrast, the methods by which lignin monomers are transported to the cell wall for polymerisation are currently not well understood (Liu, 2012).

Numerous studies have shown it possible to reduce lignin content in plants through the down-regulation of genes in the lignin biosynthetic pathway (reviewed in Li *et al.*, 2008). In general, there is a positive correlation between reduction of lignin biosynthetic gene transcription levels and an overall reduction of lignin content. It is hypothesised that the down-regulation of key genes controlling lignin biosynthesis in sugarcane will reduce the rate of lignin deposition, resulting in lignin-reduced sugarcane bagasse. In turn, this is expected to reduce the degree of pretreatment required, thus reducing the overall bioethanol production costs. It is further hypothesised that the down-regulation of specific genes within the lignin biosynthesis pathway could alter the lignin polymer composition by changing the H:G:S monomer ratio. Such an outcome may also improve pretreatment efficiency and therefore warrants further exploration. Previous research has found an increase in fermentable sugars released from the cell wall of plants with reduced lignin content (Benjamin *et al.*, 2013; Chen and Dixon, 2007; Sonbol *et al.*, 2009), providing further support for this avenue of research.

1.4 Sugarcane bagasse as a source of lignocellulosic biomass

Sugarcane is a C4 perennial grass grown in tropical and sub-tropical climates (Moore, 1987; Osabe *et al.*, 2009; Somerville *et al.*, 2010). Sugarcane accumulates and stores high concentrations of sucrose in the stem internodes which can be utilised for food or fuel production (de Souza *et al.*, 2013; Jung *et al.*, 2012; Osabe *et al.*, 2009). The vast majority of the world's sugar is produced from sugarcane, making it an important and valuable commercial crop plant (Jackson, 2013; Suprasanna *et al.*, 2011). Sugarcane grows rapidly and generates an abundance of lignocellulosic biomass (de Souza *et al.*, 2014; Lakshmanan *et al.*, 2005), which could potentially be utilised for biofuel production (Canilha *et al.*, 2012; Yuan *et al.*, 2008). An additional benefit of bagasse is that once harvested, sugarcane is transported to a central location (Pandey *et al.*, 2000), thus reducing biomass transport costs which could otherwise represent a significant cost in bioethanol production (Hill *et al.*, 2006). Large scale production of additional commercially valuable products from harvested material, such as bioethanol from lignin-reduced bagasse, will add value to a currently

underutilised product, thus increasing the sustainability and competitiveness of the global sugarcane industry (Lakshmanan *et al.*, 2005), an important contributor to the Australian economy (Canegrowers, 2010).

Although traditional breeding techniques have contributed to commercially sought improvements in sugarcane (Suprasanna *et al.*, 2011), and reducing lignin content through targeted breeding programs is an option as lignin content varies throughout cultivars (Benjamin *et al.*, 2014; Masarin *et al.*, 2011), the lengthy timeframe of traditional breeding programs limits the development of new sugarcane cultivars with new or improved agronomic traits (Suprasanna *et al.*, 2011). When combined with the large and complex polyploid genome of sugarcane (Grivet and Arruda, 2001), these unique characteristics highlight sugarcane as a suitable candidate for biotechnological improvement (Canilha *et al.*, 2012; de Souza *et al.*, 2013; Suprasanna *et al.*, 2011).

An alternative to traditionally bred, lignin-reduced sugarcane is the potential development of dedicated sugarcane varieties for energy and biofuel production, namely 'energycane' (Botha and Moore, 2014; Chong and O'Shea, 2012). These new varieties would have increased fibre or increased biomass as the focus of breeding, rather than the traditional focus of increased sucrose content (Botha and Moore, 2014; Chong and O'Shea, 2012). Additionally, energycane could potentially provide biomass year-round, a requirement of bioenergy refineries (Botha and Moore, 2014; Chong and O'Shea, 2012), as well as contributing to the alleviation of the food versus fuel debate (Yuan *et al.*, 2008). However, energycane breeding programs would still face the same obstacles as breeding for high sucrose content (Grivet and Arruda, 2001; Suprasanna *et al.*, 2011).

While sugarcane is an important crop for Australia, the impacts and benefits of second generation bioethanol production from sugarcane bagasse will be globally realised. Sugarcane currently ranks within the top ten most cultivated crops globally, being grown in nearly 200 countries (Botha and Moore, 2014; Suprasanna, *et al.*, 2011), and the harvesting of sugarcane generates a greater biomass than any other crop (Jackson, 2013). Brazil is currently the largest sugarcane producer in the world, followed by India, China and Thailand (Botha and Moore, 2014). In 2010, Brazil harvested 719 million tonnes of sugarcane, resulting in 101 million tonnes of bagasse, whereas Australia harvested 32 million tonnes of sugarcane, resulting in four million tonnes of bagasse (Botha and Moore, 2014). Based on these figures, it is predicted Brazil could produce between 11.1 - 27.2

billion litres of bioethanol from the resulting bagasse, which dwarfs Australia's predicted bioethanol yield of between 0.5 - 1.2 billion litres (Botha and Moore, 2014). Despite these figures, the production of second generation bioethanol in Australia will benefit both the sugarcane industry and the Australian economy, as well as having a positive influence on the reduction of oil-related climate change, with the development of biofuel technology having additional economic value.

1.5 Research aims

Lignin biosynthesis has been widely studied in some plant species but it has only recently been explored in sugarcane. In brief, four studies have been published concerning gene expression patterns in maturing sugarcane, which when combined, provide information on eight genes of the lignin biosynthesis pathway (PAL, C4H, C3H, CCoAOMT, HCT, F5H, CAD and COMT; Figure 1.1) (Casu *et al.,* 2004; Casu *et al.,* 2007; Kolahi *et al.,* 2013; Papini-Terzi *et al.*, 2009). Only one study has been published on the deposition of lignin, cellulose and hemicellulose in the secondary cell walls of developing sugarcane (Lingle and Thomson, 2012) and only two publications of the application of biotechnology to alter lignin deposition in genetically modified sugarcane (Jung *et al.*, 2012; Jung *et al.*, 2013). These studies are further discussed in Section 2.1. The limited number of publications suggest lignin biosynthesis and its manipulation in sugarcane are still far from being fully understood. This thesis focuses on decreasing this knowledge gap through the following three specific research aims:

1. Development of a lignin biosynthesis profile in sugarcane stem tissue by complementing lignin biosynthesis gene expression level data with cell wall compositional data in stem internodes of increasing maturity to further the understanding of monocot lignin formation.

2. Generation and analysis of transgenic sugarcane expressing two maize R2R3 MYB transcription factors known to down-regulate lignin biosynthesis in other plant species, with a goal of reducing the overall lignin content and improving saccharification after enzymatic hydrolysis.

3. Generation and analysis of transgenic sugarcane expressing RNA interference (RNAi) constructs targeting caffeoyl CoA 3-O-methyltransferase (CCoAOMT), ferulate 5-

hydroxylase (F5H) or caffeic acid 3-O-methyltransferase (COMT) genes of the lignin biosynthetic pathway to further explore the potential benefits of altering the H:G:S monomer ratio of the lignin polymer for liberation of glucose after enzymatic hydrolysis.

The outcomes of this research will contribute to the current knowledge surrounding cell wall formation and specifically lignin biosynthesis in sugarcane including the phenotypic effects resulting from manipulation of the lignin biosynthesis pathway.

Chapter 2

Literature Review

The multifaceted nature of lignin and its key role in the cell wall has led to an increasing volume of lignin-based articles published across a broad spectrum of scientific fields. This includes analysis of lignin biosynthesis and layers of regulation; the chemistry of the deposited lignin polymer; the extraction of and uses for lignin polymers; and different pretreatment approaches to improve biofuel production from lignocellulosic biomass. The scope of this literature review is focused on published research relevant to the three aims of this thesis outlined in Chapter 1. To briefly reiterate, the papers discussed will contain information and findings that contribute to the profiling of lignin monomer biosynthesis and deposition of the lignin polymer; regulatory control over lignin biosynthesis by MYB transcription factors; alterations to lignin biosynthesis through RNAi targeting of genes within the lignin biosynthesis pathway; and the saccharification of lignin altered plants, whether by GM or non-GM methods.

2.1 Expression profiles of lignin biosynthesis genes and cell wall development

Expression profiling of the genes involved in an individual pathway is an important aspect to gaining insight into plant development. In the case of lignin biosynthesis, a focused analysis of the expression patterns of genes within the pathway in tissues of different types or of different levels of maturity can provide information into the rate and location of lignin deposition. Expression profiles of lignin biosynthesis genes have now been established, or can be inferred from results in a number of different plant species and have been used to identify homologues of lignin biosynthesis pathway genes, as well as comparing the expression levels of lignin biosynthesis genes in different tissues. The compositions of cell wall components have also been profiled in a range of plant species; however, an expression profile, such as the one proposed in this thesis, which combines lignin biosynthesis pathway gene expression levels with cell wall component deposition in

sugarcane is not currently available. As interest in developing renewable biofuels from lignocellulosic biomass increases, knowledge surrounding lignin formation in sugarcane will become increasingly relevant if the bagasse produced by the Australian or global sugarcane industry is utilised for the production of second generation bioethanol.

2.1.1 Identification of lignin biosynthesis pathway gene homologues

Koutaniemi *et al.* (2007) highlighted the most likely lignin biosynthesis gene pathway in Norway spruce by comparing transcript levels of homologues of each gene involved in lignin biosynthesis between 1 year and 40 year old trees. Shang *et al.* (2012) employed semi-quantitative methods to profile seven PAL homologues in cucumber (*Cucumis sativus*) in seven different tissues. RT-PCR was performed on RNA from each tissue followed by agarose gel electrophoresis with imaging software used to measure band intensity (Shang *et al.*, 2012). The authors found that some homologues were expressed in all tissues whereas other homologues appeared tissue specific (Shang *et al.*, 2012). Overall the highest levels of PAL expression were seen in root, stem and female flower tissue, with less expression detected in cotyledons, fruit and male flower tissue and the least amount of expression detected in leaf tissue.

Shi *et al.* (2010) identified 95 putative gene sequences in poplar (*Populus trichocarpa*) representing the 10 lignin biosynthesis genes (Figure 1.1) with potential involvement in phenylpropanoid biosynthesis. Each biosynthesis gene had between three and 25 homologues. qPCR quantified the involvement of each homologue of each gene in four different tissues: differentiating xylem; differentiating phloem; shoot tip; and fully expanded leaf. Of the 10 lignin biosynthesis genes, PAL was the only one whose homologues all showed detectable expression in various tissues. Of the nine other genes there was a clear pattern of between one and three homologues showing expression in the tissues examined with many homologues showing no expression. Although some expression could be seen in other tissues, a very clear trend was seen for xylem preferred expression. That the majority of homologues showed no detectable expression may be due to tissue specifity in a tissue other than those examined; inducible expression not being triggered; having unstable transcript; or being pseudogenes (Shi *et al.*, 2010).

2.1.2 Profiling of lignin biosynthesis gene expression and lignin deposition in dicotyledonous species

Anterola *et al.* (2002) measured the expression response of lignin biosynthesis genes in *Pinus taeda* cells suspensions before and after spiking the suspensions with phenylalanine. PAL, 4CL, CCoAOMT and CCR had obvious increases in expression levels while CAD had a minimal increase in expression and C4H and C3H showed no change in expression levels (Anterola *et al.*, 2002). CAD expression was compared with lignin content in *Ginkgo biloba* stem tissue aged 1, 2, 3 and 4 years (Cheng *et al.*, 2013). CAD expression was highest in one year old (yo) tissue before decreasing in 2yo and plateauing in the 3yo and 4yo stem tissue which correlates with lignin content being lowest in the youngest tissue and steadily rising as the stem tissue matured (Cheng *et al.*, 2013). CAD expression detected in roots was slightly higher than the highest stem expression levels (in the youngest stem tissue) (Cheng *et al.*, 2013).

Next-generation sequencing has also led to identification and profiling of lignin biosynthesis gene expression patterns. Firon *et al.* (2013) sequenced and compared initial storage roots (ISR) and fibrous roots (FR) from sweet potato (Ipomea spp; *Georgia Jet*) as the development of ISRs from FRs is a key step for tuber production. Analysis of the read count of contigs found that FR had a higher number of reads than ISR for PAL, C4H, 4CL, HCT, CCoAOMT, CCR, COMT and CAD. This indicated a higher number of transcripts of lignin genes in FR than ISR which correlates with FR containing more lignin than ISR (Firon *et al.*, 2013). qPCR validation confirmed that the relative expression levels of 4CL, CCoAOMT and CAD were lower in ISR when compared with FR (Firon *et al.*, 2013). Chinese fir (*Cunninghamia lanceolata*) has also undergone next-generation sequencing with a focus on lignin biosynthesis genes (Huang *et al.*, 2012). Analysis and qPCR validation confirmed that PAL, C3H, CCoAOMT, COMT, CCR and CAD showed similar expression patterns with the highest expression occurring in xylem, followed by lignifying stem, bark and the least expression in non-lignifying stem (Huang *et al.*, 2012). C4H also showed a similar pattern but had very low expression levels overall and 4CL had little expression detected in any tissues (Huang *et al.*, 2012).

2.1.3 Profiling of lignin biosynthesis gene expression and lignin deposition in monocotyledonous species

Ma (2007) specifically examined CCR in two different wheat (*Triticum aestivum*) cultivars; lodging resistant H4564 and lodging sensitive C6001. Stem, leaf and root tissues were collected at three different developmental stages, youngest to most mature being elongation, heading, and milky. Using semi-quantitative imaging software, Northern Blot analysis found CCR expression was highest in stem tissue, lower in leaf tissue and not detectable in root tissue (Ma, 2007). During maturation H4564 showed strong CCR expression throughout the three developmental stages. C6001 showed strong CCR expression in the elongation stage, but only weak expression in the two latter stages (Ma, 2007). Both cultivars showed a decrease in CCR expression when plants entered heading stage before an increase in milky stage (Ma, 2007). The increase in milky stage CCR expression was more pronounced for H4564 (a 117% increase over elongation levels) than C6001, which only reached 34% of elongation expression levels (Ma, 2007).

When the Klason method was utilised to determine the lignin content in elongation, heading and milky stem tissue it was found that lignin content increased with tissue maturation however C6001 lignin levels only reached 73% of H4564 levels, which correlates with their respective susceptibility and resistance to lodging (Ma, 2007). Acid soluble lignin contents decreased with tissue maturation and were similar for both cultivars (Ma, 2007). Ma (2007) concluded increased CCR expression may correlate with increased lignin deposition, though not in proportion to the expression levels of CCR. Cell wall components in developing maize have also been profiled by Jung and Casler (2006). Over a period of three months during the growing season, maize plants were sampled 10 times and secondary cell wall components were quantified. Lignin content decreased during the first eight days before increasing to peak at five weeks and five days, and plateauing thereafter (Jung and Casler, 2006). Glucose content increased for five weeks before stabilising whereas hemicellulose (xylose and arabinose) content decreased for the first three weeks before plateauing (Jung and Casler, 2006).

2.1.4 Expression profiling of lignin biosynthesis genes in sugarcane

The expression profile of lignin biosynthesis in sugarcane is relatively unexplored, however, a number of groups have identified the expression profiles of some lignin biosynthesis

genes in related studies. Casu *et al.* (2004) found that PAL, CCoAOMT and COMT expression increased in maturing stem (internodes 6-11) of sugarcane when compared to young stem (apical meristem and internodes 1-3). Further work by Casu *et al.* (2007), analysing sugarcane gene expression levels in three tissues of different maturity (young stem: internodes 1-3; maturing stem: internode 8; and mature stem: internode 20) using the Affymetrix GeneChip Sugarcane Genome Array resulted in the identification of 119 differentially expressed genes. Characterisation of 23 of the 119 genes was performed using qPCR with a focus on genes from the cellulose synthase and cellulose synthase-like families. Three of the remaining 96 genes were identified as CCoAOMT, HCT and CAD. In maturing tissue CCoAOMT and HCT expression was reduced whereas CAD expression increased. This may reflect the roles of CCoAOMT and HCT in cell growth and development and CAD having an ongoing role in cell wall maintenance (Casu *et al.*, 2007).

A microarray comparison of high Brix and low Brix sugarcane by Papini-Terzi *et al.* (2009) was conducted to assess genes differently expressed between the groups which may indicate a role in sucrose accumulation. Among the 117 genes identified as differing in expression levels were five lignin genes: F5H, C4H, COMT, PAL and C3H. A comparison of immature tissue (internode 1) and mature tissue (internode 9) revealed F5H and COMT expression increased with tissue maturity, C4H and C3H showed no difference and PAL homologues showed no change or decreased (Papini-Terzi *et al.*, 2009). Kolahi *et al.* (2013) quantified PAL expression in different sugarcane tissues (leaf, sheath, stem internodes 7-8 and root) at different developmental stages: germination (1-3 months), tillering (3-7 months), grand growth (7-11 months) and maturation (11-15 months). Roots initially had low PAL expression increasing to peak in the tillering stage before steadily decreasing in grand growth and again in maturation whereas leaf and sheath showed steady expression of PAL over different stages (Kolahi *et al.*, 2013). Stem tissue was not available for germination stage. Some PAL expression was detected in the tillering plants with highest expression detected in grand growth before reducing slightly in maturation stage.

2.1.5 Cell wall composition profile during sugarcane development

To date, there has only been one paper published which focuses on the deposition of lignocellulosic components during sugarcane growth and development. Lingle and Thomson (2012) developed a cell wall composition profile in sugarcane by quantifying internode composition during two consecutive growing seasons whilst also quantifying

juice sugar content (sucrose) of the same internodes for comparison. Two different experiments were conducted. The first involved marking internode 1 on 100 sugarcane plants (20 plants x five cultivars) and harvesting one plant per cultivar at weeks 1, 2, 3, 4, 6, 8, 10 and 12 and the second experiment involved harvesting odd-numbered internodes between 1 and 11 (inclusively) from four plants of a single cultivar in July and September of the same year, and repeated again the following year. Overall findings confirmed different cultivars have different cell wall composition, but similar trends in component deposition. The rate of lignin deposition was highest in young internode tissue before increasing in smaller increments as the plants matured with lignin content continuing to increase as the plants developed (Lingle and Thomson, 2012). Cellulose content was lowest in young tissue and the rate of deposition peaked between internodes three and five, after which the rate of deposition would either plateau or slowly decline (Lingle and Thomson, 2012). Hemicellulose content showed the opposite trend to cellulose biosynthesis. Hemicellulose content was highest in young tissue and reduced until internode three to five before plateauing for remaining plant development (Lingle and Thomson, 2012).

The published research discussed in this section provides the first insights into cell wall development across a range of plant species. Although the publications are limited in number, there is a general consensus in their findings. Lignin and cellulose content has been found to increase during plant development with the rate of deposition being greater in early development than more mature tissue (Cheng *et al.*, 2013; Jung and Casler, 2006; Lingle and Thomson, 2012; Ma, 2007), whereas hemicellulose content is highest in young tissue and decreases as plant tissue matures (Jung and Casler, 2006; Lingle and Thomson, 2012). Research into the lignin biosynthesis found homologues exist for genes within this pathway (Koutaniemi *et al*., 2007; Shang *et al.*, 2012; Shi *et al.*, 2010) and the levels of gene expression was found to be higher in lignifying tissues when compared with tissues that undergo little or no lignification (Firon *et al.*, 2013; Huang *et al.*, 2012; Ma, 2007).

Within sugarcane, data has been published on the expression patterns of eight lignin biosynthesis genes (Figure 1.1). Five of these eight genes have only been mentioned in a single study. In maturing tissue, when compared with younger tissue, HCT expression was found to decrease (Casu *et al.*, 2007), CAD (Casu *et al.*, 2007), F5H (Papini-Terzi *et al.*, 2009) expression increased, and C3H and C4H show no differences in expression levels (Papini-Terzi *et al.*, 2009). The expression patterns of three sugarcane lignin pathway genes (PAL, CCoAOMT and COMT) have been reported in more than one study, with mixed findings.

COMT expression was found to increase in maturing sugarcane stem (Casu *et al.*, 2004; Papini-Terzi *et al.*, 2009). Both increased (Casu *et al.*, 2004), and decreased (Casu *et al.*, 2007) CCoAOMT expression has been reported as sugarcane tissue matures. The expression of PAL in maturing sugarcane has been found to increase (Casu *et al.*, 2004; Kolahi *et al.*, 2013), decrease (Papini-Terzi *et al.*, 2009), or show no change (Papini-Terzi *et al.*, 2009) when compared with younger tissue.

While these studies begin to shed light on sugarcane lignin biosynthesis, there is currently no literature which has conducted gene expression and lignocellulosic compositional analysis on the same tissue samples which would allow for direct comparison. The development of an expression profile of lignin formation in sugarcane which combines gene expression data with cell wall compositional data will further improve knowledge in this area. By providing insight to both spatial and temporal expression of lignin biosynthetic genes, the targeting or manipulation of these genes within different tissues can be performed with better specificity to obtain the desired plant genotype or phenotype.

2.2 MYB transcription factor regulation of lignin biosynthesis

2.2.1 R2R3 MYB transcription factors

In addition to the genes directly involved in the production of H, G and S monomers, lignin biosynthesis has also been found to be regulated by a number of transcription factor families, including the MYB family. Transcription factors are sequence-specific DNA-binding proteins that regulate gene expression through activation or repression of mRNA transcription levels (Riechmann *et al.*, 2000; Xiong *et al.*, 2005). As their regulatory influence can range from a single gene to entire biological pathways they are anticipated to play a major role in the future of GM crops (Ambawat *et al.*, 2013; Century *et al.*, 2008). MYB transcription factors were first identified in the avian myeloblastosis virus (Klempnauer *et al.*, 1982) and the first plant MYB gene was isolated from *Zea mays* (Paz-Ares *et al.*, 1987). MYB transcription factors now represent one of the largest transcription factor families identified in plants (Du *et al.*, 2012b; Wilkins *et al.*, 2009) with known roles in both developmental and biochemical pathway regulation (Patzlaff *et al.*, 2003b). A thorough review of the diversity of MYB gene functions can be found in Ambawat *et al.* (2013).

R2R3 MYB genes are the most common form found in plants, and contain two N-terminal DNA binding domains (R2 and R3) consisting of characteristic imperfect repeats of a helix-turn-helix motif (Patzlaff *et al.*, 2003a) with conserved tryptophan residues (Ogata *et al.*, 1995; Stracke *et al.*, 2001) that are thought to aid sequence-specific DNA binding (Gabrielsen *et al.*, 1991). This binding domain and its target sequence have been well characterised (Ogata *et al.*, 1996; Ogata *et al.*, 1995; Ogata *et al.*, 1994; Prouse and Campbell, 2012; Prouse and Campbell, 2013). The less conserved C-terminal region is highly divergent (Jiang *et al.*, 2004) and is responsible for the regulatory function of the translated protein (Bedon *et al.*, 2007; Jiang *et al.*, 2004; Legay *et al.*, 2007; Tamagnone *et al.*, 1998). Analysis of conserved motifs within this C-terminal region result in the identification of subgroup clusters. The total number of subgroups appears to differ between plant species and increases as genomic data sets improve and expand (Du *et al.*, 2012a; Du *et al.*, 2012b). The lignin down-regulating R2R3 MYB genes comprise subgroup G4 (Du *et al.*, 2012a; Fornalé *et al.*, 2010). Promoters of lignin biosynthetic genes contain conserved AC element motifs which act to limit lignification to xylem tissues (Bedon *et al.*, 2007; Patzlaff *et al.*, 2003b). The MYB R2R3 domain interacts with these AC elements, allowing MYB transcription factors to regulate promoter activity of lignin biosynthetic genes (Fornalé *et al.*, 2010; Lauvergeat *et al.*, 2002; Patzlaff *et al.*, 2003a; Patzlaff *et al.*, 2003b) thus making G4 MYB genes important targets for lignin modification.

2.2.2 MYB gene regulation of lignin biosynthesis

2.2.2.1 MYB gene directed up-regulation of lignin biosynthesis

A number of MYB transcription factors have been shown to promote lignin biosynthesis in various plant species through the analysis of gene expression patterns and transgenic expression studies (Table 2.1). Plants expressing these MYB transcription factors are generally characterised by increased and ectopic lignin deposition combined with increased expression of genes within the phenylpropanoid pathway.

*Antirrhinum majus Am*MYB305 increased reporter gene expression by activating the upstream *Phaseolus vulgaris* PAL promoter sequence in tobacco protoplasts (Sablowski *et al.*, 1994). Borevitz *et al.* (2000) identified an Arabidopsis gene (PAP1) whose overexpression resulted in increased lignin and anthocyanin production. Analysis found an increase in expression of PAL, but little change in expression levels of COMT and CCoAOMT.

TABLE 2.1: Summary of published results of the up-regulation of lignin biosynthesis by MYB genes

MYB	Source	Expressed in	Genes up-regulated	Promoter activated	Lignin content	S:G ratio	Reference
*Am*MYB305	*Antirrhinum majus*	*Nicotiana tabacum* protoplasts		PAL			(Sablowski *et al.*, 1994)
*At*MYB46	*Arabidopsis thaliana*	*Arabidopsis thaliana*	PAL, C4H, CAD, C3H, CCoAOMT, 4CL	PAL, CCoAOMT	Increased (OE) and decreased (DS)		(Kim *et al.*, 2012; Ko *et al.*, 2009; Zhong *et al.*, 2007)
*At*MYB58	*Arabidopsis thaliana*	*Arabidopsis thaliana*	All except F5H	4CL	Increased (OE) and decreased (DS/RNAi)	Increased	(Zhou *et al.*, 2009)
*At*MYB63	*Arabidopsis thaliana*	*Arabidopsis thaliana*	All except F5H	4CL	Increased (OE) and decreased (DS/RNAi)	Increased	(Zhou *et al.*, 2009)
*At*MYB83	*Arabidopsis thaliana*	*Arabidopsis thaliana*	4CL, CCoAOMT		Increased		(McCarthy *et al.*, 2009)
*At*MYB85	*Arabidopsis thaliana*	*Arabidopsis thaliana*			Increased (OE)		(Zhong *et al.*, 2008)
*At*MYB103	*Arabidopsis thaliana*	*Arabidopsis thaliana* mutants	F5H (down-regulated)		No change	Decreased	(Öhman *et al.*, 2013)
PAP1 (*At*MYB75/*At*MYB90)	*Arabidopsis thaliana*	*Arabidopsis thaliana; Salvia miltiorrhiza*	PAL, C4H, 4CL, CCR		Increased	No change	(Borevitz *et al.*, 2000; Kranz *et al.*, 1998; Zhang *et al.*, 2010)
*Eg*MYB2	*Eucalyptus gunnii*	*Nicotiana tabacum*	HCT, CCR, CAD, C3H, CCoAOMT, F5H, COMT	CCR, CAD		Increased	(Goicoechea *et al.*, 2005)
*Os*MYB42/85	*Oryza sativa*	*Oryza sativa*	CAD		Increased		(Hirano *et al.*, 2013)
*Os*MYB46	*Oryza sativa*	*Arabidopsis thaliana*	4CL		Increased		(Zhong *et al.*, 2011)
*Os*MYB55/61	*Oryza sativa*	*Oryza sativa*	CAD	CAD	Increased		(Hirano *et al.*, 2013)

TABLE 2.1: Summary of published results of the up-regulation of lignin biosynthesis by MYB genes (continued)

MYB	Source	Expressed in	Genes up-regulated	Promoter activated	Lignin content	S:G ratio	Reference
*Os*MYB58/63	*Oryza sativa*	*Oryza sativa*	CAD		Increased		(Hirano *et al.*, 2013)
*Os*MYB103	*Oryza sativa*	*Oryza sativa*	CAD		Increased		(Hirano *et al.*, 2013)
*Pg*MYB2	*Picea glauca*	*Picea glauca* compression wood	4CL, PAL, CAD, CCoAOMT		Increased		(Bedon *et al.*, 2007)
*Pg*MYB4	*Picea glauca*	*Picea glauca* compression wood	4CL, PAL, CAD, CCoAOMT		Increased		(Bedon *et al.*, 2007)
*Pg*MYB8	*Picea glauca*	*Picea glauca* compression wood	4CL, PAL, CAD, CCoAOMT		Increased		(Bedon *et al.*, 2007)
*Pp*MYB8	*Pinus pinaster*	*Pinus pinaster* stem protoplasts		PAL			(Craven-Bartle *et al.*, 2013)
*Pt*MYB1	*Pinus taeda*	*Nicotiana tabacum* NT-1 cells; *Picea glauca* (Moench) Voss	All except F5H (not measured)	PAL	Increased		(Bomal *et al.*, 2008; Patzlaff *et al.*, 2003b)
*Pt*MYB4	*Pinus taeda*	*Arabidopsis thaliana; Nicotiana tabacum*	C3H, CCoAOMT, COMT, CCR, CAD		Increased		(Newman *et al.*, 2004; Patzlaff *et al.*, 2003a)
*Pt*MYB8	*Pinus taeda*	*Picea glauca* (Moench) Voss	PAL, C4H, 4CL,C3H, CCR, CAD, HCT, COMT		Increased		(Bedon *et al.*, 2007; Bomal *et al.*, 2008)
*Pto*MYB216	*Populus tomentosa*	*Populus tomentosa*	PAL, 4CL, C3H, CCR		Increased	No change	(Tian *et al.*, 2013)
*Ptr*MYB2	*Populus trichocarpa*	*Arabidopsis thaliana*: plants and leaf protoplasts; *Populus alba x Populus tremula*	4CL, CCoAOMT	CCoAOMT, COMT	Increased		(Zhong *et al.*, 2013)

TABLE 2.1: Summary of published results of the up-regulation of lignin biosynthesis by MYB genes (continued)

MYB	Source	Expressed in	Genes up-regulated	Promoter activated	Lignin content	S:G ratio	Reference
*Ptr*MYB3	*Populus trichocarpa*	*Arabidopsis thaliana*: plants and leaf protoplasts	4CL, CCoAOMT	CCoAOMT, COMT	Increased		(McCarthy *et al.*, 2010; Zhong *et al.*, 2013)
*Ptr*MYB20	*Populus trichocarpa*	*Arabidopsis thaliana:* plants and leaf protoplasts	4CL, CCoAOMT	CCoAOMT	increased		(McCarthy *et al.*, 2010)
*Ptr*MYB21	*Populus trichocarpa*	*Arabidopsis thaliana*: plants and leaf protoplasts; *Populus alba x Populus tremula*	4CL, CCoAOMT	CCoAOMT, COMT	Increased		(Zhong *et al.*, 2013)
*Vv*MYB5a	*Vitis vinifera*	*Nicotiana tabacum*	C4H				(Deluc *et al.*, 2006)
*Zm*MYB46	*Zea mays*	*Arabidopsis thaliana*	4CL		Increased		(Zhong *et al.*, 2011)

OE: Over-expressed

DS: Dominant suppression

The mutant line had increased overall lignin content but no change in the S:G ratio. Comparisons against Arabidopsis databases found PAP1 to be identical to *At*MYB75 (Kranz *et al.*, 1998), with the exception of an earlier termination codon in *At*MYB75 resulting from a single base pair deletion and to have very close homology to Arabidopsis EST sequence later identified to be *At*MYB90 (Kranz *et al.*, 1998).

*At*PAP1 has since been expressed in *Salvia miltiorrhiza* to increase the natural production of salvianolic acid B (Sal B) (Zhang *et al.*, 2010), a water soluble phenolic acid used in traditional Chinese medicine to treat cardiovascular disease (Ho and Hong, 2011). The biosynthesis pathway for Sal B shares the genes PAL, C4H and 4CL with the lignin biosynthesis pathway (Figure 1.1; Zhang *et al.*, 2010). The increased activation of Sal B biosynthesis by the expression of *At*PAP1 led to increases in PAL, C4H, 4CL and CCR expression, with a decrease in COMT expression (Zhang *et al.*, 2010). This coincided with increases in lignin contents of transgenic plants when compared to controls (Zhang *et al.*, 2010). The H monomer content did not change between transgenics and controls, however there were increases in S and G monomer contents, and in some cases increased S:G ratio (Zhang *et al.*, 2010).

Over-expression of the endogenous MYB genes *At*MYB46 (Kim *et al.*, 2012; Ko *et al.*, 2009; Zhong *et al.*, 2007), *At*MYB58 and *At*MYB63 (Zhou *et al.*, 2009) and *At*MYB85 (Zhong *et al.*, 2008), and exogenous poplar *Ptr*MYB3 and *Ptr*MYB20 (McCarthy *et al.*, 2010) and pine *Pt*MYB4 (Newman *et al.*, 2004) in Arabidopsis all resulted in increased and ectopic lignin deposition highlighting their role in lignin biosynthesis. Further study found *At*MYB46 was also involved in regulating cellulose synthesis as well as lignin biosynthesis (Kim *et al.*, 2013). Ectopic lignin deposition was also found in stem tissue of transgenic poplar (*Populus tomentosa*) after isolation and over-expression of endogenous *Pto*MYB216 (Tian *et al.*, 2013). The increased lignin content coincided with changes to the expression patterns of lignin biosynthetic genes. PAL, 4CL, C3H and CCR showed increased expression compared to controls whereas F5H had decreased expression. No change was observed in expression levels of CCoAOMT, COMT or CAD (Tian *et al.*, 2013).

A number of spruce MYB genes *Pg*MYB2, *Pg*MYB4 and *Pg*MYB8 have been identified as being preferentially expressed in secondary xylem (Bedon *et al.*, 2007), suggesting a role in secondary

cell wall formation. Expression of the closely related *Pinus taeda* genes *Pt*MYB1 (Patzlaff *et al.*, 2003b) and *Pt*MYB8 (Bedon *et al.*, 2007) in spruce resulted in increased lignin deposition (Bomal *et al.*, 2008). *Pp*MYB8 from *Pinus pinaster* was transiently over-expressed in pine stem protoplasts in the presence of the GUS reporter gene driven by the *P. pinaster* PAL promoter resulting in induced GUS expression through the interaction with the AC element present in the PAL promoter sequence, highlighting *Pp*MYB8 as a regulator of PAL expression (Craven-Bartle *et al.*, 2013). The expression of grape *Vv*MYB5a (Deluc *et al.*, 2006), pine *Pt*MYB4 (Patzlaff *et al.*, 2003a) and Eucalyptus *Eg*MYB2 (Goicoechea *et al.*, 2005) in tobacco led to increased levels of lignin biosynthesis gene transcription levels as well as increased lignin deposition. Loss of function *At*MYB103 Arabidopsis had reduced F5H expression which led to a reduction in S monomers and an increase in G monomers without changing overall lignin content (Öhman *et al.*, 2013). This indicates *At*MYB103 is required for expression regulation of F5H (Öhman *et al.*, 2013).

Similar results have also been found in rice. Hirano *et al.* (2013) identified 123 transcription factors as potentially involved in secondary cell wall regulation in rice. Six transcription factors were selected for further analysis to validate the initial findings which included four rice (*Oryza sativa*) MYB genes, namely *Os*MYB58/63, *Os*MYB42/85, *Os*MYB55/61 and *Os*MYB103 (Hirano *et al.*, 2013). The over-expression of each of these genes individually resulted in plants having increases in CAD expression and increased lignin contents. Despite these common findings phenotypic variations were observed. The height of *Os*MYB58/63 expressing plants was comparable to that of controls while *Os*MYB42/85 plants were mildly dwarfed and *Os*MYB55/61 and *Os*MYB103 plants both displayed severe dwarfing of regenerated and acclimated plants (Hirano *et al.*, 2013). The *Os*MYB55/61 gene was found to increase expression of a reporter gene driven by the CAD promoter region 7-fold over the control plants (Hirano *et al.*, 2013). This highlights the changes in CAD expression and lignin content may be in part due to the regulatory influence the MYB genes may exert over lignin biosynthetic pathway genes via their promoters. These same rice MYBs were also knocked down via multiple strategies (Hirano *et al.*, 2013). *Os*MYB42/85, *Os*MYB103 and *Os*MYB55/61 knockdown lines showed less CAD expression with *Os*MYB103 and *Os*MYB55/61 knockdown lines also showing a reduction in lignin content (Hirano *et al.*, 2013). *Os*MYB42/85 and *Os*MYB103 lines were severely dwarfed whereas *Os*MYB55/61 lines displayed a less severe phenotype. *Os*MYB58/63

lines were again comparable to controls in height with no dwarfing observed (Hirano *et al.*, 2013).

Many MYB genes appear to play redundant roles. In transgenic Arabidopsis, dominant repression of *At*MYB46 resulted in decreased cell wall deposition and thickness whereas *At*MYB46 knockout lines showed no phenotypic differences to controls (Zhong *et al.*, 2007). Dominant repression of *At*MYB46 also repressed the homologues of *At*MYB46 whereas knockout of *At*MYB46 did not, thus their redundancy countered any negative effects potentially derived from the knocking out of *At*MYB46 (Zhong *et al.*, 2007). Over-expression of *At*MYB83, one such homologue, caused ectopic lignin deposition, whereas silencing of *At*MYB83 did not result in any phenotypic differences (McCarthy *et al.*, 2009). When *At*MYB46/*At*MYB83 double knockout Arabidopsis was generated, the plants displayed very stunted growth with little to no cell wall thickening in the leaves and roots (McCarthy *et al.*, 2009). These findings suggest that *At*MYB46 and *At*MYB83 act redundantly for the regulation of cell wall deposition.

Functional conservation of MYB transcription factors has been documented across several plant species. Homologues of *At*MYB46/83 have been cloned from rice (*Os*MYB46), maize (*Zm*MYB46) (Zhong *et al.*, 2011) and from poplar (*Ptr*MYB2 and *Ptr*MYB21) (Zhong *et al.*, 2013). In *At*MYB46/83 double knockout Arabidopsis all four MYB homologues, regardless of plant of origin, were able to restore normal function and phenotype to the mutant plant (Zhong *et al.*, 2011; Zhong *et al.*, 2013). When the cloned MYB homologues were individually over-expressed in wild type Arabidopsis these plants were found to display ectopic lignin deposition and increased cellulose levels in the cell wall (Zhong *et al.*, 2011; Zhong *et al.*, 2013). *Ptr*MYB3 (McCarthy *et al.*, 2010) and *Ptr*MYB21 (Zhong *et al.*, 2013) have also been characterised in transgenic poplar (Zhong *et al.*, 2013). Dominant repression of either of these genes resulted in poplar of reduced height and secondary cell wall thickening whereas over-expression of these genes individually increased the lignin content in regenerated plants (Zhong *et al.*, 2013).

2.2.2.2 MYB gene directed down-regulation of lignin biosynthesis

A number of MYB transcription factors have been shown to down-regulate the expression of genes within the lignin biosynthesis pathway which in turn correlates strongly with a reduction

in lignin deposition during cell wall formation (Table 2.2). The first evidence of R2R3 MYB transcription factors down-regulating lignin biosynthesis came after *Antirrhinum* MYB genes *Am*MYB308 and *Am*MYB330 were expressed in tobacco (Tamagnone *et al.*, 1998). *Am*MYB308 was found to down-regulate expression of lignin genes C4H, 4CL and CAD and *Am*MYB330 expressing plants had reduced 4CL expression when compared to wild type controls. Phenotypically 48% of *Am*MYB308 and 29% of *Am*MYB330 showed slow and stunted growth with leaves presenting with general chlorosis with necrotic patches (Tamagnone *et al.*, 1998). *Am*MYB308 plants had reduced S and G units, with an increase in S:G ratio, along with a 17.5% decrease in total lignin content of young stem tissue and a 56.5% lignin reduction in mature stem tissue (Tamagnone *et al.*, 1998). The lignin composition was not reported for *Am*MYB330 plants.

Expression of Eucalyptus *Eg*MYB1 in tobacco resulted in the repression of promoter activity of the CCR and CAD genes (Legay *et al.*, 2007). *Eg*MYB1 has since been expressed in Arabidopsis and poplar resulting in plants with reduced height and diameter (Legay *et al.*, 2010). Further analysis found these plants had significantly less lignin than controls but their S:G monomer ratio was unchanged (Legay *et al.*, 2010). Seven and nine different lignin biosynthetic genes had reduced expression in poplar and Arabidopsis respectively (Legay *et al.*, 2010). *Cm*MYB1 isolated from *Chrysanthemum morifolium* has also been expressed in transgenic Arabidopsis (Zhu *et al.*, 2013). It was found that *Cm*MYB1 expression reduced mRNA levels of COMT and CAD greatly, and C4H, 4CL, HCT, CCR and F5H to a lesser extent. The lignin content of the transgenic plants was reduced, and lignin monomer composition was altered in favour of G monomers.

Over-expression of *At*MYB4 in tobacco resulted in reduced expression of C4H, 4CL and CAD (Jin *et al.*, 2000), while in Arabidopsis, *At*MYB4 over-expression reduced C4H and 4CL expression (Jin *et al.*, 2000). In both lines PAL, F5H and COMT showed no change in expression levels whereas CCoAOMT showed an increase in expression (Jin *et al.*, 2000). In *At*MYB4 loss-of-function Arabidopsis, C4H showed increased levels of expression whereas CCoAOMT showed reduced expression (Jin *et al.*, 2000) indicating *At*MYB4 may have a more complex relationship with lignin biosynthesis than just gene down-regulation. *At*MYB4 was also found to repress reporter gene expression when driven by *At*C4H or *At*4CL promoters in Arabidopsis protoplasts

TABLE 2.2: Summary of published results of the down-regulation of lignin biosynthesis by MYB genes

MYB	Source	Expressed in	Genes down-regulated	Promoter repressed	Lignin content	S:G ratio	Reference
*Am*MYB308	*Antirrhinum majus*	*Nicotiana tabacum*	C4H, 4CL, CAD		Decreased	Increased	(Jackson *et al.*, 1991; Tamagnone *et al.*, 1998)
*Am*MYB330	*Antirrhinum majus*	*Nicotiana tabacum*	4CL				(Jackson *et al.*, 1991; Tamagnone *et al.*, 1998)
*At*MYB4	*Arabidopsis thaliana*	*Arabidopsis thaliana*: plants and protoplasts; *Nicotiana tabacum*	C4H, 4CL, CAD	C4H, 4CL			(Jin *et al.*, 2000)
*At*MYB68	Loss-of-function *Arabidopsis thaliana*				Increased		(Feng *et al.*, 2004)
*Cm*MYB1	*Chrysanthemum morifolium*	*Arabidopsis thaliana*	COMT, CAD, C4H, 4CL, HCT, CCR, F5H		Decreased	Reduced	(Zhu *et al.*, 2013)
*Eg*MYB1	*Eucalyptus gunnii*	*Nicotiana benthamiana; Arabidopsis thaliana; Populus tremula x P.alba*	All except HCT	CCR, CAD	Decreased	Unchanged	(Legay *et al.*, 2007; Legay *et al.*, 2010)
*Ll*MYB1	*Leucaena leucocephala*	*Nicotiana tabacum*	PAL, C4H, 4CL		Decreased		(Omer *et al.*, 2013)
P1/Ufo1	*Zea mays*	*Zea mays*	CCoAOMT, COMT		Decreased		(Robbins *et al.*, 2013)
*Ptt*MYB21a	*Populus tremula* L. x *tremuloides* Michx						(Karpinska *et al.*, 2004)
*Pv*MYB4	*Panicum virgatum*	*Arabidopsis thaliana* protoplasts; *Nicotiana tabacum*; *Panicum virgatum*	All ten genes	PAL, CCoAOMT	Decreased	Increased	(Shen *et al.*, 2012; Shen *et al.*, 2013)
*Ta*MYB4	*Triticum aestivum*	*Nicotiana tabacum*	CAD, CCR		Decreased	Increased	(Ma *et al.*, 2011)
*Zm*MYB31	*Zea mays*	*Arabidopsis thaliana*	4CL, F5H, COMT, C3H	COMT	Decreased	Unchanged	(Fornalé *et al.*, 2010; Fornalé *et al.*, 2006)
*Zm*MYB42	*Zea mays*	*Arabidopsis thaliana*	4CL, F5H, COMT, CAD, PAL, C4H, HCT	COMT	Decreased	Reduced	(Fornalé *et al.*, 2006; Sonbol *et al.*, 2009)

(Jin *et al.*, 2000). Reporter gene expression driven by *At*CCoAOMT was unchanged by *At*MYB4, indicating that its increased expression may not be directly caused by *At*MYB4 (Jin *et al.*, 2000). The expression of *Ta*MYB4 (from wheat) in transgenic tobacco caused a decrease in CAD and CCR expression but not in COMT (Ma *et al.*, 2011). This led to a reduction in lignin content with an increase in S:G ratio when compared to controls (Ma *et al.*, 2011). Over-expression of *Leucaena leucocephala* (*Ll*MYB1) in tobacco resulted in decreased transcript levels of PAL, C4H and 4CL and plants which were stunted with reduced lignin levels (Omer *et al.*, 2013).

The *maize pericarp color1* (P1) gene is a MYB transcription factor whose expression can be influenced by the dominant modifying factor *Unstable factor for orange1* (UFO1) gene (Chopra *et al.*, 2003; Robbins *et al.*, 2013). The expression of UFO1 increases the expression of P1 (Chopra *et al.*, 2003) but has been associated with unfavourable developmental characteristics (Robbins *et al.*, 2013). To explore this further, Robbins *et al.* (2013) crossed transgenic maize plants expressing P1 with maize plants that had either an expressing modifier (UFO1 E) or a mutated non-expressing modifier (*ufo1 NE*) and compared the two lines by proteomic analysis. Expression of lignin biosynthetic genes PAL, C4H and CAD was increased whereas expression levels of COMT and CCoAOMT was decreased in pericarp tissue of P1/UFO1 E plants when compared with P1/*ufo*1 *NE* plants. In internode tissue P1/UFO1 E plants showed a reduction in COMT protein levels but not in mRNA levels, indicating post-transcriptional silencing, which led to a decrease in lignin content when compared to P1/*ufo1 NE* and wild type controls (Robbins *et al.*, 2013). Histochemical analysis of the COMT down-regulated P1/UFO1 E plants suggested a reduction in lignin S monomers however this was not confirmed with quantitative methods.

Shen *et al.* (2012) identified a switchgrass (*Panicum virgatum*) MYB (*Pv*MYB4) with the closest homology to *At*MYB4. When co-expressed in Arabidopsis protoplasts, *Pv*MYB4 reduced expression of a reporter gene driven by the *At*PAL or *At*CCoAOMT promoter sequences (Shen *et al.*, 2012). When expressed in stably transformed tobacco, *Pv*MYB4 lines had stunted stature and reduced lignin content, with an increase in S:G ratio due to a greater reduction of G monomers over S monomers. Transgenic switchgrass expressing *Pv*MYB4 had reduced height and increased tillering with a reduction in lignin content but no change in S:G monomer ratio (Shen *et al.*, 2012; Shen *et al.*, 2013). qPCR analysis of the stably transformed PvMYB4 tobacco and switchgrass revealed that mRNA levels of F5H,

HCT and CCoAOMT were greatly reduced, with less extreme reduction in expression of the remaining seven lignin biosynthetic genes (Shen *et al.*, 2012).

The regulatory roles of some MYB genes have been assumed based on alternative data. An *At*MYB68 loss-of-function mutant Arabidopsis was found to have increased lignin content in the roots suggesting that *At*MYB68 would normally act to down-regulate lignin (Feng *et al.*, 2004). However, this was not confirmed by a gain-of-function experiment. Similarly, the down-regulation of *Ptt*MYB21a in hybrid aspen resulted in increased levels of CCoAOMT, suggesting that *Ptt*MYB21a may be a repressor of CCoAOMT expression (Karpinska *et al.*, 2004). Although it can be proposed that the down-regulation of CCoAOMT by increased *Ptt*MYB21a expression may lead to repressed lignin biosynthesis, it remains unconfirmed.

2.2.3 *Zm*MYB31 and *Zm*MYB42 expression down-regulates lignin biosynthesis

Alignment of animal MYB sequences and the first cloned plant MYB gene (Paz-Ares *et al.*, 1987) highlighted conserved regions to which primers were designed and six individual MYB genes were cloned from a developing flower bud cDNA library of *Antirrhinum majus* (snapdragon) (Jackson *et al.*, 1991). An alignment of these MYB genes highlighted two *Am*MYB genes with conserved R2R3 domains but different C-terminal sequences (Tamagnone *et al.*, 1998). It was hypothesised that these two MYB genes (*Am*MYB308 and *Am*MYB330) would bind the same DNA sequence but have different regulatory effects (Tamagnone *et al.*, 1998). Using degenerate primer sequences designed from *Am*MYB308 and *Am*MYB330, Fornalé *et al.* (2006) identified and cloned five different maize R2R3 MYB genes and expressed them in Arabidopsis. Plants expressing *Zm*MYB31 and *Zm*MYB42, which were found to be closely related to *Am*MYB308 and *Am*MYB330, showed an 8-fold and 2.1-fold suppression in lignin biosynthesis respectively when corrected for fresh tissue weight (Fornalé *et al.*, 2006). Both *Zm*MYB31 and *Zm*MYB42 exerted down-regulatory control over 4CL, F5H and COMT expression (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009). Further research found that *Zm*MYB31 also down-regulated C3H expression but slightly increased CAD expression, whereas *Zm*MYB42 down-regulated CAD expression, as well as PAL, C4H and HCT expression levels (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009). A summary of the down-regulatory effects of the *Am*MYB (Tamagnone *et al.*, 1998) and *Zm*MYB genes (Fornalé *et al.*, 2006) is shown in Table 2.3.

TABLE 2.3: Summary of the regulatory effects[1] over lignin biosynthesis gene expression levels after expression of Antirrhinum MYB genes in tobacco (Tamagnone *et al.*, 1998) and *Zea mays* MYB genes in Arabidopsis (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009).

	*Am*MYB308	*Am*MYB330	*Zm*MYB31	*Zm*MYB42
PAL	-			↓
C4H	↓		-	↓
4CL	↓	↓	↓	↓
HCT				↓
C3H			↓	-
CCoAOMT			-	-
CCR				-
F5H			↓	↓
COMT			↓	↓
CAD	↓		~↑	↓

[1]Arrows represent influence on gene expression levels. A dash represents no change in expression levels. A blank cell indicates transcription levels for that gene were not reported.

Arabidopsis plants harbouring *Zm*MYB31 were found to have up to a 70% reduction in lignin content (Fornalé *et al.*, 2010). The S:G monomer ratio did not change, but H monomers increased four-fold. These changes did not alter the carbohydrate composition within the cell wall (Fornalé *et al.*, 2010). These results reflect the general suppression activity of *Zm*MYB31 over the lignin biosynthetic pathway promoters (Fornalé *et al.*, 2010). Similar to these results, Arabidopsis expressing *Zm*MYB42 was found to have a 60% reduction in lignin content (Sonbol *et al.*, 2009). *Zm*MYB42 Arabidopsis plants had a reduced S:G ratio, with a 50% reduction of S units, a 70% increase in G units and an increase in H monomers. Further analysis found the cellulose content was not altered in the *Zm*MYB42 plants but there was an increase in other cell wall carbohydrates (hemicelluloses) (Sonbol *et al.*, 2009).

The synchronous down-regulation of several lignin biosynthetic pathway enzymes by MYB overexpression is an attractive solution to avoid detrimental metabolic pooling effects that may result when an individual gene is targeted (Besseau *et al.*, 2007; Weng *et al.*, 2008) as the altered expression of lignin biosynthesis genes will in turn alter the flow of metabolites through the lignin biosynthesis pathway (Baxter and Stewart, 2013). Although negative phenotypes including slow and stunted growth have been observed in lignin reduced plants over-expressing R2R3 MYB transcription factors, (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009; Tamagnone *et al.*, 1998), this is not uncommon when lignin content is significantly reduced (Vanholme *et al.*, 2010). While the 60% - 70% lignin reduction in these studies is severe, a more modest reduction in lignin may still improve cell wall digestibility while allowing plants to maintain normal development and growth (Vanholme *et al.*, 2010). Modest lignin reductions are achievable through the optimisation of constructs for transgenic expression, through promoter choice and the addition of targeting sequences (Vanholme *et al.*, 2008). Such optimisations can better control the levels and tissue specificity of expression which may help to establish a spectrum of tolerable lignin reductions in plants, thus allowing for normal or increased biomass production, with the added benefit of enhanced digestibility.

Zea mays MYB31 and MYB42 have been specifically demonstrated to reduce total lignin content by down-regulating the lignin biosynthetic pathway in Arabidopsis (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009). In addition, Fornalé *et al.* (2006) found that MYB31 and MYB42 expression down-regulated the maize COMT promoter (Capellades *et al.*, 1996). Arabidopsis expressing either MYB31 or MYB42 were crossed with an Arabidopsis line expressing green fluorescent protein (GFP) (Chiu *et al.*, 1996) under the control of the maize COMT promoter. The resulting F1 progeny had significant decreases in GFP expression which was attributed to the down-regulation of the COMT promoter by the maize MYB genes (Fornalé *et al.*, 2006). Research suggests that the lignin biosynthetic pathway is conserved to an extent between species (Fornalé *et al.*, 2006; Marita *et al.*, 2003) and as sugarcane and maize are genetically closely related (Selman-Housein *et al.*, 1999), the success in modifying expression of the maize COMT gene promoter (Fornalé *et al.*, 2006) supports the expectation that lignin can be successfully down-regulated in sugarcane through the over-expression of the maize transcription factors MYB31 and MYB42.

2.3 RNAi targeting of individual lignin biosynthesis genes

2.3.1 RNA silencing

Small non-coding RNA sequences are an integral part of RNA silencing pathways (Frizzi and Huang, 2010). Within plants these pathways double as a method of gene regulation (Vaucheret, 2006) and as an 'immune system' by targeting double stranded RNA (dsRNA) which can be indicative of viral infection or transposon activity (Margis *et al.*, 2006; Morel *et al.*, 2002). Silencing can occur as transcriptional gene silencing (TGS) by deactivation of genomic DNA through methylation (Mette *et al.*, 2000) or as post transcriptional gene silencing (PTGS) by mRNA degradation (Hamilton and Baulcombe, 1999). PTGS, commonly referred to as RNA interference (RNAi) (Vaucheret, 2006), can be utilised as a powerful molecular tool for gene silencing.

Double stranded RNA can originate from converted single stranded RNA (ssRNA) by an RNA dependent RNA polymerase (Dalmay *et al.*, 2000), or by inverted repeats forming a dsRNA hairpin conformation (Wesley *et al.*, 2003). Both endogenous and exogenous dsRNAs are indiscriminately recognised and cleaved by Dicer, an RNase III nuclease (Bernstein *et al.*, 2001) which produces short interfering RNA (siRNA) fragments of approximately 21 – 25 nucleotides in length (Hamilton *et al.*, 2002). The siRNAs are specifically methylated by the RNA methyltransferase HEN1 (Boutet *et al.*, 2003; Saito *et al.*, 2007) which serves to stabilise the siRNA (Li *et al.*, 2005; Yang *et al.*, 2006). RNA helicase (Filipowicz, 2005) unwinds the siRNA and one strand is selectively incorporated into an RNA induced silencing complex (RISC) (Frizzi and Huang, 2010). AGO1, an argonaute protein is an integral part of plant RISCs (Fagard *et al.*, 2000; Morel *et al.*, 2002) allowing it to bind specifically to ssRNA complementary to its incorporated siRNA fragment (Hammond *et al.*, 2000). Once the RISC is bound, the ssRNA can be cleaved through the activity of Slicer, a ribonuclease (Baumberger and Baulcombe, 2005; Tolia and Joshua-Tor, 2007). This silencing signal can move systematically throughout the plant (Himber *et al.*, 2003; Mlotshwa *et al.*, 2002), allowing for wide-spread suppression of specific gene sequences.

2.3.2 RNA silencing for plant biotechnology

By utilising the RNA silencing pathway in plants, constructs can be designed to produce dsRNA complementary to specific genes, thereby directing the plant to silence specific

genes of interest. Although sense and antisense constructs have been found in the past to result in degrees of PTGS (reviewed in Frizzi and Huang, 2010), RNAi constructs designed with both sense and complementary antisense gene sequences separated by a spacer region (intron) can dramatically increase the levels of gene silencing in transgenic plants (Wesley *et al.*, 2004; Wesley *et al.*, 2003). These constructs are commonly referred to as hairpin RNA (hpRNA) due to the conformational shape of the corresponding mRNA transcript (Wesley *et al.*, 2003). Although a large body of research exists for sense, antisense and RNAi targeted gene down-regulation of the lignin biosynthetic pathway in a variety of plant species (reviewed in Hisano *et al.*, 2009; Li *et al.*, 2008; and Simmons *et al.*, 2010) this literature review will specifically focus on the down-regulation of three lignin biosynthetic genes: CCoAOMT, F5H and COMT. Osabe *et al.* (2009) first provided evidence for the effectiveness of RNAi silencing in sugarcane by targeting the phytoene desaturase (PDS) gene within the β-carotene biosynthesis pathway. Regenerated sugarcane expressing the RNAi construct showed near complete loss of PDS mRNA expression (Osabe *et al.*, 2009). Two further studies (Jung *et al.*, 2012; Jung *et al.*, 2013) support the effectiveness of RNAi mediated gene silencing in sugarcane and provide evidence lignin biosynthesis genes can be down-regulated using an RNAi approach. This further highlights RNAi as an attractive tool for the manipulation of lignin biosynthetic genes with the intent of reducing lignin content, altering the lignin monomer ratio and improving overall bagasse digestibility.

2.3.3 Targeting CCoAOMT, F5H and COMT for RNAi down-regulation

2.3.3.1 Transgenic down-regulation of CCoAOMT

Caffeoyl CoA 3-O-methyltransferase (CCoAOMT) is responsible for the 3' methylation of caffeoyl-CoA to produce feruloyl-CoA (Hisano *et al.*, 2009; Raes *et al.*, 2003), a key step in the production of G and S lignin monomers (Figure 1.1; Barrière *et al.*, 2004). Only three studies in which CCoAOMT has been down-regulated using an RNAi mediated approach have been reported (Chen *et al.*, 2006; Li *et al.*, 2013; Wagner *et al.*, 2011). Analysis of CCoAOMT RNAi down-regulated alfalfa plants by Chen *et al.* (2006) and Chen and Dixon (2007) found a reduction in total lignin (approximately by 20%) in which G monomers were reduced whereas S monomers remained at wild type levels, thus increasing the S:G ratio. RNAi down-regulation of CCoAOMT in *Pinus radiata* reduced the lignin contents of transgenics by 5-20% of that of the controls with an increase in the H:G ratio due to decreased G-monomers (Wagner *et al.*, 2011). No S:G ratio was available as coniferous

gymnosperms, which include *Pinus radiata*, lack S-type lignin monomers (Wagner *et al.*, 2011). Similar results were found after RNAi targeting of CCoAOMT expression in transgenic maize (Li *et al.*, 2013). Transgenics had an average of 22% reduction in Klason lignin with a 57% increase in the S:G ratio when compared to wild type controls (Li *et al.*, 2013). Further analysis found the transgenic maize had an average increase of 23% in cellulose content but no differences in hemicellulose content when compared with controls (Li *et al.*, 2013).

Sense and antisense suppression of CCoAOMT in alfalfa (Guo *et al.*, 2001a; Marita *et al.*, 2003; Nakashima *et al.*, 2008) support these findings, as do results in other plant species engineered for down-regulated CCoAOMT expression. Increased S:G ratios with a 12 – 40% reduction in lignin content has been reported in flax (Day *et al.*, 2009), poplar (Meyermans *et al.*, 2000; Zhong *et al.*, 2000) and tobacco (Pinçon *et al.*, 2001; Zhao *et al.*, 2002). Further analysis found improved rumen digestibility and increased cellulose content in CCoAOMT down-regulated alfalfa (Guo *et al.*, 2001a; Guo *et al.*, 2001b; Marita *et al.*, 2003). Rumen digestibility is another indication of cell wall recalcitrance as higher lignin content reduces the digestibility and therefore the forage value of crops (Guo *et al.*, 2001b). Mutant Arabidopsis deficient in CCoAOMT expression also provide support for the transgenic studies (Van Acker *et al.*, 2013). A 21-25% reduction in lignin content was found in *ccoaomt* mutant lines with decreases in G monomers and increases in S monomers leading to an overall increase to the S:G ratio (Van Acker *et al.*, 2013). Mutant *ccoaomt* lines did not show any differences in crystalline cellulose content but had increases in rhamnose and galactose polysaccharides in the hemicellulose matrix (Van Acker *et al.*, 2013). In general, no phenotypic differences were observed between transgenic plants and controls (Day *et al.*, 2009; Meyermans *et al.*, 2000; Van Acker *et al.*, 2013; Zhao *et al.*, 2002), with the exception of one report of reduced plant height (Pinçon *et al.*, 2001).

Targeting CCoAOMT for down-regulation may also have a secondary effect on the cell wall structure itself. CCoAOMT, along with C3H are hypothesised to be important control points for cell wall lignification by acting as part of the ferulate production pathway (Barrière *et al.*, 2004). The feruloyl residues aid in cross-linking within the cell wall and may increase the resistance of the cell wall to hydrolysis by adding to its structural stability (Barrière *et al.*, 2004; Grabber, 2005). Therefore, a plant with reduced or impaired ferulate content or cross-linking may be more susceptible to enzymatic hydrolysis due to a lessening in the cell wall stability (Buanafina, 2009; Grabber, 2005; Li *et al.*, 2008).

2.3.3.2 Transgenic down-regulation of F5H

Ferulate 5-hydroxylase (F5H) is a cytochrome P450-dependent monooxygenase responsible for the 5' hydroxylation of coniferaldehyde and coniferyl alcohol (Hisano *et al.*, 2009; Raes *et al.*, 2003). As such it plays a key role in the production of S lignin as the hydroxylation of the 5' position is an essential step in the formation of the syringyl units (Figure 1.1; Weng and Chapple, 2010). There is only one report of F5H being down-regulated by RNAi, however it was included with COMT as half of a chimeric RNAi construct (Bhinu *et al.*, 2009). Regenerated canola plants were found to have up to a 26% reduction in lignin content with no apparent phenotypic differences, however it is not known if the altered lignin content is caused by reduced expression of F5H, COMT, or a combination of both. Expression of an antisense F5H construct in alfalfa (Chen and Dixon, 2007; Chen *et al.*, 2006; Nakashima *et al.*, 2008; Reddy *et al.*, 2005) resulted in reduced S monomers and, therefore a decreased S:G ratio, but did not reduce the overall lignin content.

In a comparison of two *f5h* mutant Arabidopsis lines, one had increased lignin content and the other had no changes to lignin content (Van Acker *et al.*, 2013). There was no change to the H monomer content of the lignin polymer however increases in G monomers (99.3% of lignin polymer was composed of G monomers in mutants) and a decrease in S monomers (none were detected in mutant lines) led to a decrease in the S:G ratio (Van Acker *et al.*, 2013). Crystalline cellulose was reduced in both mutant lines however no changes were detected in hemicellulose polysaccharides (Van Acker *et al.*, 2013). The *f5h* mutants did not show any phenotypic differences when compared with wild type controls (Van Acker *et al.*, 2013).

Stewart *et al.* (2009) analysed transgenic poplar over-expressing the F5H gene. The results revealed a drastic increase in S monomers (and therefore in S:G ratio) in transgenic plants when compared to controls (from 65% to over 90% of lignin was composed of S units). This provides additional support to previous findings of F5H expression positively correlating with S monomer concentration (whether increased or decreased) in the lignin polymer. This makes F5H an attractive gene target for suppression as the lignin content will potentially remain the same with an alteration in the lignin monomer ratios being the major impact of the genetic manipulation.

2.3.3.3 Transgenic down-regulation of COMT

Caffeic acid 3-O-methyltransferase (COMT) is responsible for the methylation of 5-hydroxy-coniferaldehyde and 5-hydroxy-coniferyl alcohol into sinapaldehyde and sinapyl alcohol (Hisano *et al.*, 2009; Raes *et al.*, 2003) and is therefore essential for the production of S lignin units (Figure 1.1; Bonawitz and Chapple, 2010; Weng and Chapple, 2010). RNAi suppression of COMT has decreased S units in canola (Bhinu *et al.*, 2009) and alfalfa (Chen and Dixon, 2007; Chen *et al.*, 2006) with a lignin reduction of up to 40% (Bhinu *et al.*, 2009). *Brassica napus* harbouring COMT-RNAi constructs were found to have residual COMT enzyme activities of 21% - 31% of controls which correlated with 15.67% - 35.09% reductions in lignin content of transgenic plants (Oraby and Ramadan, 2014). RNAi suppression of COMT expression has also been analysed in switchgrass with total lignin reductions of 6.4% - 14.7% and decreases to G and S monomers and to the overall S:G ratio (Fu *et al.*, 2011a). Further studies in COMT RNAi switchgrass have found Klason lignin reductions of 14% and a decreased S:G ratio (Samuel *et al.*, 2014). Fu *et al.* (2011a) found no changes to cellulose content or crystallinity and Samuel *et al.* (2014) found slightly higher xylose content but did not report if the changes were significant.

Two of the RNAi-COMT switchgrass lines previously generated (Fu *et al.*, 2011a) were field acclimatised and analysed over two growing seasons (2011-2012) (Baxter *et al.*, 2014). COMT transcript levels in leaf tissue of growing plants was found to be reduced in both RNAi lines by up to 97% over the 2011 and 2012 harvest (Baxter *et al.*, 2014). Samples were taken mid-season (green tissue) and during end-of-season harvesting (senesced tissue) during 2011 and 2012 and analysed for cell wall composition and saccharification potential. Total lignin content was reduced in green and senesced tissue in both RNAi lines by 8.4% – 14.5% over the two growing seasons with S:G reductions seen in all tissues analysed (Baxter *et al.*, 2014). No change in cellulose content was seen for either line over either growing season though both lines showed increased hemicellulose in 2012 season (Baxter *et al.*, 2014). The 2011 green tissue released 9% - 11.7% more sugar than controls and senesced tissue released 0.0% - 5.6% more sugar with no differences in cellulose crystallinity seen (Baxter *et al.*, 2014). In the 2012 harvest both lines had 16% - 18.6% decreases in cellulose crystallinity with 14.7% - 18.7% increase in sugar release from green tissue and 32% - 34.2% increase in release from senesced tissue (Baxter *et al.*, 2014).

COMT expression has also been RNAi targeted (using native COMT sequence) in transgenic sugarcane (Jung *et al.*, 2012; Jung *et al.*, 2013), providing the first insights into biotechnological mediated lignin down-regulation in sugarcane. In glasshouse grown COMT-RNAi sugarcane with approximately 65% - 95% reductions in COMT expression, decreases of 3.9% - 13.7% in total lignin were found when compared with controls (Jung *et al.*, 2012). These plants had a reduction to the S:G ratio due to a decrease in S monomers (Jung *et al.*, 2012). In field conditions, the COMT RNAi sugarcane continued to show greatly reduced COMT expression after six months in the field which resulted in reductions in lignin content by 5.5% - 12% (Jung *et al.*, 2013). While G monomer content remained unchanged in transgenics, S monomer content was reduced, reducing the S:G ratio (Jung *et al.*, 2013). Further analysis of two transgenic lines revealed no differences in glucose (indicative of cellulose) or arabinose, though one line showed an increase in xylose (Jung *et al.*, 2013).

Sense and antisense constructs have also been employed for COMT down-regulation. Overall the results showed lignin reductions of 10 – 31% with reductions in S units approaching near total loss in some reports in alfalfa (Guo *et al.*, 2001a; Guo *et al.*, 2001b; Marita *et al.*, 2003; Nakashima *et al.*, 2008), poplar (Jouanin *et al.*, 2000), tobacco (Pinçon *et al.*, 2001; Zhao *et al.*, 2002) and maize (He *et al.*, 2003; Piquemal *et al.*, 2002). Interestingly, a reduction of S-monomer units was also detected in aspen that had reduced levels of COMT (Tsai *et al.*, 1998), but there was no change in total lignin content. Reductions of S units reflect the positioning of COMT within the lignin biosynthesis pathway (Figure 1.1). COMT down-regulated plants were reported to have relative increases in cellulose (Marita *et al.*, 2003) and improved rumen digestibility (Guo *et al.*, 2001b; He *et al.*, 2003; Piquemal *et al.*, 2002).

Mutant *comt* Arabidopsis showed no differences in lignin content when compared against controls (Van Acker *et al.*, 2013). Further analysis found that the lignin polymer in the mutant plants had increased G monomers and reduced S monomers to the extent that the lignin polymer was composed of approximately 95% - 98% G monomers (Van Acker *et al.*, 2013). Both *comt* mutant lines showed no changes in matrix polysaccharide content, though both lines had decreases in crystalline cellulose with one line also showing a decrease in xylose and the other an increase in galactose (Van Acker *et al.*, 2013). No detrimental phenotypes were reported for any plants having reduced levels of COMT expression (Bhinu *et al.*, 2009; Fu *et al.*, 2011a; Jouanin *et al.*, 2000; Oraby and Ramadan, 2014; Pinçon *et al.*, 2001; Van Acker *et al.*, 2013; Zhao *et al.*, 2002) except in field-grown

sugarcane where lignin reductions of 8% - 12% resulted in reduced biomass (Jung *et al.*, 2012; Jung *et al.*, 2013). Sugarcane with lignin reductions of 6% or less did not show any phenotypic differences or loss of biomass (Jung *et al.*, 2012; Jung *et al.*, 2013). Interestingly, one RNAi-COMT switchgrass line analysed by Baxter *et al.* (2014) showed no phenotypic differences in the 2011 season but had increases in tiller height, diameter and total biomass yield in 2012 (Baxter *et al.*, 2014). The other RNAi-COMT line had decreases in tiller height, diameter and total biomass yield in 2011 but no differences were observed after the 2012 harvest (Baxter *et al.*, 2014).

2.3.4 Targeting of CCoAOMT, F5H and COMT

CCoAOMT, F5H and COMT have been chosen for down-regulation in sugarcane as their location within the lignin biosynthetic pathway suggests that their down-regulation may not only reduce overall lignin content, but may also alter the lignin monomer ratio, thereby altering the lignin polymer composition. This combination of altered and reduced lignin deposition can positively influence plant digestibility (Grabber, 2005; Sonbol *et al.*, 2009). A change in the lignin monomer composition can also reduce recalcitrance and improve digestibility (Chen and Dixon, 2007; Grabber, 2005; Huntley *et al.*, 2003; Simmons *et al.*, 2010), potentially through a change in the linkage pattern found throughout the lignin polymer (Jackson *et al.*, 2008). Production of more homogeneous lignin polymers would reduce the number of different bond types, thus simplifying the lignin polymer and in turn reducing the severity of required pretreatment (Weng *et al.*, 2008) without negatively affecting plant growth or development (Vanholme *et al.*, 2008).

The degree of lignin polymer cross-linking is determined by the S:G monomer ratio of the lignin polymers with an increase in G monomers resulting in a greater degree of cross-linking (Ferrer *et al.*, 2008). This has led to speculation that a reduction in S monomers (with a relative or actual increase in G monomers) will increase biomass recalcitrance (Kiyota *et al.*, 2012; Ziebell *et al.*, 2010). Although there is some support that a reduction in G monomers (increase in S:G ratio) improves saccharification (Chen and Dixon, 2007; Studer *et al.*, 2011; Xu *et al.*, 2011) there are also a number of reports that have found a reduction in S units (decrease in S:G ratio) correlates with saccharification improvements (Baxter *et al.*, 2014; Chen and Dixon, 2007; Fornalé *et al.*, 2012; Fu *et al.*, 2011a; Fu *et al.*, 2011b; Jung *et al.*, 2012; Jung *et al.*, 2013; Sonbol *et al.*, 2009). The results suggest that

saccharification efficiency may not be solely dependent on lignin content or polymer monomer ratio (Fornalé *et al.*, 2010; Fu *et al.*, 2011b; Studer *et al.*, 2011).

Targeting genes early in the lignin biosynthesis pathway can lead to undesirable phenotypes (Baucher *et al.*, 2003; Besseau *et al.*, 2007; Coleman *et al.*, 2008; Wagner *et al.*, 2007); targeting genes later in the pathway is expected to reduce the chance of the generated transgenic plants showing undesirable phenotypes (Chen and Dixon, 2007; Reddy *et al.*, 2005) though this is not always the case (Prashant *et al.*, 2011; Thévenin *et al.*, 2011; Wang *et al.*, 2012). As previously discussed, CCoAOMT down-regulation reduced the G monomers and F5H and COMT down-regulation reduced the S monomers within the lignin polymer, but no decreases in biomass were reported (Chen and Dixon, 2007; Reddy *et al.*, 2005). This further supports the selection of these genes for down-regulation in sugarcane as high crop productivity is an important requirement by the industry (Hisano *et al.*, 2009), so genes that reduce crop biomass, such as HCT and C3H (Chen and Dixon, 2007) may not be the most suitable candidates for down-regulation in sugarcane. With the discovery of a larger suite of promoters for specific spatial and temporal gene expression in sugarcane, additional genes in the lignin biosynthetic pathway may become more attractive for misregulation.

2.4 Saccharification improvements through altered lignin composition

2.4.1 Saccharification improvements in non-transgenic lignin modified plants

There are an increasing number of papers finding that a reduction in lignin content, whether through classical breeding or biotechnology, correlates strongly with an increase in saccharification of structural cell wall carbohydrates. Negative correlations between lignin content and saccharification have been found in wild type switchgrass (Shen *et al.*, 2009) and poplar (Studer *et al.*, 2011). Arabidopsis lines with mutations in eight lignin biosynthetic pathway genes were enzymatically hydrolysed before and after pretreatment (Van Acker *et al.*, 2013). Mutant plants for *c4h*, *4cl*, *ccoaomt* and *ccr* with reductions in lignin all showed increases in cellulose conversion to glucose (Van Acker *et al.*, 2013). Mutant *comt* and *f5h* plants did not have any differences in lignin content but released more glucose during enzymatic hydrolysis without pretreatment (Van Acker *et al.*, 2013). Interestingly, after pretreatment only the *comt* mutants released more glucose; the *f5h* mutants were comparable to wild types (Van Acker *et al.*, 2013). Mutant plants for *pal* and *cad* did not

show any differences in lignin content and did not yield more glucose than controls after enzymatic hydrolysis either before or after pretreatment.

2.4.2 Saccharification improvements after transgenic expression of MYB transcription factors

Only two research groups have shown improved yield of fermentable sugars after enzymatic hydrolysis in plants with reduced lignin content due to the expression of MYB transcription factors. *Pv*MYB4 expression in transgenic switchgrass (*Panicum virgatum*) reduced lignin content and improved saccharification three-fold in transgenic plants compared with controls (Shen *et al.*, 2012; Shen *et al.*, 2013). Reductions in Arabidopsis lignin content by expression of *Zm*MYB31 (Fornalé *et al.*, 2010) and *Zm*MYB42 (Sonbol *et al.*, 2009) improved enzymatic release of monosaccharides by 14% and 68% respectively when compared with controls (Fornalé *et al.*, 2010; Sonbol *et al.*, 2009).

2.4.3 Saccharification improvements after expression of sense/antisense/RNAi constructs

Increased glucose yield after enzymatic hydrolysis has been reported in both dicot and monocot plants with altered lignin contents resulting from specific down-regulation of lignin biosynthetic pathway gene expression levels. Antisense down-regulation of C4H, HCT, C3H, CCoAOMT, F5H and COMT in transgenic alfalfa led to all lines except F5H having reduced lignin (Chen and Dixon, 2007). In turn, all lines except F5H had improved enzymatic hydrolysis efficiency over controls after pretreatment (Chen and Dixon, 2007). Alfalfa with antisense down-regulation of CAD and CCR were found to have improved saccharification and rumen digestibility (Jackson *et al.*, 2008) and RNAi targeting of CAD improved enzymatic hydrolysis release of glucose in switchgrass (*Panicum virgatum* L.) (Fu *et al.*, 2011b; Saathoff *et al.*, 2011) and maize (Fornalé *et al.*, 2012).

Interestingly, although CAD RNAi switchgrass lines had significantly reduced lignin content (Fu *et al.*, 2011b; Saathoff *et al.*, 2011) CAD RNAi maize did not show reduced total lignin in stem tissue, though a decrease in S:G ratio and an increase in H monomers was observed (Fornalé *et al.*, 2012). This may be due to the enzyme activity of CAD in the transgenic maize being reduced by 66% (Fornalé *et al.*, 2012) whereas the transgenic switchgrass had CAD enzyme activity reductions of 61% - 88% (Fu *et al.*, 2011b) and by 70% - 95% in the

majority of plants analysed (Saathoff *et al.*, 2011). Sense and antisense down-regulation of COMT in tall fescue (*Festuca arundinacea* Schreb.) also led to improved rumen digestibility (Chen *et al.*, 2004). RNAi down-regulation of COMT (Baxter *et al.*, 2014; Fu *et al.*, 2011a) and 4CL (Xu *et al.*, 2011) in switchgrass reduced lignin content and improved glucose yield after enzymatic hydrolysis (Baxter *et al.*, 2014; Fu *et al.*, 2011a; Xu *et al.*, 2011). Antisense targeting of 4CL in poplar reduced lignin content however this did not correlate with an increase in sugar released during enzymatic hydrolysis (Voelker *et al.*, 2010).

2.4.4 Saccharification improvements in wild type and transgenic sugarcane

A negative correlation between lignin content and saccharification yield has also been observed in both wild type and transgenic sugarcane. 115 non-commercial varieties of sugarcane produced by precision or classical breeding were analysed for lignocellulosic compositions and saccharification efficiencies (Benjamin *et al.*, 2013). Sugarcane with a low lignin content of 14% - 16% released more glucose after enzymatic hydrolysis than the varieties with intermediate lignin contents of 18% - 20% however the intermediate plants had better agronomic productivity (biomass content) than the low lignin plants (Benjamin *et al.,* 2013). Six sugarcane varieties analysed in Benjamin *et al.* (2013) had cell wall composition and saccharification potentials examined over the 2009 and 2011 harvesting seasons (Benjamin *et al.*, 2014). Some negative correlations were seen between lignin content and glucose liberated after enzymatic hydrolysis in the 2009 harvested material but not in 2011 across a range of pretreatment severities (Benjamin *et al.*, 2014). The 2011 harvested material also saw small increases in lignin content for four of the six varieties which the authors suggest contributed to some slight overall decreases in glucose yields after hydrolysis (Benjamin *et al.*, 2014). Lower than average rainfall during a period of drought in 2010 was considered at least partially responsible for these changes in the 2011 harvested material (Benjamin *et al.*, 2014).

Masarin *et al.* (2011) selectively bred low-lignin content sugarcane varieties to assess their performance in the production of bioethanol. Each of the 11 new varieties had lignin contents of 17% - 21.5% which was significantly less than the controls (Masarin *et al.*, 2011). After 72 hours of enzymatic digestion the 11 new varieties and controls segregated into three groups. The two varieties with the lowest lignin contents of 16.8% and 18.6% had the highest rates of cellulose conversion to glucose, 31% and 25% respectively (Masarin *et al.*, 2011). Plants with intermediate lignin contents of 18.6% - 20.6%, had intermediate

cellulose conversion rates of 19% - 22% and the varieties and controls with higher lignin contents of 20.2% - 24.5% had the lowest levels of cellulose conversion of 12% - 14% (Masarin *et al.*, 2011). When the new varieties were assessed for agronomic performance the results were mixed with no clear relations between lignin content, biomass production and sucrose content (Masarin *et al.*, 2011). Interestingly the variety with the lowest lignin content had the highest biomass of all varieties and the fourth highest sucrose content (Masarin *et al.*, 2011). The variety with the second lowest lignin content had the second highest sucrose content but was the penultimate variety in biomass yield (Masarin *et al.*, 2011).

Transgenic sugarcane with significantly reduced lignin content through RNAi targeting of COMT have also been assessed for saccharification performance (Jung *et al.*, 2012; Jung *et al.*, 2013). The glasshouse grown transgenic sugarcane lines with lignin reductions of 3.9% - 13.7% released up to 29% more glucose without pretreatment and 34% more glucose after pretreatment than the control plants (Jung *et al.*, 2012). The COMT RNAi sugarcane (Jung *et al.*, 2012) were transferred and grown in field conditions for seven months (Jung *et al.*, 2013). Analysis found a 6% lignin reduction improved glucose release by enzymatic hydrolysis by 19% - 23% without phenotypic or agronomic penalties or differences (Jung *et al.*, 2013). This is in contrast to sugarcane plants with an 8% - 12% lignin reduction (Jung *et al.*, 2013). Though these plants had a 28% - 32% increase in saccharification efficiency over controls, the plants also had reduced biomass after harvesting (Jung *et al.*, 2013).

2.5 Summary and implications

Worldwide oil consumption and demand are increasing at an unsustainable rate (Murray and King, 2012). This trend has been strongly linked with climate change which has influenced increased research into alternative, environmentally friendly energy sources (Limayem and Ricke, 2012; Murray and King, 2012; Yang *et al.*, 2013). As well as limiting further environmental damage, oil production is also slowing and being a finite resource makes finding and developing alternative energy sources imperative and unavoidable (Murray and King, 2012). One part of the solution is to increase the energy sourced from renewable materials, such as second generation bioethanol from waste plant biomass.

The sugarcane industry produces large amounts of sugarcane bagasse (de Souza *et al.*, 2014; Lakshmanan *et al.*, 2005) which is currently underutilised and would be a suitable

lignocellulosic source of fermentable sugars for bioethanol production (Canilha *et al.*, 2012; Yuan *et al.*, 2008). However the recalcitrant nature of the lignin present in the bagasse increases the bioethanol production costs (Benjamin *et al.*, 2013; Chen and Dixon, 2007) to levels which would not be economically competitive with oil as an energy source (Benjamin *et al.*, 2013; Yang *et al.*, 2013; Yuan *et al.*, 2008). The production costs could be reduced if the lignin content of the sugarcane bagasse was reduced.

The complexity of the sugarcane genome, long timeframes for traditional breeding (Grivet and Arruda, 2001; Suprasanna *et al.*, 2011), and complex combinations of desired commercial traits, such as high sucrose yield and improved cell wall properties for biofuels, highlights sugarcane as a strong candidate for biotechnological attention (Canilha *et al.*, 2012; de Souza *et al.*, 2013; Suprasanna *et al.*, 2011) which is the focus of this thesis. If biotechnological approaches can be utilised to reduce the lignin content of sugarcane plants, and thus the bagasse, this would be a step towards the production of cost-competitive second generation bioethanol from this bagasse which would benefit the environment, the Australian economy and Australian sugarcane farmers (Canegrowers, 2010; Lakshmanan *et al.*, 2005), as well as having a wider global impact.

The research described in this thesis aims to increase the limited available knowledge of lignin development and modification in sugarcane through the completion of three aims. The first aim is the development of a profile of lignin biosynthesis and the deposition of secondary cell wall components in sugarcane by complementing lignin biosynthesis gene expression level data with cell wall compositional data. This thesis further aims to explore the increased release of glucose from sugarcane bagasse genetically modified to have a reduced or altered lignin composition. This will contribute to second generation bioethanol research using sugarcane bagasse as a source of fermentable monosaccharides. This in turn will contribute to the widening body of research aimed at using plant biomass to reduce the use and reliance on oil as a source of energy. A negative correlation between lignin content and saccharification efficiency has been shown in wild type sugarcane (Benjamin *et al.*, 2013; Masarin *et al.*, 2011) supporting the hypothesis that genetically modified lignin-reduced sugarcane will have increased levels of glucose released during enzymatic hydrolysis. Two different approaches will be taken to reduce lignin content in sugarcane.

Aim two utilises the transgenic expression of transcription factors to overall reduce the expression of multiple lignin biosynthetic genes. Maize transcription factors MYB31 and

MYB42 have been previously shown to reduce lignin content in Arabidopsis by down-regulating the expression of genes within the lignin biosynthesis pathway (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009) which positively correlated with an increase in saccharification after enzymatic hydrolysis (Fornalé *et al.*, 2010; Sonbol *et al.*, 2009). Although there are no available reports of MYB31 or MYB42 transcription factors being expressed in a monocot species they are expected to function with similar results in sugarcane as the MYB genes will be isolated from maize which is closely related to sugarcane (Selman-Housein *et al.*, 1999). The expression of these transcription factors in sugarcane and the observed effects on lignin biosynthesis and enzymatic hydrolysis performance comprises the second aim of this thesis.

Aim three involves RNAi mediated down-regulation of expression levels of specifically targeted lignin biosynthesis genes, CCoAOMT, F5H and COMT, to reduce lignin content and potentially alter the lignin monomer ratio. As discussed, these genes have been down-regulated in both dicots and monocots though mutations or sense/antisense/RNAi constructs (Section 2.3.3). The findings that a down-regulation in CCoAOMT (Chen and Dixon, 2007; Van Acker *et al.*, 2013), F5H (Van Acker *et al.*, 2013) and COMT (Fu *et al.*, 2011a; Jung *et al.*, 2012; Jung *et al.*, 2013) expression led to improved saccharification efficiency supports the hypotheses that the down-regulation of these genes in sugarcane will yield improved glucose release by enzymatic hydrolysis. That this has been shown in COMT RNAi down-regulated sugarcane (Jung *et al.*, 2012; Jung *et al.*, 2013) offers further support that this result can be replicated in an important Australian sugarcane cultivar.

The outcomes of this research will contribute to the development and commercialisation of cost-competitive second generation bioethanol from sugarcane bagasse. This will reduce the reliance on oil as a source of energy and in turn will have positive benefits for the global environment and economy. Sugarcane is an important crop in Australia and the utilisation of bagasse will benefit Australian sugarcane farmers as it will add value to what is currently considered a waste by-product of sugar production.

Chapter 3

General Methods

3.1 General Statements

3.1.1 Chemicals

All chemicals used through the methods section were purchased from Sigma (MO, USA), Merck (NJ, USA) or Univar (WA, USA) unless otherwise noted.

3.1.2 Water

All water used for buffers and washes was ultra-pure Millipore Milli-Q Plus (MA, USA) filtered water. All water used for polymerase chain reaction (PCR) and quantitative real-time polymerase chain reaction (qPCR) reactions, pellet resuspensions and all general molecular biology procedures was sterile RNase and DNase free water (Promega, WI, USA).

3.1.3 Buffers

All buffers were made to volume with ultra-pure Millipore Milli-Q Plus (MA, USA) filtered water.

TAE buffer: 40mM Tris-acetate, 1mM EDTA
TPS buffer: 100mM Tris, 1M KCl, 10mM Na_2EDTA
Citrate buffer (50mM): 9.61g/L citric acid, pH 4.8
Sodium acetate buffer (50mM) with 0.02% (w/v) sodium azide: 4.1g/L sodium acetate, pH 5.0 with 10mL of 2% (w/v) sodium azide solution/L added

3.1.4 Media

All media used in the methods section are listed in Appendix A.

3.1.5 Design and synthesis of oligodeoxyribonucleic acids

All oligodeoxyribonucleic acids (primers) used for PCR and qPCR were designed using Primer3 (Rozen and Skaletsky, 2000) and by manual identification. All primers were synthesised by Geneworks (SA, AUS) and diluted as appropriate prior to use in PCR or qPCR reactions.

3.1.6 Polymerase chain reaction cyclers

All endpoint PCR was performed using a BioRad (CA, USA) DNA Engine Peltier Thermal Cycler. All qPCR 100 well gene discs (Qiagen, Limburg, NLD) were run and analysed on a Qiagen Rotor-Gene Q (Limburg, NLD). Program information for the different polymerases used in this research can be found in Sections 3.5.5 and 3.6.2.

3.1.7 Electrophoresis of nucleic acids

Endpoint PCR products and digested plasmids were electrophoresed through a 1% (product ≥200bp) or 2% (product ≤200bp) agarose (Bioline, London, UK) TAE gel with the addition of SYBR safe DNA gel stain (Life Technologies, CA, USA) at 80V and 120V, respectively, and visualised using a G-Box (Syngene, Cambridge, UK). For all agarose gels Molecular Weight Marker X (Roche, Basel, CHE) was used a guide for determining the size of electrophoresed bands. An equal mix by volume of yellow food colouring (Queen, QLD, AUS), glycerol and sucrose (solid) was used as a loading dye for all samples.

3.1.8 Computer software

Vector NTI (Life Technologies, CA, USA) was utilised for the creation and management of all cloning sequences, and for viewing sequence chromatographs after BigDye Terminator reactions (Life Technologies, CA, USA) and SnapGene Viewer v2.8 (GSL Biotech LLC, IL, USA) was used for creating plasmid maps (Appendix E and Appendix J). Sequences obtained after BLAST searching (www.ncbi.nlm.nih.gov) were aligned using Multiple Sequence Alignment by CLUSTALW (Kyoto University Bioinformatics Centre; www.genome.jp/tools/clustalw). DNA electrophoresed through agarose gel was visualised within a G-Box fluorescent gel imaging system (Syngene, Cambridge, UK) and images taken with Genesnap software (Syngene, Cambridge, UK). The 100 well gene discs (Qiagen, Limburg, NLD) for qPCR were

pipetted out by a CAS1200 precision liquid handling system (Qiagen, Limburg, NLD) using Robotics 4 software. The results from qPCR were analysed using the associated Qiagen (Limburg, NLD) Rotor-Gene Q software and Microsoft Excel 2010 (WA, USA). All HPLC runs were analysed using Waters (MA, USA) Empower Pro software. Statistical analysis of experimental results was performed using GraphPad Prism v6.04 (GraphPad Software, CA, USA) and Minitab 16 (Minitab, Inc., PA, USA). This thesis was prepared using Microsoft Word 2010 (WA, USA).

3.1.9 Statistical analysis of results

Statistical analysis was carried out using either a two-tailed t-test assuming unequal variance, $p = 0.05$, comparing transgenic plants to transgenic controls or a one-way ANOVA with Tukey post-hoc analysis, $p = 0.05$, as appropriate. All data for one-way ANOVA analysis satisfied the Brown-Forsythe test for equal variance (GraphPad Software, CA, USA).

3.2 Preparation and transformation of competent *Escherichia coli*

3.2.1 Preparation of heat shock competent *E.coli*

Heat shock competent *E.coli* (XL1 Blue *Escherichia coli*, Agilent Technologies, CA, USA) stocks were generated using the methods published by Inoue *et al.* (1990) with minor modifications. In brief, an aliquot of previously prepared competent XL1 Blue *E.coli* were grown overnight in LB with tetracycline (52.875µM) at 37°C with 200rpm. The following morning a 100µL aliquot was subcultured into 4mL LB with tetracycline (53µM) and grown at 37°C with 200rpm. After six hours a 500µL aliquot was added to 250mL SOB and grown for two nights at 18°C with 200rpm. The cultures were pelleted by centrifugation and resuspended in TB with added DMSO (70µL/mL) before 100µL aliquots were snap frozen in liquid nitrogen and stored at -80°C.

3.2.2 Heat shock transformation of competent XL1 Blue *E.coli*

The prepared heat shock competent *E.coli* were transformed using the method published by Inoue *et al.* (1990) with minor modifications. In brief, 5µL of ligation mixture mixed with 50µL of competent *E.coli* was rested on ice for 20 minutes before incubation (42°C for 45 seconds) to initiate heat shock uptake of the ligated plasmids. After two minutes resting on

ice, 950µL of LB was added to the mixture and incubated at 37°C for 1.5 hours with shaking (200rpm). A 100µL aliquot of the mixture was then spread onto an LB agar plate with appropriate antibiotic.

3.2.3 Preparation of LB agar plates for blue/white selection

LB agar plates prepared for blue/white selection (Vieira and Messing, 1982) contained 2.4mg isopropyl-β-D-thiogalactopyranoside (IPTG), 1mg 5-bromo-4-chloro-3-indolyl-β-D-galactoside (x-gal), and the appropriate selection antibiotic. Antibiotic was added to media after autoclaving but before pouring into petri dishes. After LB/antibiotic plates had set, they were wrapped and stored at 4°C until use. IPTG and x-gal were spread onto individual LB/antibiotic petri dishes and allowed to be absorbed at room temperature for approximately 1.5 hours before the LB/antibiotic/IPTG/x-gal plates were inoculated with heat shock transformed *E.coli* and incubated at 37°C overnight. White *E.coli* colonies were screened for the plasmid with ligated insert whereas blue colonies were assumed to contain empty plasmids.

3.2.4 Subculturing of transformed *E.coli* colonies for plasmid extraction

After overnight incubation, white *E.coli* colonies were inoculated into 5ml liquid LB containing appropriate antibiotic selection and incubated overnight at 37°C with shaking (200rpm). The subcultured colonies then underwent plasmid extractions. Cultures for long term storage were mixed 1:1 with 80% glycerol and stored at -80°C.

3.3 Recovery, digestion and sequencing of recombinant plasmids from transformed competent *E.coli*

3.3.1 Extraction of plasmids from transformed *E.coli*

Plasmids for routine screening were purified from transformed *E.coli* colonies via alkaline lysis (Bimboim and Doly, 1979) following the modifications to the solutions and protocols seen in Sambrook and Russell (2001) and Thomas *et al.* (1988). Briefly, two milliliters of inoculated LB (Section 3.2.4) was centrifuged (14,800rpm, 1 minute, room temperature), and the resulting pellet was resuspended in 100µL Solution 1. Solution 2 (200µL) was added

and mixed gently to lyse cells before the addition of 150µL of Solution 3. Tubes were centrifuged (14,800rpm, 3 minutes, room temperature) to pellet gDNA and cellular debris. The supernatant was transferred to a fresh tube and the addition of 100% ethanol precipitated the pDNA before centrifugation (14,800rpm, 1 minute, room temperature). The resulting pDNA pellet was washed with 70% ethanol before air drying (room temperature) for five minutes. Purified pDNA pellets were resuspended in 30µL DNase free water containing 0.6µg RNase (Roche, Basel, CHE). Plasmids for sequencing, cloning or microprojectile bombardment were purified with the Wizard Plus SV Miniprep DNA purification system (Promega, WI, USA) following manufacturer's instructions.

3.3.2 Restriction digests

All restriction digests were performed at the recommended temperature and time according to manufacturer's instructions. All restriction enzymes used were supplied by either Roche (Basel, CHE) or New England Biolabs (MA, USA).

3.3.3 BigDye Terminator sequencing of DNA

All sequencing was conducted using BigDye Terminator (Life Technologies, CA, USA) PCR reactions and prepared for sequencing using the sodium acetate/EDTA/ethanol clean-up methods using manufacturer's protocol with modifications noted in Griffith University DNA sequencing facility (GUDSF) clean-up protocol (GUDSF, 2005). Briefly, 2µL of purified plasmid (Wizard Plus SV Miniprep DNA purification system, Promega, WI, USA) was added to a solution containing 1µL BigDye, 4µL sequencing buffer, 1µL primer (final concentration of 3.2pM) and 12µL water. Upon completion of the sequencing PCR (96°C x 1 minute, then 30 cycles of: 96°C x 10 seconds, 50°C x 5 seconds, 60°C x 4 minutes) the samples underwent a clean-up process. This involved mixing the PCR reaction, 2ul of 3M sodium acetate pH 5.2 and 2ul of 125mM EDTA pH 8.0 in a 1.5mL microfuge tube before the addition of 50µL 100% ethanol to precipitate nucleic acids. After 15 minutes incubation, tubes were centrifuged (14,800rpm, 20 minutes, room temperature) and the supernatant was removed. The remaining pellet was washed in 70% ethanol and centrifuged (14,800rpm, 5 minutes, room temperature) before the supernatant was removed and tubes air dried for one hour. All sequencing was performed at either GUDSF (Griffith University, QLD) or CARF (QUT, QLD).

3.4 Generation and transformation of sugarcane callus

3.4.1 Generation of sugarcane callus for transformation

Sugarcane callus (KQ228 cultivar) was generated from transverse leaf whorl sections from field grown plants (BSES Limited, Meringa, QLD, AUS) starting at 1cm above the meristem. Whorls were incubated in the dark on MSC_3 media (Bower *et al.*, 1996) at 28°C for 6 weeks. All calli was subcultured onto fresh media every 10-12 days throughout the initiation, transformation, selection and regeneration periods.

3.4.2 Preparation of gold microcarrier particles for microprojectile bombardment

Preparation of gold microcarrier particles was performed using in-house methods. Briefly, 250mg of 1µm gold microcarrier particles (BioRad, CA, USA) were vortex mixed for three minutes continuously in 1mL 100% ethanol before centrifugation (3000rpm; 5 minutes; room temperature). The supernatant was discarded and the process twice repeated. The process was again repeated in triplicate, substituting sterile Millipore water for ethanol. The pellet was resuspended in 2mL sterile 50% glycerol and 25µL aliquots were taken with regular vortex mixing to ensure no settling of the stock gold particles. The aliquots were stored at -20°C.

On the day of bombardment, 1µL plasmid (1µg/µL) containing the transgene, 1µL plasmid (1µg/µL) containing the selection gene, 5µL spermidine (0.1M) and 25µL $CaCl_2$ (1M) were added to an aliquot of gold particles and mixed by vortex 20 seconds per minute for five minutes. The reaction was allowed to rest for 10 minutes on ice before 22µL of supernatant was removed. The samples were mixed by vortex and 5µL aliquots were used per bombardment.

3.4.3 Microprojectile based transformation of sugarcane callus

Prior to transformation, callus generated from leaf whorls was subcultured onto osmotic MSC_3 media (Bower *et al.*, 1996) and placed in the dark at 28°C for four hours. Transformations were performed via particle inflow gun microprojectile bombardment (Bower and Birch, 1992) using 1500kPa helium pressure in -90kPa vacuum at a firing distance of eight centimetres. Each plate of callus was covered with a stainless steel mesh

baffle and bombarded twice, rotating the plate 90° between shots. The co-bombarded selection gene plasmid contained a *Zm*Ubi-nptII-nos cassette which codes for the neomycin phosphotransferase II (nptII) gene and confers geneticin resistance to callus for selection. Untransformed control callus did not undergo any microprojectile bombardment. After bombardment, callus was rested on osmotic media for four hours before subculturing onto fresh MSC_3 media and maintained in the dark at 28°C.

3.4.4 Regeneration of transgenic sugarcane plants

Four days after microprojectile bombardment transformation, the callus was transferred to MSC_3 media containing 50mg/L G418 disulfate salt solution (a form of geneticin; Roche, Basel, CHE) to select for transformed cells and maintained for four weeks in the dark at 28°C. After this period, callus was transferred to MS media containing 50mg/L G418 disulfate salt solution and maintained at 25°C with a 16 hour photoperiod for shoot and root development. The untransformed control callus was regenerated into wild-type sugarcane for control purposes on media without selection antibiotics. After 6 - 8 weeks one plant per calli clump (with visible roots) was used for acclimatisation to ensure only individual transformation events were analysed.

3.4.5 Acclimatisation and harvest of transgenic sugarcane plants

Tissue culture plantlets selected for acclimatisation were transferred from media to potting mix (Searles Real Premium, QLD, AUS) and placed in growth rooms under a 16 hour photoperiod at 25°C with watering every second day. Initially plants were covered, with periodic uncovering to minimise acclimatisation stress. The cover was permanently removed on the fifth day. When the plants reached approximately 30cm in height they were transferred to the QUT Carseldine Glasshouse Facility and grown in plastic pots (4.5L, 19cm height, 19cm diameter) with potting mix (Searles Real Premium, QLD, AUS) and Osmocote (Scotts, NSW, AUS) fertiliser pellets at 27°C ± 3°C under natural light. Plants were watered to saturation twice per week and fertilised with Aquasol (Yates, NSW, AUS) once per month with regular removal of tillers. Potted plants underwent periodic randomised position rotation within the glasshouse to minimise positional effects.

When the glasshouse sugarcane plants were harvested, tissue for qPCR analysis was immediately stored on dry ice and at -80°C in the laboratory. Tissue for cell wall

compositional analysis and enzymatic hydrolysis was stored on ice after harvesting, dried in a convection drying oven at 40°C for seven days and stored at room temperature. Tissue for juice analysis was stored on dry ice after harvest and at -20°C in the laboratory.

3.5 Nucleic acid extraction and use

3.5.1 DNA extractions from plant samples

All DNA extractions were performed using the Rapid Release method (Thomson and Henry, 1995). Briefly, samples were ground in 100uL of TPS buffer and incubated for 10 minutes at 95°C before being washed with chloroform: isoamyl alcohol (24:1) (Sambrook and Russell, 2001) and centrifuged (14,800rpm, 5 minutes, room temperature). A 1:4 dilution of the supernatant was used for subsequent PCR reactions.

3.5.2 RNA extraction from plant samples

RNA was extracted from all tissue samples using Tri Reagent (Sigma-Aldrich, MO, USA) following the manufacturers' protocol using tissue ground under liquid nitrogen. Briefly, tissue was incubated at room temperature with Tri Reagent for five minutes before the addition of 200µL chloroform. Tubes were vigorously mixed and incubated at room temperature for a further five minutes. Reactions were centrifuged (12,000*g*; 15 minutes; 4°C) and the supernatant collected and mixed with an equal volume of isopropanol. After 10 minutes incubation at room temperature reactions were again centrifuged (12,000*g*; 10 minutes; 4°C) and the RNA pellet subsequently washed in 75% ethanol before a final centrifugation (7500*g*; 5 minutes; 4°C). The supernatant was removed and the pellet dried for 10 minutes before being resuspended in 30µL water and incubated at 60°C for 10 minutes. The extractions were kept on ice and used for cDNA synthesis immediately or stored at -80°C. RNA concentrations were quantified with a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, MA, USA).

3.5.3 DNase digestion of RNA extractions

Extracted RNA (1µg) was digested with RQ1 RNase-free DNase (Promega, WI, USA) following manufacturer's methods with the 37°C incubation being increased to one hour.

After digestion the samples were either used immediately for cDNA synthesis or stored at -80°C.

3.5.4 First strand cDNA synthesis from DNase digested RNA

DNase-treated RNA was used as a template for first strand cDNA synthesis in the M-MLV Reverse Transcription system (Promega, WI, USA) following manufacturer's instructions. Briefly, 0.5µg of DNase-treated RNA was primed for reverse transcription by the addition of 1µL of 50µM oligo-dT primer (5' TTTTTTTTTTTTTTTTTTVN 3') and 10.75µL water followed by incubation at 70°C for five minutes. After quenching the reactions on ice, 1µL M-MLV reverse transcriptase (Promega, WI, USA), 5µL M-MLV buffer (Promega, WI, USA), 0.5µL RNasin ribonuclease inhibitor (Promega, WI, USA) and 1.25µL nucleotide mix (10µM each) (Roche, Basel, CHE) were added. Reactions were incubated at 42°C for 60 minutes then 70°C for 15 minutes. cDNA was stored at -20°C. Reverse transcriptase negative samples were prepared alongside each reverse transcriptase positive samples by replacing reverse transcriptase with water. PCR and qPCR screening of the reverse transcriptase negative samples would ensure any amplification seen in the reverse transcriptase positive samples was due to synthesised cDNA acting as template and not due to any residual contaminating gDNA from the RNA extraction.

3.5.5 PCR amplification of products from cDNA and gDNA and recovery after electrophoresis

KAPA HiFi DNA polymerase (Kapa Biosystems, MA, USA) was used for amplification of templates for cloning purposes using the manufacturer's instructions. Each reaction contained: 1µL pDNA (Wizard Plus SV Miniprep DNA purification system, Promega, WI, USA); 1µl KAPA HiFi DNA Polymerase (1U); 4µL 5x Kapa Hifi buffer; 0.6µL 10mM KAPA dNTP Mix; 0.6µL 10µM forward primer; 0.6µL 10µM reverse primer; and 12.2µL water. Each reaction underwent the following PCR conditions: 95°C x 3 minutes, then 30 cycles of: 98°C x 20 seconds, 55°C x 15 seconds, 72°C x 30 seconds, before a final extension of 72°C x 3 minutes. When necessary, A-tailing was carried out by adding 1µL of Taq DNA polymerase (Roche, Basel, CHE) and 1µL nucleotide mix (10µM each) (Roche, Basel, CHE) and incubating the reaction at 72°C for 20 minutes. GoTaq Green Master Mix (Promega, WI, USA) was used for screening the gDNA of putatively transgenic plants following the manufacturer directions. Each reaction contained: 10µL 2x GoTaq Green master mix, 1µL gDNA, 1µL

10µM forward primer, 1µL 10µM reverse primer, 7µL water. Each reaction underwent the following PCR conditions: 95°C x 2 minutes, then 30 cycles of: 95°C x 30 seconds, 55°C x 30 seconds, 72°C x 1 minute, before a final extension of 72°C x 5 minutes. After electrophoresis bands required for further reactions were excised and DNA recovered using the Quantum Prep Freeze 'N Squeeze DNA gel extraction spin columns (BioRad, CA, USA).

3.6 qPCR primer design and optimisation and run specifications

3.6.1 qPCR primer design

Primers were designed using sequences available from the NCBI database (Table 3.1). Not all genes had annotated accessions available and consensus sequences were assembled from the sugarcane EST database after BLAST analysis with the equivalent maize gene as a reference sequence. Maize was used as sugarcane and maize are genetically closely related (Selman-Housein *et al.*, 1999). The final consensus sugarcane sequences were created using only sugarcane EST sequences. Both the genes with GenBank accessions and those constructed from EST alignments are seen in Appendix B. β-tubulin was used as the housekeeping gene (Rodrigues *et al.*, 2009). Amplicons of all lignin biosynthesis genes were cloned and two insert-positive *E.coli* colonies were sequenced to determine primer specificity before use in qPCR.

Hydroxycinnamoyl transferase was not included in qPCR analysis as a specific sequence could not be confidently identified. Only one published accession for sugarcane hydroxycinnamoyl transferase was found (Casu *et al.*, 2007; accession: CA210265). This sequence was identified by Casu *et al.* (2007) to be HCT as the EST sequence (CA210265, Vettore *et al.*, 2001) used for the Affymetrix GeneChip Sugarcane Genome Array probe corresponded with an entry (EC 2.3.1.99) in the Kyoto Encyclopedia of Genes and Genomes (Kanehisa and Goto, 2000) identifying as HCT. When analysed by BLAST, accession CA210265 showed very close alignment with *Zea mays* anthranilate N-benzoyltransferase (Soderlund *et al.*, 2009; accession: NM_001153992). This was confirmed by aligning CA210265 with the reverse complimentary of the sequence provided in the NM_001153992 NCBI accession. This alignment showed 628 out of 785 amino acids (80%) aligned between the two sequences across the length of CA210265. Further BLAST searching in the sugarcane nucleotide and EST databases of NCBI with alternative HCT sequences from maize (Barrière *et al.*, 2007; accessions: AY109546, DR807341) and from

TABLE 3.1: qPCR primers designed for the quantification of expression levels of *Zm*MYB transgenes and lignin biosynthesis pathway genes in stably transformed sugarcane

Amplicon	Forward (5'-3')	Reverse (5'-3')	Size (bp)	Maize GenBank Accession/EST numbers	Sugarcane GenBank Accession/EST numbers
β-Tubulin[1]	GGAGGAGTACCCTGACAGAATGA	CAGTATCGGAAACCTTTGGTGAT	68		CA222437[A]
MYB31	TCTTCCGGCTGGAGGACGAG	GTGGCTGTGGCTCTGGCTCTG	80	NM_001112479[B]	
MYB42	ATCAAGGCCGAGGAGACGG	AGAGGTCCAGGTTGAGGTCAG	64	NM_001112539[B]	
PAL	GACATCCTGAAGCTCATGTCG	ACCGACGTCTTGATGTTCTCC	92		EF189195[C]
C4H	GTTCACCGTGTACGGCGACCACT	GAAGAAGGGCACCGTCATGATCC	61	AY104175[D]	CA131376; CA146299; CA196076; CA137884; CA263105
4CL	CTTCCCGACATCGAGATCAACAAC	CTCATCTTCCCGAAGCAGTAGGC	62	AY566301[E]; AX204868[F,G]	CA184118; CA215779; CA136560; CA176600; CA135257
C3H	GTCGACGAGCAGGTCTTCAAAGC	CGTGCTCCTCCATGATCTTCAC	73	AY107051[D]; BT086560[H]	CA262303; CA247763
CCoAOMT	ACCTCATCGCAGACGAGAAGAAC	AGCCGCTCGTGGTAGTTGAGGTAG	91	AJ242980[I]; EU952463[J]; NM_001158013[G]	5' end: CA168805; CA071322; CA159865; CA180815 3' end: CA159865; CF575000[K]; CA279207; CA179873
CCR	AGCAGCCGTACAAGTTCTCG	GAAGGTTCTTCACCGTGTCG	96		AJ231134[L]
F5H	GGTTCATCGACAAGATCATCGAC	GTCGGGGCTCTTCCCGCGCTTCAC	53	AX204869[F,M]	5' end: CA185931; CA134666; CA135938 3' end: CA287472; CA278023; CA253395; CA103877
COMT	TACGGGATGACGGCGTTCGAGTAC	GTGATGATGACCGAGTGGTTCTT	92		AY365419[N]; AJ231133[L]
CAD	ATCAGCTCGTCGTCCAAGAAG	ACCGTGTCGATGATGTAGTCC	128		AJ231135[L]

TABLE 3.1: qPCR primers designed for the quantification of expression levels of *Zm*MYB transgene and lignin biosynthesis pathway genes in stably transformed sugarcane (continued)

[1] Housekeeping primer sequences from Rodrigues *et al.* (2009).

[A] Vettore *et al.* (2001). All EST sequences with the prefix '*CA*' are from the Sugarcane Expressed Sequence Tag project (SUCEST) (Vettore *et al.*, 2001); [B] Fornalé *et al.* (2006); [C] Que (2006); [D] Gardiner *et al.* (2004); [E] Sivasankar (2004); [F] Puigdomenech *et al.* (2001); [G] Andersen *et al.* (2008); [H] Soderlund *et al.* (2009); [I] Civardi *et al.* (1999); [J] Alexandrov *et al.* (2009); [K] Casu *et al.* (2004); [L] Selman-Housein *et al.* (1999); [M] Chen *et al.* (2010); [N] Yang (2003).

Accessions from Que (2006), Sivasankar (2004) and Yang (2003) all come from unpublished data and were direct submissions to the NCBI database. The accessions were found by searching within the NCBI database itself.

MAIZEWALL (Guillaumie *et al.*, 2007; accessions : 2478084.2.1_REV, 2619423.2.1), *Medicago sativa* L. (alfalfa) (Shadle *et al.*, 2007; accession AJ507825), *Capsicum* spp. (Mazourek *et al.*, 2009; accession EU616565), *Nicotiana benthamiana* (Hoffmann *et al.*, 2004; accession: AJ555865), *Coffea arabica* (Salmona *et al.*, 2008; accession: AM116757) and *Triticum aestivum* L. (wheat) (Bi *et al.*, 2011; accessions: CK193498, CK199765) did not highlight any potential sugarcane HCT sequences. Alignment of these HCT sequences did not highlight any conserved regions of sufficient length to design primers (standard or degenerate) for potential use in sugarcane. qPCR reactions were optimised to attain suitable R^2 and PCR efficiency values (Livak and Schmittgen, 2001; Taylor *et al.*, 2011). Primers were then validated against the housekeeping primers to ensure comparable rates of product amplification (Livak and Schmittgen, 2001; Taylor *et al.*, 2011). qPCR reactions utilised GoTaq qPCR Master Mix (Promega, WI, USA) in a 20µL total reaction volume with 20ng of cDNA template using 10mM forward and reverse primers.

3.6.2 qPCR run specifications

Samples were prepared by a CAS1200 robot (Qiagen, Limburg, NLD) into 100 well gene discs (Qiagen, Limburg, NLD), heat sealed with Rotor Disc heat sealing film (Qiagen, Limburg, NLD), and run using a Qiagen Rotor-Gene Q (Qiagen, Limburg, NLD) with the following cycle: 50°C x 2 minutes; 95°C x 2 minutes, then 40 cycles of: 95°C x 15 seconds, 55°C x 30 seconds (acquiring); 72°C x 5 seconds (acquiring); 82°C x 5 seconds (acquiring) followed by a 10 minute melt curve ramping from 72°C to 99°C raising by 1°C each step. Fluorescence was acquired at three different temperatures based on primer optimisation. Relative transcript levels were quantified using delta critical threshold values (ΔCt) as follows: $\Delta Ct = 2^{-(Ct\ gene\ of\ interest - Ct\ housekeeping\ gene)}$ (Levy *et al.*, 2004).

3.7 Compositional cell wall and enzymatic hydrolysis analyses

3.7.1 Determination of cell wall chemical composition and structure

After harvest, tissue for cell wall compositional analysis was prepared as per Hames *et al.* (2008) using convection oven drying at 40°C. Dried samples were milled with a IKA Labortechnik (BW, DE) MFC mill (2mm screen) and packed into Whatman cellulose extraction thimbles (GE Healthcare, BUX, UK) for successive overnight soxhlet washes with

water and ethanol respectively to remove extractives (Sluiter *et al.*, 2008c). Samples were again dried using convection oven drying at 40°C and stored at room temperature. A sample of this prepared material was dried overnight in a convection oven set at 105°C and used to determine the total solids of the bagasse (Sluiter *et al.*, 2008a).

Lignin, cellulose and hemicellulose were quantified by a modified acid hydrolysis method (Sluiter *et al.*, 2008b). Briefly, 0.125g of bagasse was reacted in 1.5mL 72% sulfuric acid for one hour at 30°C inside a pressure tube with regular stiring. Following the addition of 42mL of water, the reaction was autoclaved at 121°C for one hour before samples were vacuum filtrated through a crucible. Acid soluble lignin was determined by UV-Vis spectrophotometry and acid insoluble lignin was measured gravimetrically (Sluiter *et al.*, 2008b). Hydrolysis liquor containing soluble cell wall carbohydrates was analysed using High Performance Liquid Chromatography (HPLC) (Sluiter *et al.*, 2008b). A Waters (MA, USA) e2695 Separations Module and Showa Denko (Bavaria, DE) Shodex SP-0810 sugar column (85°C) with micro-guard de-ashing columns (BioRad, CA, USA) equipped with a Waters (MA, USA) 2414 Refractive Index Detector was employed. Samples for HPLC analysis were neutralised by addition of $CaCO_3$ (50mg/mL) and syringe filtered through a 0.45µm leur lock nylon syringe filter (Banksia, PA, USA) into autosampler vials (Waters, MA, USA). 20µL of prepared sample was injected for analysis with water (0.2µm filtered and degassed) as the eluent (0.5mL/minute) for a run time of 35 minutes. The analysis calculations include a step to convert the sugars from their monomeric form (as measured by HPLC) to their polymeric form as would be present in the original tissue.

3.7.2 Pretreatment of bagasse for enzymatic hydrolysis

Bagasse was ground to a fine powder using a McCrone micronising mill (IL, USA) with ethanol (100%) before undergoing a mild pretreatment. This involved 1% (w/w) sulfuric acid being added in a ratio of 10:1 with bagasse followed by autoclaving (130°C for 30 minutes). Samples were then washed with water (3 x 50mL). The total solids content (%) (Sluiter *et al.*, 2008a) was determined before use for enzymatic hydrolysis. This involved weighing bagasse before drying the bagasse for 24 hours at 105°C to remove all water content and weighing again after drying. The total solids content (%) was calculated from the two weights. Non-pretreated bagasse samples were used as a control during enzymatic hydrolysis to confirm the effectiveness of pretreatment. Samples then underwent acid

hydrolysis as described above (Sluiter *et al.*, 2008b) to quantify lignin, cellulose and hemicellulose contents of the pretreated samples.

3.7.3 Microplate enzymatic hydrolysis of bagasse

Enzymatic hydrolysis of transgenic and control bagasse was performed using Accellerase 1500 (Genencor, CA, USA). Before use the filter paper units (FPU) and protein concentrations were determined. The FPU activity of Accellerase 1500 was calculated using the methods by Adney and Baker (2008) using 50mM citrate buffer and was determined to be 46.8 FPU/mL. The protein content of Accellerase 1500 was determined using a Bradford assay (Bradford, 1976) using BioRad (CA, USA) Protein Assay Dye Reagent Concentrate following the supplied protocols for microtiter plates and was found to be 22.87 mg/mL (±0.38 SEM). Bovine serum albumin (Thermo Fisher Scientific, MA, USA) was used to develop a protein standard curve.

Enzymatic hydrolysis was performed in 200µL tubes following published methods (Harrison *et al.*, 2013). Briefly, ground bagasse samples were mixed with 50mM sodium acetate + 0.02% (w/v) sodium azide to a concentration of 1.3% cellulose (w/v) and rotated overnight by a Suspension Mixer (Ratek, VIC, AUS) at 4°C. A 2x enzyme master mix was prepared containing Accelerase 1500 and *Aspergillus niger* β-glucosidase (Megazyme, Wicklow, IRE) to ensure complete hydrolysis of cellobiose to glucose. The final reaction concentration of Accelerase 1500 was 6 FPU (2.93 µg/g cellulose) and β-glucosidase was 50µg/g cellulose. A 100µL aliquot of the bagasse suspension was mixed with 100µL of the 2x enzyme master mix resulting in a final cellulose concentration of 0.65% (w/v). A low FPU in combination with the mild pretreatment was considered the best approach to highlight any enzymatic performance differences, including subtle differences, due to structural changes in the cell walls of transgenic plants when compared to controls (Van Acker *et al.*, 2013).

Both pretreated and non-treated control bagasse samples were digested in triplicate at 50°C with rotation (Hybaid Shake ‘n’ Stack Hybridisation Oven, Thermo Fisher Scientific, MA, USA) for 72 hours with samples being taken at 0, 6, 12, 24, 48 and 72 hours. Reactions were quenched in liquid nitrogen and stored at -80°C. The glucose released in each sample was analysed using a D-Glucose Assay (GOPOD Format) (Megazyme, Wicklow, IRE) following manufacturer’s instructions and plates were read using a Beckman Coulter AD200C Plate Reader (Beckman Coulter, Inc., CA, USA).

3.7.4 Determination of cellulose crystallinity index in bagasse

A sample was taken from the bagasse that had been finely ground in a McCrone (IL, USA) micronising mill for enzymatic hydrolysis (Section 3.7.2) and did not undergo any pretreatment. This tissue was used to determine cellulose crystallinity index following the methods of Segal *et al.* (1959). X-ray diffraction patterns of cellulose samples were recorded with a Bruker (WI, USA) AXS D8 Advance X-ray diffractometer at room temperature from 10° to 40° using Cu/Kα_1 irradiation (1.54Å) at 40kV and 40mA. The scan speed was 15 seconds/step with a step size of 0.05. Crystallinity Index (CI) was obtained from the relationship between the intensity of the 002 peak for cellulose I (I_{002}) and the minimum dip (I_{am}) between the 002 and the 101 peaks using the equation: CI(%) = (($I_{002} - I_{am}$) / I_{002}) x 100 where I_{002} = intensity at 22.7°A and I_{am} = 18°A. The divergence slit and anti-scatter slit were 3.722°. The program XRD commander (Bruker, WI, USA) was used to collect and analyse the data from the diffractometer.

3.8 Juice extraction and component quantification

Juice was hot-water extracted from internodes ground under liquid nitrogen following the methods of Inman-Bamber *et al.* (2008) before being stored at -20°C until required for analysis. Juice samples were diluted according to ICUMSA method GS7/8/4-24 using lactose as an internal standard and quantified using High Performance Ion Chromatography (HPIC). A Waters (MA, USA) e2695 Separations Module and Dionex CarboPac PA1 HPLC column with guard column (27°C) (Thermo Fisher Scientific, MA, USA) equipped with a Waters (MA, USA) 2465 Electrochemical detector was employed. 20µL of prepared sample was injected for analysis with 150mM NaOH (0.2µm filtered and degassed) as the eluent (1.0mL/minute) for a run time of 15 minutes.

Chapter 4

Profiling of Sugarcane Lignin Biosynthesis

The deposition of the secondary cell wall is an important stage in terrestrial plant development (Cosgrove, 2005; Rogers and Campbell, 2004; Weng and Chapple, 2010). This involves the ordered deposition of cellulose and hemicellulose polysaccharides followed by the impregnation of lignin polymers into this polysaccharide matrix during lignification (Bourquin *et al.*, 2002; Plomion *et al.*, 2001; Vogel, 2008). Lignin polymers are composed of H, G and S monolignols, which are produced via the lignin biosynthesis pathway (Boerjan *et al.*, 2003; Liu, 2012). Lignin biosynthesis is part of the much larger phenylpropanoid pathway (Boerjan *et al.*, 2003) and as a result genes within the lignin biosynthesis pathway can be multifunctional with additional roles outside of lignin biosynthesis. PAL, C4H and 4CL, the enzymes catalysing the first three steps of lignin biosynthesis, are fundamental in the biosynthesis of all phenylpropanoids (Bonawitz and Chapple, 2010; Vogt, 2010), and CCoAOMT and C3H are also hypothesised to have roles in the ferulate production pathway (Barrière *et al.*, 2004; Grabber, 2005).

Due to the importance of lignin in structural stability and water transportation (Ma, 2007; Weng and Chapple, 2010) the role and function of each gene within the lignin biosynthesis pathway is well established (reviewed in Bonawitz and Chapple, 2010). The more recent definition of the relationship between lignin and efficiency of second generation bioethanol production has led to increased focus and research into lignin biosynthesis and manipulation (Section 2.4), and further realises the possibility of cost-competitive bioethanol being produced from lignin-altered sugarcane bagasse. Given the influence lignin has on cell wall digestibility, an expression profile that aligns lignin gene expression levels with cell wall chemical composition is an important contribution to further current understanding of the role gene expression has on lignin biosynthesis and deposition in sugarcane. This information will also be applicable for the genetic modification of plants to specifically alter lignin characteristics.

To date, there has only been minimal gene expression profiling in sugarcane (Casu *et al.*, 2004; Casu *et al.*, 2007; Kolahi *et al.*, 2013; Papini-Terzi *et al.*, 2009) and only one paper analysing lignin deposition during plant maturation (Lingle and Thomson, 2012). The limited expression data available for sugarcane suggest that not all lignin biosynthesis genes have plateaued expression levels during plant development but instead may have increased or decreased expression as tissue matures. RNA levels for CCoAOMT, COMT, CAD and F5H have been found to increase with tissue maturity (Casu *et al.*, 2004; Casu *et al.*, 2007; Papini-Terzi *et al.*, 2009) and PAL, C4H and C3H showed stable expression during maturation (Papini-Terzi *et al.*, 2009). PAL expression has also been found to decrease during plant development (Kolahi *et al.*, 2013; Papini-Terzi *et al.*, 2009), as has the expression of CCoAOMT and HCT (Casu *et al.*, 2007). Only one paper has explored the changes in biomass composition throughout sugarcane development. Lingle and Thomson (2012) examined cell wall compositional data for sugarcane and found that lignocellulosic composition also varies during plant growth. As plants matured, lignin deposition continually increased, whereas hemicellulose levels plateaued and cellulose levels decreased after initial increases in younger tissue (Lingle and Thomson, 2012).

A comprehensive and lignin-focussed profile consisting of gene expression and cell wall compositional data has not been established in sugarcane and only partial lignin profiles have been established or can be inferred from the literature (Cheng *et al.*, 2013; Firon *et al.*, 2013; Huang *et al.*, 2012; Ma, 2007; Shang *et al.*, 2012; Shi *et al.*, 2010). The availability of both gene expression and cell wall compositional data will contribute a useful assessment of lignin formation in sugarcane. From a biotechnological viewpoint this work may highlight specific genes for targeting to affect rate of lignin deposition. Additionally, the identification and cloning of promoters of genes with desirable spatial or temporal expression patterns may enable transgene expression to be focused to a specific tissue type or developmental stage.

4.1 SPECIFIC METHODS

4.1.1 Generation and harvesting of wild type sugarcane for expression profile development

Wild type sugarcane were generated from callus and acclimatised in growth chambers before being transferred to the glasshouse (Section 3.4). It was decided to use tissue

culture generated sugarcane grown in a glasshouse so that plant generation and growth conditions were controlled throughout plant development. Ten plants were grown for nine months before five were randomly selected and destructively harvested for analysis. At this stage, plants had between 15 and 16 internodes. Each harvested plant was divided into five different sections (A-E) (Table 4.1) to represent increasing tissue maturity, with section A being the youngest tissue and section E being the most mature tissue.

All plants were watered to saturation two days before harvesting to avoid results being affected by any potential drought related stress response. Harvesting occurred between 10am and 2pm in a single session to minimise light or circadian related fluctuations in gene expression levels (Pan *et al.*, 2009; Rogers *et al.*, 2005). Before dividing the stems into sections (Table 4.1), all leaf tissue and sheaths were removed. The length of the stalk was measured (internode one to the final internode) to represent plant height, the internodes were counted as per van Dillewijn (1952) and diameter of internodes 2, 4, 6, 8, 10, 12 and 14 was recorded using callipers. For all stem analyses only internode tissue was used. After sectioning the stem, the root ball was washed in water to remove potting mix and ten buttress roots (Moore, 1987) were collected from each plant for qPCR analysis. Plant roots also contain lignin (Bonawitz and Chapple, 2010), though few published lignin expression profiles include analysis of root tissue. Roots were included in the development of this profile to begin to gain a general understanding of overall lignin biosynthesis in this tissue.

4.1.2 Analysis of harvested material

Each stem section (Table 4.1) was represented by five individually analysed tissue samples from five individual plants. Each sample underwent qPCR analysis (Section 3.6) in triplicate following cDNA synthesis (Section 3.5) with primers described previously for nine of the ten lignin biosynthesis genes (Table 3.1). Root tissue was also analysed by qPCR but cell wall composition was not quantified due to limited tissue availability. Cell wall compositional analysis was performed in duplicate on all stem tissue. Grubbs' test for outliers (GraphPad Software, CA, USA) was used to confirm no plants analysed had outlying values for height or number of internodes. ANOVA analysis with Tukey post-hoc analysis (p = 0.05) was employed to assess any statistical differences in internode diameter, gene expression in stem sections and roots, as well as any differences in cell wall composition.

TABLE 4.1: Tissue harvested for development of lignin biosynthesis profile. Each stem was divided into five sections (A-E) with section A representing the youngest tissue and section E representing the most mature tissue as well as collecting root tissue. Each section was comprised of three internodes for qPCR analysis (Method Section 3.6.2) and cell wall compositional analysis (Method Section 3.7.1)

Section	Internode number	Analysis performed
A	1	qPCR gene expression
	2-3	Cell wall compositional analysis
B	4	qPCR gene expression
	5-6	Cell wall compositional analysis
C	7	qPCR gene expression
	8-9	Cell wall compositional analysis
D	10	qPCR gene expression
	11-12	Cell wall compositional analysis
E	13	qPCR gene expression
	14-15	Cell wall compositional analysis
Roots		qPCR gene expression

4.2 RESULTS

4.2.1 Phenotypic comparison of harvested plants

The sugarcane plants used in this profiling experiment were regenerated from a single batch of callus and grown side-by-side from tissue culture generation until harvest, therefore it was expected that they would have similar, though not identical phenotypes. There was some variation in height and number of internodes (Table 4.2) though statistical analysis did not identify any plants as having outlying values. Nor were any statistically significant differences found between plants when internode diameters were compared (Table 4.2). As the five plants presented with similar phenotypes they were considered suitable for the expression profiling.

4.2.2 qPCR expression profiles of lignin biosynthesis genes

Expression profiles for the nine lignin biosynthesis genes (Table 3.1) were established after qPCR analysis of the five stem sections and the root tissue (Figures 4.1, 4.2 and 4.3). The

TABLE 4.2: Phenotypic measurements of five wild type sugarcane plants used for developmental profile. Height (cm), number of internodes, and internode diameter (mm) were recorded at time of harvest.

WT plant number	Height (cm)	Number of internodes	Internode (IN) diameter (mm)						
			IN-2	IN-4	IN-6	IN-8	IN-10	IN-12	IN-14
1	130	15	11.86	13.78	14.48	14.83	13.90	13.04	11.51
2	136	16	12.40	13.42	12.71	11.95	11.16	11.41	12.72
3	138	15	12.17	12.23	13.65	13.73	13.63	12.87	12.55
4	142	16	11.17	13.34	13.76	13.68	13.39	12.31	10.77
5	148	16	13.14	14.06	13.43	12.70	11.59	10.81	12.45

ΔCt values were normalised against section A (Figures 4.1, 4.2 and 4.3; Appendix C). Values were normalised to allow for easier comparison of changes in expression in relation to young tissue for each gene. The raw ΔCt values (Appendix C) show that in section A, PAL is expressed at levels greater than the other eight lignin biosynthesis genes analysed. CCoAOMT, COMT and CCR also had greater expression levels in Section A than CAD, 4CL, C4H, F5H and C3H. These trends in expression levels in section A are also seen throughout the remaining stem sections and root tissue for each gene (Appendix C).

The genes analysed in stem tissue fell into two categories, those with highest expression in section A which then decreased as stem tissue matured (PAL, CCR, 4CL, COMT and CAD) (Figure 4.1, Appendix C), and those whose expression showed little change as tissue aged (C3H, F5H, C4H and CCoAOMT) (Figure 4.2, Appendix C). The different patterns of expression do not appear to fall in a specific pattern based on their position within the lignin biosynthesis pathway (Figure 1.1 and Figure 4.3).

Of the group whose expression is highest in section A before decreasing with stem tissue maturation (Figure 4.1, Appendix C), this trend is strongest in PAL and CCR, with expression in section B decreased by 70% - 80% and by more than 90% in section E. 4CL and COMT are less pronounced in this trend with expression dropping approximately 65% from section A to section B. Expression of 4CL and COMT decrease 65% - 75% in sections C, D and E when compared to expression in section A (Figure 4.1). CAD, the final gene in this group, is the least pronounced in this trend. A 40% decrease in expression between section A and section B is followed by expression levelling out between 25% and 40% of section A in

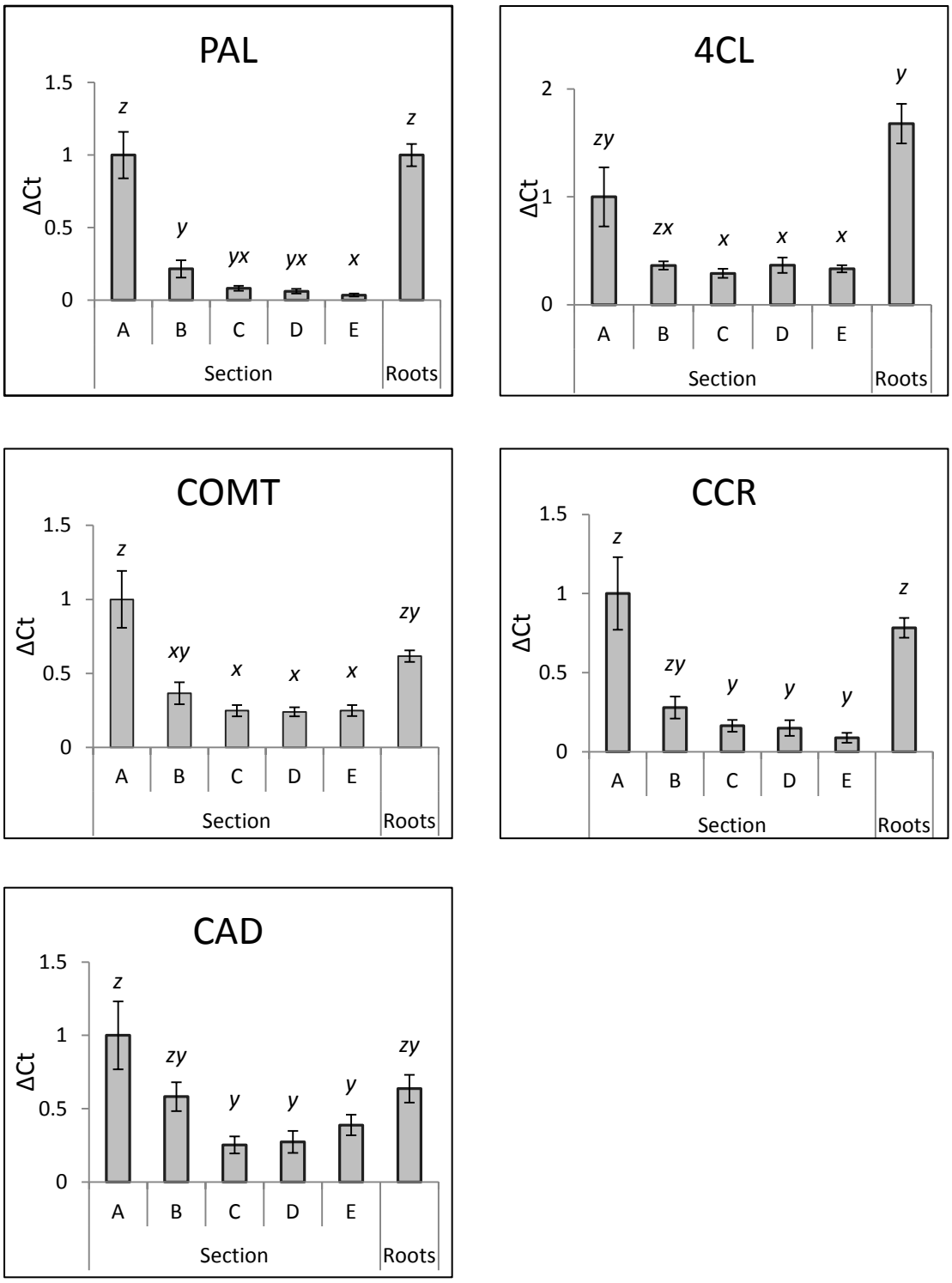


FIGURE 4.1: Genes showing highest expression in section A (young stem tissue) with decreased expression in more mature stem regions. ΔCt expression levels of lignin biosynthesis genes from the five stem sections and roots (n = 5 individual plants per tissue section) normalised against section A for each individual gene is shown with standard error of the mean. Statistical differences are noted by different letters above bars (*x*, *y* and *z*) after ANOVA analysis with Tukey post-hoc analysis ($p = 0.05$).

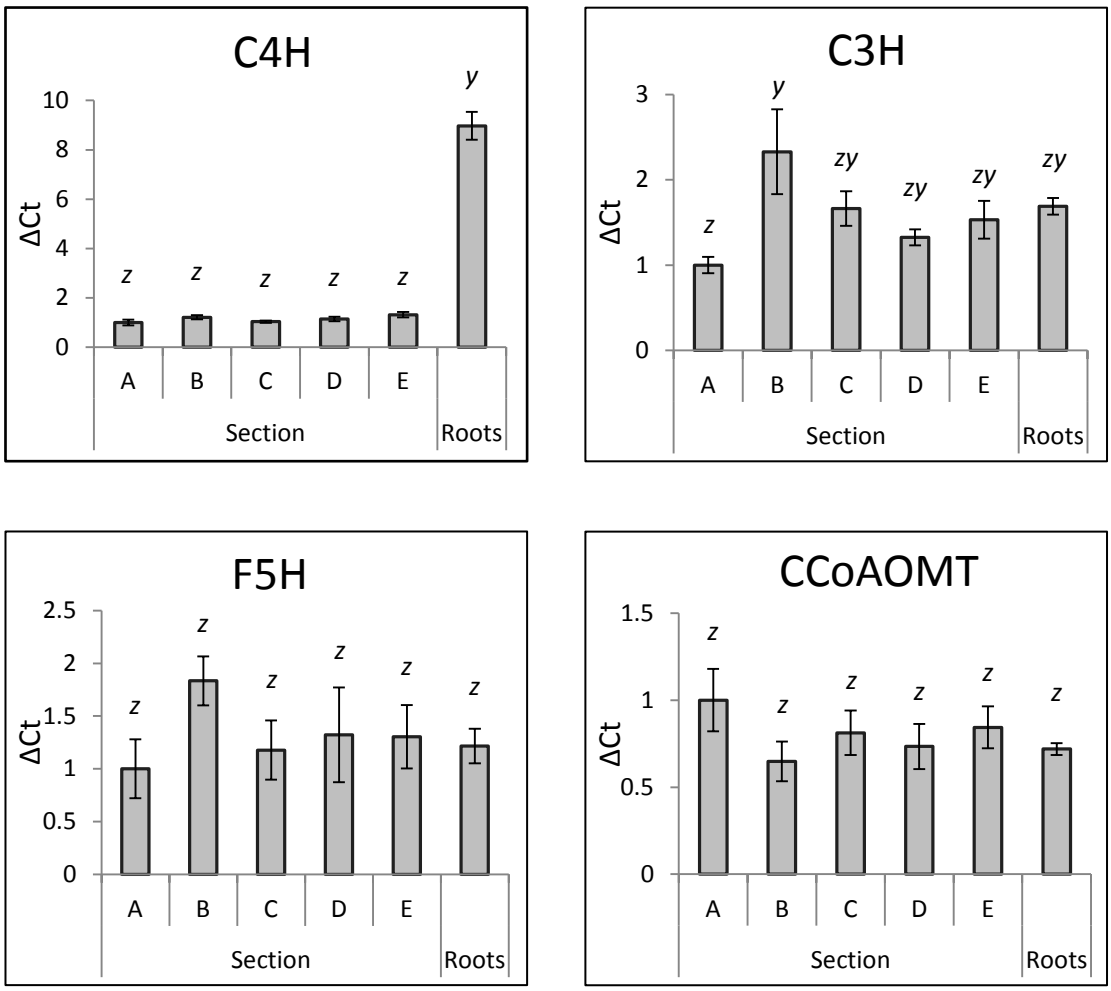


FIGURE 4.2: Genes showing consistent expression across the five stem sections. ΔCt expression levels of lignin biosynthesis genes from the five stem sections and roots (n = 5 individual plants per tissue section) normalised against section A for each individual gene is shown with standard error of the mean. Statistical differences are noted by different letters above bars (*x*, *y* and *z*) after ANOVA analysis with Tukey post-hoc analysis ($p = 0.05$).

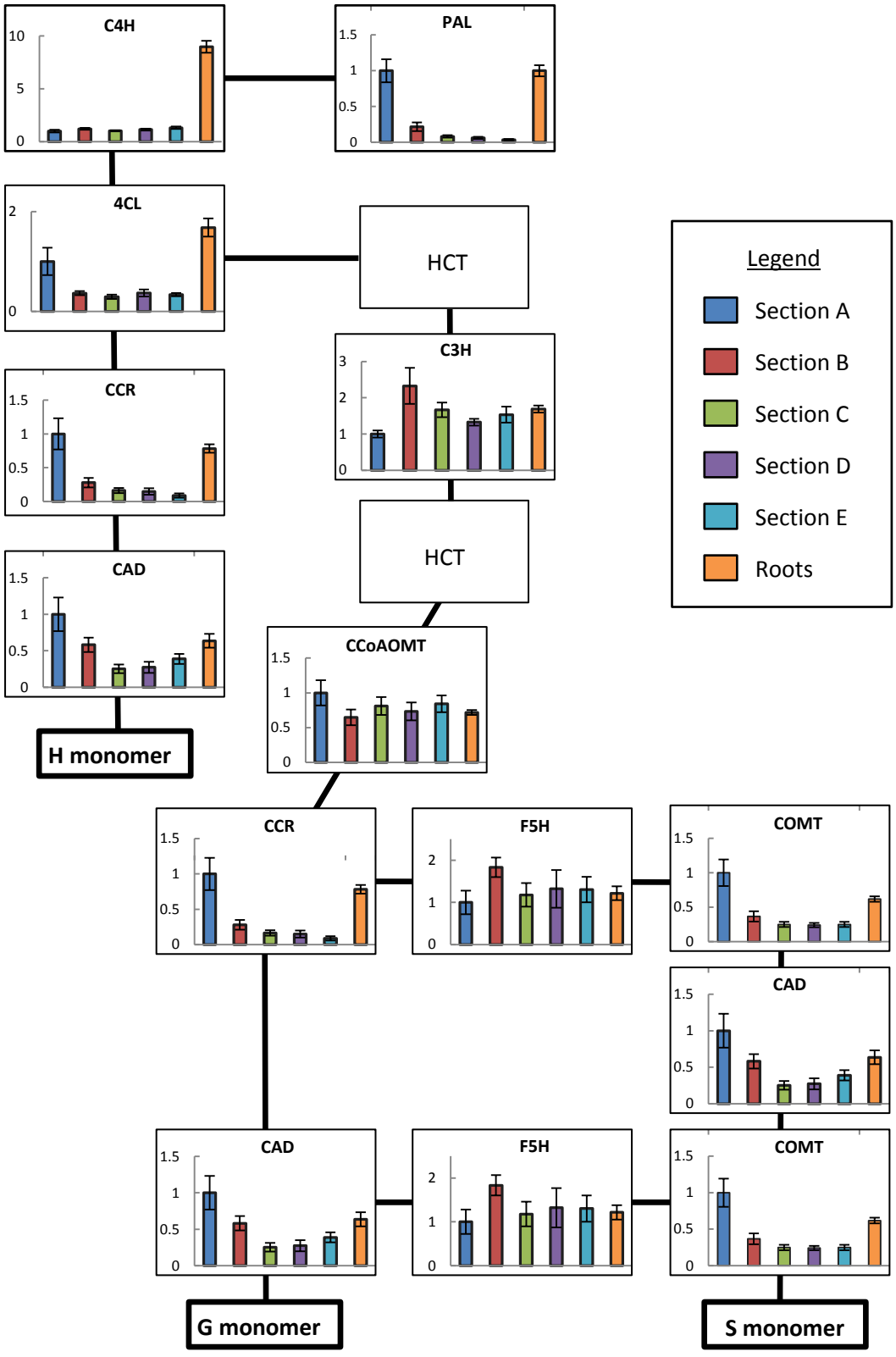


FIGURE 4.3: qPCR expression levels of lignin biosynthesis genes in stem sections A – E and root tissue as seen in Figures 4.1 and 4.2 laid out as lignin biosynthesis pathway based on Hisano *et al.* (2009). Expression levels are normalised to section A for each gene.

sections C, D and E (Figure 4.1). Of the five genes within this group only PAL and COMT show a significant reduction in expression in section B when compared to section A (Figure 4.1). All five genes have significantly less expression in section C, D and E when compared to section A expression levels (Figure 4.1). The second group shows similar expression across all five sections of stem tissue analysed (Figure 4.2, Appendix C). C4H, CCoAOMT and F5H all show no significant differences in expression levels across the five stem sections. C3H shows a significant increase in expression levels between section A and section B before stabilising in sections C, D and E (Figure 4.2).

Expression levels of the nine lignin biosynthesis genes being analysed were also established in root tissue (Figures 4.1 and 4.2, Appendix C). C3H, CCoAOMT, F5H and CAD expression in root tissue was not significantly different to any stem section (A-E) (Figures 4.1 and 4.2). Expression of PAL and 4CL in root tissue were not significantly different to section A, but were significantly higher than sections B-E (Figure 4.1). CCR and COMT showed a similar pattern being not significantly different to sections A or B but significantly higher than expression in sections C-E (Figure 4.1). Expression of C4H in root tissue was approximately 9-fold higher than in any stem section (Figure 4.2).

4.2.3 Cell wall compositional analysis

Secondary cell wall components were quantified in stem sections (Table 4.3). Section A had significantly less lignin than the more mature stem internodes, though levels appear to stabilise after section B. This significant difference was due to lower acid insoluble lignin in section A as there are no significant differences in acid soluble lignin levels across the five stem sections (Table 4.3). There were no significant differences across the five sections for glucose, xylose or galactose amounts, though galactose represents only a very minor component of the cell wall. Section A had significantly more arabinose than the remaining sections, though as for galactose, arabinose also only represents a very small proportion of the cell wall. All mass balances were within 10% of starting material during cell wall compositional analysis.

TABLE 4.3: Accumulation of individual cell wall components in stem regions of increasing maturity. The percentage of each component of the total composition is shown with the standard error of the mean. Values in bold type are significantly different (ANOVA analysis with Tukey post-hoc analysis, p = 0.05) to section A for each component. No significant differences were observed between sections B, C, D or E for any component. Each stem section for each component was represented by five individual plant samples.

Section	Total lignin		Acid insoluble lignin		Acid soluble lignin		Glucose		Xylose		Galactose		Arabinose	
	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-
A	20.76	0.52	15.57	0.48	5.19	0.11	49.26	0.53	20.58	0.35	0.32	0.19	2.28	0.09
B	**22.35**	**0.27**	**17.14**	**0.32**	5.21	0.06	47.93	0.50	20.01	0.29	0.00	0.00	**1.63**	**0.05**
C	**23.39**	**0.13**	**18.29**	**0.10**	5.10	0.08	48.09	0.49	20.36	0.35	0.00	0.00	**1.58**	**0.04**
D	**23.24**	**0.29**	**18.08**	**0.33**	5.16	0.06	48.00	0.54	20.90	0.27	0.00	0.00	**1.58**	**0.06**
E	**22.49**	**0.24**	**17.40**	**0.30**	5.09	0.06	47.34	0.18	21.58	0.46	0.10	0.09	**1.81**	**0.12**

4.3 DISCUSSION

A comparison of lignin biosynthetic gene transcription levels with cell wall composition during maturation is not currently available for sugarcane. As discussed, data exist for a range of plant species which compare some aspects of this developmental relationship, though a focused analysis which combines gene expression levels for the entire lignin biosynthesis pathway with lignin, cellulose and hemicellulose deposition in any plant species is yet to be published. The aim of this research was to fill this knowledge gap for sugarcane, an important crop to the global economy and a potentially important crop for second generation bioethanol production (Canilha *et al.*, 2012; Yuan *et al.*, 2008). Sugarcane plants with lower lignin contents (wild type and genetically modified) have been found to release increased levels of glucose compared with sugarcane plants with higher lignin contents (Benjamin *et al.*, 2013; Jung *et al.*, 2012; Jung *et al.*, 2013; Masarin *et al.*, 2011). Improved understanding of lignin biosynthesis and deposition in sugarcane will be of great value when deciding the most appropriate approaches to facilitate the development of commercial lines with increased saccharification potential.

4.3.1 Expression profile of lignin biosynthesis pathway genes

The trends in the stem expression data (Figures 4.1 and 4.2) dichotomise the lignin biosynthesis genes: expression decreases with tissue age (PAL, CCR, 4CL, COMT and CAD) or expression remains constant during maturation (C3H, F5H, C4H and CCoAOMT). The position of the genes within the lignin biosynthesis pathway and their expression pattern do not appear linked (Figure 4.3) suggesting the function of the gene may influence its expression more than its location within the biosynthetic pathway.

4.3.1.1 Gene expression decreases with stem tissue age

The trend of decreased gene expression as tissue matures is most strongly seen in PAL and CCR (Figure 1.1). This finding is supported by Papini-Terzi *et al.* (2009) who found some PAL homologues showed decreased expression in maturing sugarcane stem tissue when compared to younger tissue, but is in contrast to findings by Casu *et al.* (2004) and Kolahi *et al.* (2013) that PAL expression in sugarcane increased with stem maturity. Phenylpropanoid biosynthesis, which includes lignin biosynthesis, begins after the deamination of phenylalanine by PAL (Bonawitz and Chapple, 2010; Ferrer *et al.*, 2008; Halpin, 2004). As

PAL catalyses the entry of metabolites into the lignin biosynthesis pathway (Liu, 2012; Weng and Chapple, 2010) its high level of expression in younger tissue found in this study (Figure 4.1, Appendix C) may represent an initial metabolic flux to provide a burst of metabolites for the various phenylpropanoid pathways including lignin biosynthesis.

As opposed to PAL, CCR functions in the final stages of lignin biosynthesis and is considered a committed step which is key in the production of H, G and S lignin monomers (Vogt, 2010; Weng and Chapple, 2010). Given its positioning in the lignin biosynthesis pathway (Figure 1.1) CCR may additionally act as a regulating control point for directing the metabolic flux into lignin monomer production (Lacombe *et al.*, 1997). As the high expression of PAL in young tissue may be to stimulate metabolic flux into phenylpropanoid production, the high expression of CCR in young tissue may be to ensure a high level of metabolite commitment into lignin biosynthesis, which is fundamentally important for healthy plant development (Ma, 2007; Weng and Chapple, 2010).

To a lesser degree 4CL and COMT share the same expression patterns as PAL and CCR. 4CL is an essential enzyme in phenylpropanoid biosynthesis (Bonawitz and Chapple, 2010) and represents an important branch point where metabolites are directed into lignin biosynthesis or alternative phenylpropanoid biosynthesis pathways (Vogt, 2010; Weng and Chapple, 2010). Its position allows for direct metabolite contribution into H monomer biosynthesis or redirection of metabolites for G or S monomer biosynthesis (Vogt, 2010). The high level of 4CL expression in young tissue may reflect its response to the metabolic flux into the phenylpropanoid pathway initiated by PAL.

In this study COMT expression was found to decrease with tissue maturity. This is contrary to the findings of Casu *et al.* (2004) and Papini-Terzi *et al.* (2009) in which COMT expression was found to increase as sugarcane stem maturity increased. COMT is the last of two enzymes entirely responsible for the production of the S lignin monomer within the lignin biosynthesis pathway (Bonawitz and Chapple, 2010; Weng and Chapple, 2010). The increased expression of COMT in young tissue in this research may be to ensure S monomer production during the availability of the initial metabolic flux.

The final gene showing a reduction in expression as stem tissue matures was CAD, though the trend was not as strong as the previously discussed genes. Consistent with our results, Cheng *et al.* (2013) found CAD expression decreased before plateauing with stem tissue age

in *Ginkgo biloba*, whereas Casu *et al.* (2007) found CAD expression in sugarcane increased with tissue age. CAD represents the final enzyme in the lignin biosynthesis pathway catalysing the production of precursor monolignols and committing them to H, G and S lignin monomer synthesis (Ferrer *et al.*, 2008; Halpin, 2004; Vogt, 2010). The initial high expression of CAD in young tissue may relate to the increased metabolic flux through the lignin biosynthesis pathway. The continued production of lignin throughout plant development (Rogers and Campbell, 2004) may require CAD to retain a certain level of activity which could explain why its expression does not decrease to the levels of PAL, CCR, 4CL or COMT (Figure 4.1). It was hypothesised that the high expression levels of PAL, CCR, 4CL, COMT and CAD in young tissue (section A) (Figure 4.1) was to initiate lignin biosynthesis in the developing sugarcane internodes. This is supported by the finding that lignin content significantly increased in section B, relative to section A (Table 4.3), which may be a result of increased lignin biosynthesis gene expression levels in section A.

To reiterate, this research found PAL, COMT and CAD have high expression in young stem tissue, which then decreases as the stem tissue matures. This is in contrast to findings in sugarcane that PAL (Casu *et al.,* 2004; Kolahi *et al.*, 2013), COMT (Casu *et al.*, 2004; Papini-Terzi *et al.*, 2009), and CAD (Casu *et al.*, 2007) expression increased with stem maturity. The discrepancies between the current and published research may be a result of various experimental differences between the current research and published findings. Different results may have arisen from the use of different sugarcane cultivars (no published research analysed KQ228) or data collection methods as the current profile used qPCR, whereas Casu *et al.* (2004); Casu *et al.* (2007); and Papini-Terzi *et al.* (2009) used microarrays to generate expression data. Although qPCR is commonly used to validate microarray data, the results produced by these two methods do not always correlate well (Dallas *et al*., 2005; Git *et al*., 2010; Morey *et al*., 2006). This research selected qPCR for quantifying gene expression levels as the genes being analysed are from a well-known and studied pathway, the lignin biosynthesis pathway, and qPCR allowed for flexibility in primer design to suit the specific requirements of this research.

Different growing conditions may also have affected results as the sugarcane studied by Kolahi *et al.* (2013) and Papini-Terzi *et al.* (2009) was field grown in Iran and Brazil respectively. The sugarcane analysed by Casu *et al.* (2004) and Casu *et al.* (2007) was field grown in Queensland, however, there was a minimum of six years difference between the sugarcane harvesting and sampling dates between Casu *et al.* (2007) and the current

research, which may have resulted in different seasonal weather patterns, such as sunlight levels, during periods of plant growth. Comparing glasshouse grown sugarcane with the field grown sugarcane in the published studies may also introduce differences to growth conditions. Within the glasshouse, plants are protected from weather conditions, such as wind, and therefore may require less lignin to remaining upright, whereas, the field grown sugarcane may require additional lignin content to prevent lodging (Ma, 2007; Weng and Chapple, 2010), therefore requiring the genes of the lignin biosynthesis pathway to remain active for longer during plant maturation. The glasshouse also acts to protect plants from insects, which can wound plant stems and induce lignification (Rogers and Campbell, 2004) which is a potential hazard in field grown plants, and may complicate analysis if insect wounded plants were used.

4.3.1.2 Gene expression constant throughout stem development

Four genes within the lignin biosynthesis pathway presented with a plateaued expression pattern across stem sections of increasing age (Figure 4.2), namely C4H, C3H, F5H and CCoAOMT. Papini-Terzi *et al.* (2009) also found C3H and C4H showed no difference in expression levels in young and maturing sugarcane stem, however in their study, expression levels of F5H were higher in maturing sugarcane stem when compared to younger stem tissue. Other discrepancies with published results include the expression levels of CCoAOMT in sugarcane being found to be increased in maturing sugarcane stem when compared to younger stem tissue (Casu *et al.*, 2004). However, another study by Casu *et al.* (2007) found in the maturing tissue of sugarcane CCoAOMT expression was reduced when compared to young stem tissue. In contrast, our results show expression levels of CCoAOMT did not change with stem maturation. Although plants of a similar age and same cultivar were analysed, the differences in findings by Casu *et al*. (2004) and Casu *et al*. (2007) may result from the use of different microarrays, requiring different sample preparation and analysis methods. Additional reasons for the discrepancies between current and published data have been previously discussed in Section 4.3.1.1.

C4H, C3H and F5H are cytochrome P450 monooxygenases (Liu, 2012; Weng and Chapple, 2010). P450 enzymes are versatile catalysts acting as oxygenases by incorporating molecular oxygen into reactions (Porter and Coon, 1991) which is a key step in assembling complex molecules (Nelson, 2006). As a result, P450 enzymes are involved in many different biosynthesis pathways including lignin biosynthesis (Nelson, 2006). C4H catalyses

the first aromatic hydroxylation of the phenylpropanoid skeleton after its deamination by PAL (Ferrer *et al.*, 2008; Weng and Chapple, 2010). This step (along with the reactions catalysed by PAL and 4CL) is one of the three essential steps in phenylpropanoid biosynthesis (Bonawitz and Chapple, 2010; Vogt, 2010). The steady expression of C4H throughout tissue maturation may aid in continuing to funnel metabolites into the phenylpropanoid biosynthesis pathways. F5H, another hydroxylase, is paired with COMT in the S monomer branch of the lignin biosynthesis pathway and its expression is necessary for the production of S lignin monomers (Weng and Chapple, 2010). The expression pattern of F5H may ensure continued flow of metabolites for S lignin monomer production.

C3H, related to C4H, catalyses the second aromatic hydroxylation reaction in the lignin biosynthesis pathway (Weng and Chapple, 2010) and is an important hub in controlling metabolic flux into G and S lignin monomer synthesis (Barrière *et al.*, 2004). Caffeoyl CoA 3-O-methyltransferase (CCoAOMT) is responsible for the 3' methylation of caffeoyl-CoA to produce feruloyl-CoA (Hisano *et al.*, 2009; Raes *et al.*, 2003), a key step in the production of G and S lignin monomers (Figure 1.1) (Barrière *et al.*, 2004). CCoAOMT, along with C3H are hypothesised to be important control points for cell wall lignification by acting as part of the ferulate production pathway (Barrière *et al.*, 2004). The feruloyl residues aid in cross-linking within the cell wall and may increase the resistance of the cell wall to hydrolysis by adding to its structural stability (Barrière *et al.*, 2004; Bonawitz and Chapple, 2010; Grabber, 2005). The steady expression of CCoAOMT and C3H within the maturing sugarcane stem may reflect their continued requirement for feruloyl residue production for ongoing cell wall lignification and not just their role in lignin monomer biosynthesis.

4.3.1.3 Gene expression levels in root tissue

Lignin is present in root tissue to aid in water and nutrient transport through its hydrophobic properties (Bonawitz and Chapple, 2010), but is not often examined during the development of lignin expression profiles. One reason for this may be that the majority of lignin research focuses on lignin in light of pulping or biofuels, in which case, the stem or stalk of the plant is the tissue of interest, rather than the roots. Firon *et al.* (2013), Ma (2007), Kolahi *et al.* (2013) and Cheng *et al.* (2013) represent the only reports in current literature that report lignin gene expression levels in root tissue and only Firon *et al.* (2013) examined the expression of more than one lignin biosynthesis gene. Cheng *et al.* (2013) found CAD expression in *Ginkgo biloba* roots to be slightly higher than the expression levels

in young stem tissue, similar to the expression of CAD found in the current study. PAL expression was detected in sugarcane roots of a similar age to this research and was comparable to PAL expression levels in young stem tissue (Kolahi *et al.*, 2013) which also supports the current findings. Firon *et al.* (2013) detected the expression of PAL, C4H, 4CL, CCoAOMT, CCR, COMT and CAD in sweet potato root. The same genes were also found to be expressed in sugarcane roots in this research. In contrast to the current research, Ma (2007) was unable to detect CCR expression in the root tissue of wheat. Although this research only determined lignin biosynthesis gene expression levels in combined root tissue representing all developmental stages, and therefore is not a complete expression profile within this tissue, these results do provide an insight into the expression levels in sugarcane roots after growing for nine months in a glasshouse. A more comprehensive profile of lignin gene expression patterns in root tissue may involve dividing the root into different developmental stages before analysis.

There are no significant differences in expression levels between root tissue and the five stem sections (A-E) for C3H, CCoAOMT, F5H and CAD (Figures 4.1 and 4.2). Interestingly C3H, CCoAOMT and F5H are all in the group with plateaued gene expression during development and may highlight the promoters of these three genes as potential biotechnological tools to drive continuous and even expression of transgenes in sugarcane stem and root tissue. Further research would be required to assess the expression of C3H, CCoAOMT, F5H and CAD during different stages of root development, as well as in additional tissues, such as leaves. Expression of PAL, 4CL, CCR and COMT in root tissue was not significantly different to section A (youngest stem sample) but was significantly higher than expression in more mature stem tissues (Figure 4.1). The only gene with an unexpected level of expression was C4H which had approximately 9-fold higher expression in roots than in any stem section (Figure 4.2). Although no literature was found which could shed light on this finding, this result does suggest that the promoter of C4H may be useful for root preferential expression of transgenes in sugarcane, however, further analysis, including the functionality of this promoter in additional tissue types, such as leaves, would need to be assessed.

4.3.2 Cell wall compositional changes during sugarcane stem maturation

The composition of the cell wall changes as a plant matures as a result of secondary cell wall deposition. After cell elongation has ceased, the secondary cell wall is formed

(Cosgrove, 2005; Rogers and Campbell, 2004; Weng and Chapple, 2010) which involves the initial deposition of cellulose and hemicellulose, followed by the lignification of this polysaccharide matrix (Bourquin *et al.*, 2002; Plomion *et al.*, 2001; Vogel, 2008). Within sugarcane, rapid elongation of young internode cells precedes cell wall thickening, including lignification, which is indicative of internode maturation (Casu *et al.*, 2007). As the internode diameters (Table 4.2) were not significantly different this would suggest cell expansion had already occurred before samples were collected. No significant differences were seen in glucose, xylose or galactose levels throughout sections (Table 4.3) indicating that the deposition of structural carbohydrates into the secondary cell wall had also occurred before harvesting of samples. Although arabinose decreased with stem maturity, the maximum difference in levels only accounted for 0.7% of the secondary cell wall content (Table 4.3). The significant increase in lignin content between section A and section B (Table 4.3) may reflect the order of the formation of the secondary cell wall. As lignin is incorporated after polysaccharide deposition, the secondary cell walls in section A may have been under construction when harvested, compared to mature cell walls in sections B through E. Polysaccharide components had been fully deposited but lignin deposition was ongoing. Results suggest the lignin deposition was complete by section B as lignin content plateaued and no differences were detected between sections B, C, D and E (Table 4.3).

In the youngest tissue (section A) there was significantly less lignin than sections B – E (Table 4.3). Cheng *et al.* (2013), Ma (2007) and Jung and Casler (2006) also found that lignin content increased with tissue maturity in *Ginkgo biloba,* wheat and maize. However unlike the current study in which lignin content plateaued, Cheng *et al.* (2013) found that lignin content in *Ginkgo biloba* continued to increase with stem maturity. The literature also presents some different patterns of lignification than that seen in the current study. Ma (2007) found acid soluble lignin content in wheat decreased as stem tissue matured however this study found acid soluble lignin content to be stable throughout stem maturation. Acid insoluble lignin is precipitated during acid hydrolysis, whereas acid soluble lignin tends to be composed of low molecular weight phenolic compounds derived from lignin and is not precipitated from the hydrolysis solution (Schwartz and Lawoko, 2010; Yasuda *et al*., 2001). Before increasing and plateauing, Jung and Casler (2006) initially found maize lignin content decreased, which was not seen in this research.

Two different experiments by Lingle and Thomson (2012) involved either sampling a marked sugarcane internode over a period of 12 weeks as it matures from being internode

1 to approximately internode 13, or, harvesting odd-numbered sugarcane internodes between 1 and 11 (inclusively) at single time points in July and September of the same year. The first experiment found lignin content increased continuously over time, whilst the second experiment found lignin content was highest in internode 1 but significantly decreased in internode 3 before slowly increasing again with tissue age to approximately 75% of the initial content of internode 1 (Lingle and Thomson, 2012). As a possible explanation for this pattern of lignification, Lingle and Thomson (2012) cited findings by Jung and Casler (2006). Young maize tissue is comprised of a higher percentage of lignified protoxylem vessels than more mature tissue, which initially results in a high lignin content in young tissue (Jung and Casler, 2006). The development of non-lignified tissues dilute the initial lignin concentration, causing a decrease in lignin content, before increasing as the lignified secondary cell wall is deposited (Jung and Casler, 2006). When comparing patterns of lignin deposition, neither experiment by Lingle and Thomson (2012) support the current findings in which lignin significantly increased in section B, relative to section A, before plateauing.

The significant increase in lignin content in section B (Table 4.3) may be a result of the increased metabolic flux detected at the gene expression level in section A (Figures 4.1 and 4.2). As the expression levels of five of the lignin biosynthesis genes decrease with tissue maturity (Figure 4.1) so too may the rate of lignin biosynthesis and deposition resulting in no significant increases or decreases in lignin content in more mature stem sections (Table 4.3). The cell wall compositional results in this current study represent the proportions (%) of bagasse dry weight that the cell wall component accounts for, and not the actual amounts (Table 4.3). This presents another alternative explanation as to the plateauing of cell wall components after section B in this study (Table 4.3) when compared with increasing lignin content in maturing tissue seen in published studies (Cheng *et al.*, 2013; Lingle and Thomson, 2012; Ma, 2007; Jung and Casler, 2006). A possibility exists that both lignin and the structural carbohydrates are increasing in the maturing internodes in proportion to each other, while the water content of the maturing stem decreases. Therefore, while lignin content is increasing, the percentage each cell wall component accounts for in dried bagasse does not change, hence the plateauing of components after section B (Table 4.3).

Apart from lignin content, there is little change in the composition of cell wall carbohydrates throughout stem development as glucose and xylose levels do not change significantly across the five sections analysed (Table 4.3). This is in contrast to published

findings in sugarcane (Lingle and Thomson, 2012) and maize (Jung and Casler, 2006). Both experiments by Lingle and Thomson (2012) found cellulose peaked in week three (Experiment 1) and in internode five (Experiment 2) before slowly declining or plateauing as tissue matured whereas hemicellulose was highest in young tissue in both experiments before reducing and plateauing as tissue matured (Lingle and Thomson, 2012). The decrease in cellulose and hemicellulose coincided with an increase in sucrose concentration, which may be due to the carbon which was originally partitioned for structural polysaccharides instead being used in sucrose production (Lingle and Thomson, 2012). Jung and Casler (2006) found similar results in maize in which glucose content increased as tissue matured before plateauing, and hemicellulose (xylose and arabinose) decreased as tissue matured before also plateauing. Jung and Casler (2006) found the decrease in xylose and arabinose coincided with an increase in ferulates, and suggested that the ferulates may be replacing the xylose and arabinose within the cell wall, hence their decrease during tissue maturation.

The disparity between the current results and those of Lingle and Thomson (2012) for lignin, cellulose and hemicellulose deposition may be due to a number of factors. Although the internodes harvested were similar (internodes 1-13 and 1-11 in Experiments 1 and 2 respectively (Lingle and Thomson, 2012), and internodes 2-15 in the current research (Table 4.1)), there are two main differences between the current research and Lingle and Thomson (2012). The first is that Lingle and Thomson (2012) analysed North American cultivars whereas this research focused on an Australian cultivar. The second is Lingle and Thomson (2012) sampled field grown sugarcane whereas this experiment used glasshouse grown sugarcane. Lingle and Thomson (2012) found a significant difference in lignin content between the July and September harvests in 2009 but not in 2008, and suggested the environment may have a greater influence on lignin content than growth stage. If correct, the glasshouse environment may affect lignin content differently to field conditions and may have influenced the differences in lignin deposition found in this research and published findings. An experiment comparing glasshouse and field conditions on growth and development would be necessary to confirm this hypothesis.

4.3.3 Limitations and future research

When the results found in this study are compared with the current literature there are some published results that support the current findings (Cheng *et al.*, 2013; Jung and

Casler, 2006; Kolahi *et al.*, 2013; Papini-Terzi *et al.*, 2009) and others that have found differing patterns of gene expression and cell wall deposition (Casu *et al.*, 2004; Casu *et al.*, 2007; Jung and Casler, 2006; Lingle and Thomson, 2012; Papini-Terzi *et al.*, 2009). Because biosynthesis pathways for phenylpropanoid production, including lignin, undergo complex regulation (reviewed in Gray *et al.*, 2012) measuring gene expression levels in static intervals (harvesting then freezing tissue) may not fully elucidate the complexity of lignin gene expression. It is also likely that different plant species will have different lignin biosynthesis expression patterns, which may also contribute to differences between current findings and published findings in the literature.

Another potential limitation of this research can be summarised by the experiment by Bi *et al.* (2011) in which 32 gene homologues representing the ten lignin biosynthesis genes in wheat were assessed for their roles in lignin production during plant development and when challenged with pathogens. This research aimed to highlight homologues specifically involved in lignin biosynthesis during plant growth as potential targets for down-regulation to decrease lignin content without increasing pathogen susceptibility. Bi *et al.* (2011) found that different homologues express at different levels during the different conditions and to compare homologue gene expression levels with lignin deposition would be very difficult. Given the complexity of the sugarcane genome (Grivet and Arruda, 2001) it would not be surprising if homologues exist for sugarcane lignin biosynthesis genes. As the complete sugarcane genome for cultivar KQ228 is not currently available it was not possible to design qPCR primers to bind all homologues or particular homologues so it is currently unknown if the results in this experiment are a reflection of individual genes or groups of gene homologues. The steps taken when designing primers are outlined in Section 3.6.1. One potential solution that was unavailable due to time restrictions would be to design multiple sets of qPCR primers per lignin biosynthesis gene and confirm if a similar expression pattern is observed. Another option would be to sequence the cloned qPCR amplicons for each gene and examine sequences for mixed transcripts. Although amplicons were sequenced during primer design (Section 3.6.1) to determine primer specificity, only two *E.coli* colonies were sequenced per amplicon. This was considered too few colonies to make any conclusions concerning the potential for gene homologues.

In addition, the glasshouse used to grow the sugarcane for this profile did not have adequate ceiling height to allow the plants to grow to their full height. The plants were grown as tall as possible without the ceiling interfering with their growth. It is expected the

results are still reflective of full height plants and given that the expression patterns of the biosynthesis genes and the levels of secondary cell wall components had stabilised after Section B, the usefulness of additional data that could be obtained is not clear. Using setts or mature field grown plants would be an alternative solution. It was decided to use sugarcane generated by tissue culture and grown in a glasshouse with controlled conditions so that plant origins and growth conditions were comparable, the same of which may not be true for field grown plants or the setts grown from these plants. In addition, glasshouse experiments allowed for control of watering and fertilisation regimes as well as exclusion of biotic and abiotic damage. Further research may involve comparing tissue culture generated and field grown plants to determine if tissue cultured sugarcane grown in glasshouse conditions is comparable with field conditions. Additional research may also examine enzyme activity levels of lignin biosynthesis genes as an extra set of data linking gene expression levels with the deposition of secondary cell wall components. Gene expression and enzyme activity levels of cellulose and hemicellulose biosynthesis pathways would provide additional information into the production of structural polysaccharides during plant development and maturation.

Chapter 5

MYB gene expression in sugarcane

MYB transcription factors have a wide spectrum of regulatory influences over metabolism and development including the regulation of the lignin biosynthetic pathway (Ambawat *et al.*, 2013; Fornalé *et al.*, 2010). Many MYB genes have been identified with up-regulatory roles in lignin deposition (Table 2.1) and a modest number of MYB genes with down-regulatory influences have also been identified (Table 2.2). Of the MYB genes from the latter category two were further analysed in this research: *Zm*MYB31 and *Zm*MYB42.

Previous research identified the maize *Zm*MYB31 and *Zm*MYB42 transcription factors (Fornalé *et al.*, 2006) using degenerate primers designed from *Antirrhinum majus* MYB genes *Am*MYB308 and *Am*MYB330, which had previously been shown to have down-regulatory influence over lignin biosynthesis when expressed in tobacco (Tamagnone *et al.*, 1998). It was hypothesised that the maize MYB genes would also negatively influence lignin deposition (Fornalé *et al.*, 2006).

When *Zm*MYB31 and *Zm*MYB42 were independently expressed in transgenic Arabidopsis they were found to reduce the expression of genes within the lignin biosynthesis pathway which led to significant decreases in the lignin content of the transgenic Arabidopsis (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009). During enzymatic hydrolysis the reduced lignin content allowed for a greater release of fermentable sugars from the *Zm*MYB31 and *Zm*MYB42 expressing plants when compared to controls (Fornalé *et al.*, 2010; Sonbol *et al.*, 2009), thus improving the bioethanol potential of these plants. Recent research found the over-expression of *Pv*MYB4 in switchgrass reduced lignin content through the down-regulation of lignin biosynthesis genes which increased the saccharification of these plants three-fold (Shen *et al.*, 2012; Shen *et al.*, 2013). Phylogenetic analysis of *Pv*MYB4 found it was most closely related to *Zm*MYB31, *Zm*MYB42 and *Zm*MYB38 than the other MYB transcription factors comprising subgroup G4 (Du *et al.*, 2012a; Fornalé *et al.*, 2010). This further highlights the potential of MYB transcription factors in the production of second generation bioethanol.

If the expression of *Zm*MYB31 and *Zm*MYB42 in sugarcane emulates the findings of Sonbol *et al.* (2009) and Fornalé *et al.* (2010) and improves saccharification, then this characteristic would benefit the production of second generation bioethanol from, and increase the monetary value of sugarcane bagasse. The assessment of this possibility required the completion of a number of specific aims:

1) Clone MYB31 and MYB42 genes from maize and construct plasmids for the transformation and regeneration of sugarcane plants harbouring and expressing either *Zm*MYB31 or *Zm*MYB42
2) Assess any regulatory effects the expression of *Zm*MYB31 or *Zm*MYB42 may have on the expression levels of genes within the lignin biosynthesis pathway
3) Quantify secondary cell wall components (lignin, cellulose and hemicellulose) in sugarcane plants showing down-regulation of lignin biosynthesis pathway genes
4) Perform enzymatic hydrolysis on plants with reduced lignin contents to determine if this increases the release of fermentable glucose monomers
5) Assess whether juice sucrose levels are influenced by the expression of *Zm*MYB31 or *Zm*MYB42

5.1 SPECIFIC METHODS

5.1.1 Generation of constructs harbouring maize MYB genes

5.1.1.1 Maize embryo rescue

Kernels were excised from an ear of corn (*Zea mays*) purchased at local markets. After surface sterilisation, the embryos were excised using aseptic techniques and germinated in liquid MS regeneration media at 27°C with a 16 hour photoperiod. Leaf and shoot tissue was harvested after two weeks of growth and used for the cloning of the MYB transcription factors (MYB31 and MYB42).

5.1.1.2 PCR amplification and cloning of *Zm*MYB31 and *Zm*MYB42

RNA extracted from combined maize leaf and stem tissues was used for cDNA synthesis. Primers were designed from available sequence information to amplify MYB31 and MYB42

genes from the maize cDNA (Table 5.1). Two amplicons were cloned for each MYB gene to either include or exclude adjacent sequence of the 5' and 3' untranslated regions (UTR) as there is evidence that retaining these sequences can regulate and improve gene expression (Hughes, 2006; Mignone *et al.*, 2002; Pesole *et al.*, 2001; Sharma and Sharma, 2009; Wilkie *et al.*, 2003). 'MYB UTR' refers to DNA with the inclusion of untranslated regions and 'MYB ORF' refers to the open reading frame only. MYB31 UTR and MYB42 UTR had 128bp and 64bp of 5' UTR and 251bp and 32bp of 3' UTR respectively (Mignone *et al.*, 2002).

TABLE 5.1: Primers designed from GenBank accessions for PCR amplification of *Zm*MYB31 and *Zm*MYB42 genes from maize cDNA.

Amplicon	Accession	Forward (5'-3')	Reverse (5'-3')	Size (bp)
MYB31 ORF	NM_001112479	ATGGGGAGGTCGCCGTGCTG	TCATTTCATCTCGAGGCTTCT	819
MYB31 UTR	NM_001112479	ACAGCAGCAGCAACAACAAC	TGGCGATGGTGATTACAGAG	1198
MYB42 ORF	NM_001112539	ATGGGGCGGTCGCCGTGCTGC	TCACTTCATCTCCAGGCCTCT	780
MYB42 UTR	NM_001112539	ACTCGCTGCCTTCTCAAATC	GGAGAAGAAAGGACGTGTGG	876

The resulting PCR products were A-tailed and cloned into the pGEM-T Easy Vector System (Promega, WI, USA). These plasmids were transformed into XL1 Blue *E. coli* and white colonies screened for the insert (Table 5.1) by restriction digest and agarose gel electrophoresis. The pGEM-T Easy Vector System plasmids (Promega, WI, USA) were purified from positive colonies and subsequently sequenced. Inserts (Table 5.1) were subcloned into the *SmaI* site of the existing alkaline phosphatase (Roche, Basel, CHE) treated *Zm*Ubi-iUbi-nos/pBlueScript (Agilent Technologies, CA, USA) entry vector containing the maize Ubiquitin promoter (Ubi) and intron (iUbi) (Christensen and Quail, 1996; Christensen *et al.*, 1992) and the nopaline synthase (nos) terminator (Bevan *et al.*, 1983a), resulting in four *Zm*Ubi-iUbi-MYB-nos/pBS constructs. Sequences for the maize Ubiquitin promoter (Ubi) and intron (iUbi) and the nopaline synthase (nos) terminator are seen in Appendix D. A vector map for MYB expression plasmids is seen in Appendix E. All cloning junctions have been sequenced as above.

5.1.2 Generation of transgenic sugarcane with MYB gene constructs

5.1.2.1 Transformation of callus with *Zm*Ubi-iUbi-MYB-nos/pBS constructs

Callus was co-bombarded with individual *Zm*Ubi-iUbi-MYB-nos/pBS constructs and *Zm*Ubi-nptII-nos/pUC19. Neomycin phosphotransferase II (nptII) confers G418 disulfate salt (Roche, Basel, CHE) resistance to callus for selection (Bevan *et al.*, 1983b; Colbére-Garapin *et al.*, 1981). Transformation control callus was co-bombarded with one microgram *Zm*Ubi-GFP-nos/pUC19 constructs with one microgram of *Zm*Ubi-nptII-nos/pUC19 to confirm that the resistance gene (nptII) did not affect lignin biosynthetic genes. GFP is the gene encoding the green fluorescence protein isolated from *Aequorea victoria* (Chalfie *et al.*, 1994; Chiu *et al.*, 1996; Prasher *et al.*, 1992). GFP was included in the bombardment of transgenic control callus as the detection of transient GFP expression in callus provided an indication that the microprojectile bombardment conditions had not adversely affected the expression cassettes. Untransformed control (wild type) callus did not undergo microprojectile bombardment and would allow for further analysis and comparisons of the effects of the MYB transcription factors over the lignin biosynthesis pathway. Regeneration of callus was carried out as previously described, after which individual events were transferred to growth chambers for continued development.

5.1.2.2 Screening regenerated events for transgene presence and expression

Individual leaves were excised from MYB-transformed sugarcane and controls two to four weeks after acclimatisation in the growth chamber for gDNA screening for the presence of the MYB gene cassette. All samples underwent Rapid Release gDNA extractions before endpoint PCR. Primers were designed to bind within the 3' end of the Ubi promoter and 5' end of the MYB coding sequence of the gene of interest. This determined both the presence of the MYB gene as well as the general integrity of the expression cassette. At six weeks after acclimatisation, endpoint PCR positive plants for the MYB transgene cassette had leaf tissue sampled for RNA extraction, cDNA synthesis and qPCR quantification of *Zm*MYB31 or *Zm*MYB42 transcript level as well as the expression of nine lignin biosynthesis pathway genes (Table 3.1). Genomic DNA and cDNA extracts from wild type and GFP plants were included as controls and water replaced DNA in no template controls for both endpoint and qPCR reactions respectively. Primers used for both endpoint and qPCR are seen in Table 5.2.

TABLE 5.2: Primers designed from GenBank accessions for leaf gDNA screening for *Zm*MYB31 and *Zm*MYB42 gene cassettes and qPCR quantification of MYB gene expression in young and maturing internodes.

	Amplicon	Forward (5'-3')	Reverse (5'-3')	Size (bp)
gDNA screening	MYB31ORF	GCGGTCGTTCATTCGTTCTA	GATCAGGGACCACTTGTTGC	699
	MYB31UTR	GCGGTCGTTCATTCGTTCTA	GATCAGGGACCACTTGTTGC	827
	MYB42ORF	GCGGTCGTTCATTCGTTCTA	GATGAGCGACCACTTGTTCC	699
	MYB42UTR	GCGGTCGTTCATTCGTTCTA	GATGAGCGACCACTTGTTCC	763
qPCR quantification	MYB31ORF	TCTTCCGGCTGGAGGACGAG	GTGGCTGTGGCTCTGGCTCTG	80
	MYB31UTR	TCTTCCGGCTGGAGGACGAG	GTGGCTGTGGCTCTGGCTCTG	80
	MYB42ORF	ATCAAGGCCGAGGAGACGG	AGAGGTCCAGGTTGAGGTCAG	64
	MYB42UTR	ATCAAGGCCGAGGAGACGG	AGAGGTCCAGGTTGAGGTCAG	64

5.1.2.3 Glasshouse acclimatisation and harvesting of qPCR positive plants

Events harbouring the MYB gene and having detectable levels of MYB gene expression were transferred to the glasshouse and grown for nine months before being destructively harvested for analysis. Growing tissue culture generated sugarcane under glasshouse conditions allowed for plant generation and growth conditions to be controlled throughout plant development, and therefore, any changes in lignin biosynthesis could be attributed to the MYB transgenes and not to the growing environment, which would not be possible in field conditions. All plants were watered to saturation two days before harvesting to avoid results being affected by any potential drought related stress response. Harvesting occurred between 10am and 5pm and occurred over four consecutive days. Longer harvesting hours than previously used (Section 4.1.1) were required to harvest all plants over as few days as possible. Before measuring and cutting, all leaf tissue and sheaths were removed, and the internodes were counted as per van Dillewijn (1952). The length of the stalk was measured (internode 1 to the final internode), the number of internodes was counted and the diameter of internodes 2, 4, 6, 8, 10, 12 and 14 was recorded. Average internode length was calculated by dividing height by total number of internodes. For all

analyses only internode tissue was used. The internodes collected for each assay can been seen in Table 5.3, alongside the General Methods Chapter sections with the associated protocols. These internodes were selected after discussion with the supervisory team and based on required biomass for each proposed experiment.

TABLE 5.3: Internodes harvested for MYB sugarcane analyses.

Internode number	Analysis conducted	Method section with protocol
1	qPCR gene expression analysis in young tissue	3.6.2
7	qPCR gene expression analysis in maturing tissue	3.6.2
8-10	Cell wall and enzymatic hydrolysis analysis	3.7
14	Juice analysis	3.8

5.1.3 Selection of plants and order of analyses

Analyses on glasshouse harvested sugarcane plants were performed in a predetermined order. Initially cDNA was extracted from the internodes harvested for young and maturing stem for qPCR analysis. Seven individual plants per MYB construct (28 MYB plants in total) were selected for this analysis based on these plants having the greatest number of lignin biosynthesis genes down-regulated after initial qPCR analysis of leaf tissue. qPCR was also carried out on nine wild type and three GFP control plants. These same plants also underwent acid hydrolysis to quantify their cell wall composition. Although analyses of subsequent sugarcane generations would provide information into the stability of MYB transgene expression, it was decided to analyse the initial regenerated sugarcane as the timeframe of this research and the lengthy timeframe of analysing multiple sugarcane generations were not compatible. Additionally, the space available at Carseldine glasshouse was insufficient for the growth of multiple sugarcane generations and therefore analysing the first generation of sugarcane was considered to be the best approach for this research.

Enzymatic hydrolysis was performed on three plants per MYB line with four wild type and three GFP transgenic control plants over 72 hours with sampling at six different time points after pretreatment of bagasse. The three MYB plants per line were selected based on having the lowest lignin contents after compositional analysis. The four wild type controls were selected at random and all GFP control plants were used. A random selection of plants with sufficient tissue also underwent enzymatic hydrolysis without pretreatment beforehand. A comparison of the results with and without pretreatment allowed the effectiveness of the pretreatment method selected to be assessed. Quantification of juice sugar components was performed on the plants selected for enzymatic hydrolysis only.

Statistical analysis involved either a two-tailed t-test assuming unequal variance, $p = 0.05$, or a one-way ANOVA with Tukey post-hoc analysis, $p = 0.05$, as appropriate, comparing transgenic plants to transgenic controls. As phenotypic measurements could only be made once per transgenic plant, the number of standard deviations (*z* scores) for each MYB plant measurement were calculated against the GFP transgenic controls. Measurements were considered different to controls if a *z* score greater than 2 or -2 was calculated.

5.2 RESULTS

5.2.1 Amplification and cloning of *Zm*MYB31 and *Zm*MYB42

5.2.1.1 DNA and amino acid alignments of the cloned and published sequences of *Zm*MYB31 ORF and *Zm*MYB31 UTR

The cloned MYB31 ORF and MYB31 UTR nucleotide sequences showed 100% identical matches over the coding sequence. When this coding sequence was aligned with the published sequence (Fornalé *et al.*, 2006) there was 100% alignment within the sequence coding for the R2 domain and a single base pair difference was seen in the R3 domain. Alignment of the translated amino acid sequences showed this to be a silent mutation. Downstream of the R2 and R3 domains there were three additional SNPs which were also found to be silent mutations after translation and an in-frame deletion of nine nucleotides in the cloned sequences, resulted in the deletion of three amino acids within the C-terminus end of the translated protein. Within the nucleotide sequence of the cloned MYB31 5' UTR there were two consecutive SNPs and a single SNP when aligned with the

published sequence. Within the 3' UTR there were two insertional mutations and two individual SNPs when compared to the published sequence. Nucleotide and amino acid alignments are seen in Appendix F.

5.2.1.2 DNA and amino acid alignments of the cloned and published sequences of *Zm*MYB42 ORF and *Zm*MYB42 UTR

The cloned MYB42 ORF and MYB42 UTR nucleotide sequences showed complete alignment except for a single SNP within the coding region of the R3 domain. This SNP was found to be a silent mutation and the translated amino acid sequences of the cloned MYB42 sequences aligned 100%. The cloned MYB42 ORF and UTR nucleotide sequences showed one and two SNPs respectively within the R3 domain when compared against the published nucleotide sequence (Fornalé *et al.*, 2006). Both SNPs were found to be silent mutations as the R2 and R3 domains showed 100% alignment when the cloned and published nucleotide sequences are translated into amino acids and aligned. Downstream of the R2 and R3 nucleotide coding sequences there were six SNPs and three consecutive in-frame base pairs deletions seen in the cloned MYB42 sequences when compared with the published sequence. After amino acid alignment, three SNPs were found to be silent mutations and the remaining three resulted in amino acid substitute mutations. The missing base pairs resulted in the deletion of a single amino acid in the C-terminus of the translated protein. The cloned 5' and 3' MYB42 UTR nucleotide sequences aligned 100% with the published sequence. Nucleotide and amino acid alignments are seen in Appendix F.

5.2.2 Regeneration and analysis of stably transformed sugarcane plants harbouring maize MYB genes

The four MYB constructs were successfully bombarded into sugarcane callus, along with the control construct (*Zm*Ubi-GFP-nos/pUC19). Total plants regenerated can be seen in Table 5.4. Each plant regenerated was considered an independent transgenic event as only one plant per callus piece was subcultured during tissue culture. Endpoint PCR screening of leaf tissue gDNA identified sugarcane plants containing the MYB gene of interest. Three wild type and three GFP controls plants were used as controls during endpoint PCR screening, with no MYB gene cassettes detected in any control. Endpoint PCR confirmed incorporation of the *Zm*Ubi-GFP-nos/pUC19 cassette into gDNA of transgenic controls. Endpoint PCR positive plants underwent qPCR screening of leaf tissue cDNA to confirm the number of

TABLE 5.4: Total number of MYB sugarcane analysed at each stage of analysis. This includes total plants regenerated after microprojectile bombardment; plants identified as MYB transgene positive after endpoint PCR screening of leaf gDNA (shown as positive plants/total number of plants analysed); plants positively expressing the MYB transgene after qPCR screening of leaf cDNA (shown as positive plants/total number of plants analysed). The number of plants analysed post maturation and harvest is also shown.

Plant	Total plants regenerated	Initial leaf PCR	Initial leaf qPCR	Harvested plants qPCR	Cell wall analysis	Enzymatic hydrolysis	Juice analysis
WT	20	3	6	9	4	4	3
GFP	11	3	5	3	3	3	3
MYB 31 ORF	17	17/17	14/17	7	3	3	3
MYB 31 UTR	36	30/36	19/30	7	3	3	3
MYB 42 ORF	30	26/30	10/26	7	3	3	3
MYB 42 UTR	38	36/38	13/36	7	3	3	3

transgenic events showing MYB transgene expression (Table 5.4). All plants showing MYB transgene expression also had the expression levels of nine lignin biosynthesis genes quantified. The control plants showed no *Zm*MYB transgene expression for either MYB31 or MYB42. The qPCR positive MYB expressing plants and controls were transferred to the glasshouse for further growth and maturation. During acclimatisation, only three GFP controls survived and were able to be harvested for subsequent analyses (Table 5.4).

5.2.3 qPCR analysis of lignin biosynthesis pathway genes in MYB expressing sugarcane

qPCR analysis was performed on cDNA synthesised from young and maturing sugarcane internodes for each control and transgenic plant. Seven individual plants for each MYB construct were analysed. They were selected as they had the highest number of lignin biosynthesis genes being down-regulated after the initial qPCR analysis of leaf tissue (Appendix G). The MYB expressing plants were run alongside nine randomly selected wild type and three GFP control plants. This was to assess the down-regulatory effect the

expression of the MYB genes may be having in young tissue and in more developed tissue. It also allowed comparison of MYB transgene expression levels in different tissues to assess if any transgene silencing was occurring.

The majority of MYB31 and MYB42 (Figure 5.1) plants continued to express the MYB transgene in both young and maturing internode tissue, although the levels of expression for some plants, MYB31 UTR11 and MYB42 UTR16 for example, are very low. Of the plants that did not show detectable levels of MYB expression it appears that this was only in young tissue and MYB transgene expression was detectable in maturing tissue of the same plants. There also appeared to be a trend of increased MYB transgene expression in maturing tissue when compared with young tissue. No transgene expression was detected in any tissue of any control plant. After confirming the MYB transgene was still being expressed in each plant, expression levels of the lignin biosynthesis genes were quantified by qPCR from the same cDNA samples (Appendix G). A summary of the regulatory effects that each MYB construct had over the lignin biosynthesis pathway can be seen in Table 5.5. The lignin gene expression levels in MYB31 and MYB42 expressing plants analysed by enzymatic hydrolysis are presented (Figure 5.2) as a representative sample of the total number of plants analysed (Appendix G).

The up- and down-regulatory trends discussed are overall impressions. Although it appears some genes are down-regulated by MYB31 and MYB42, these trends are not consistently supported by statistical analysis (Figure 5.2). For example, C3H appears to be down-regulated by both MYB genes, but no statistical differences were found (Figure 5.2). This was considered to be due to the large variation in expression levels amongst the UKN control group. The overall trends suggest that *Zm*MYB31 down-regulated more genes than *Zm*MYB42 when constitutively expressed in sugarcane (Table 5.5, Figure 5.2, Appendix G). Plants expressing MYB31 had a down-regulatory pattern which is spread across genes that are both early and late in the lignin biosynthesis pathway (Figure 1.1), whereas MYB42 expression appeared to down-regulate the early pathway genes more so than the later pathway genes (Table 5.5). In common, both MYB genes down-regulated C4H and C3H and showed an increase in CAD expression. Although not seen in MYB31 plants, MYB42 plants also showed increases in CCoAOMT, CCR and COMT expression. As well as down-regulating more genes of the lignin biosynthesis pathway overall, MYB31 also appeared to be more consistent in gene regulation across the three different tissue types when compared to MYB42.

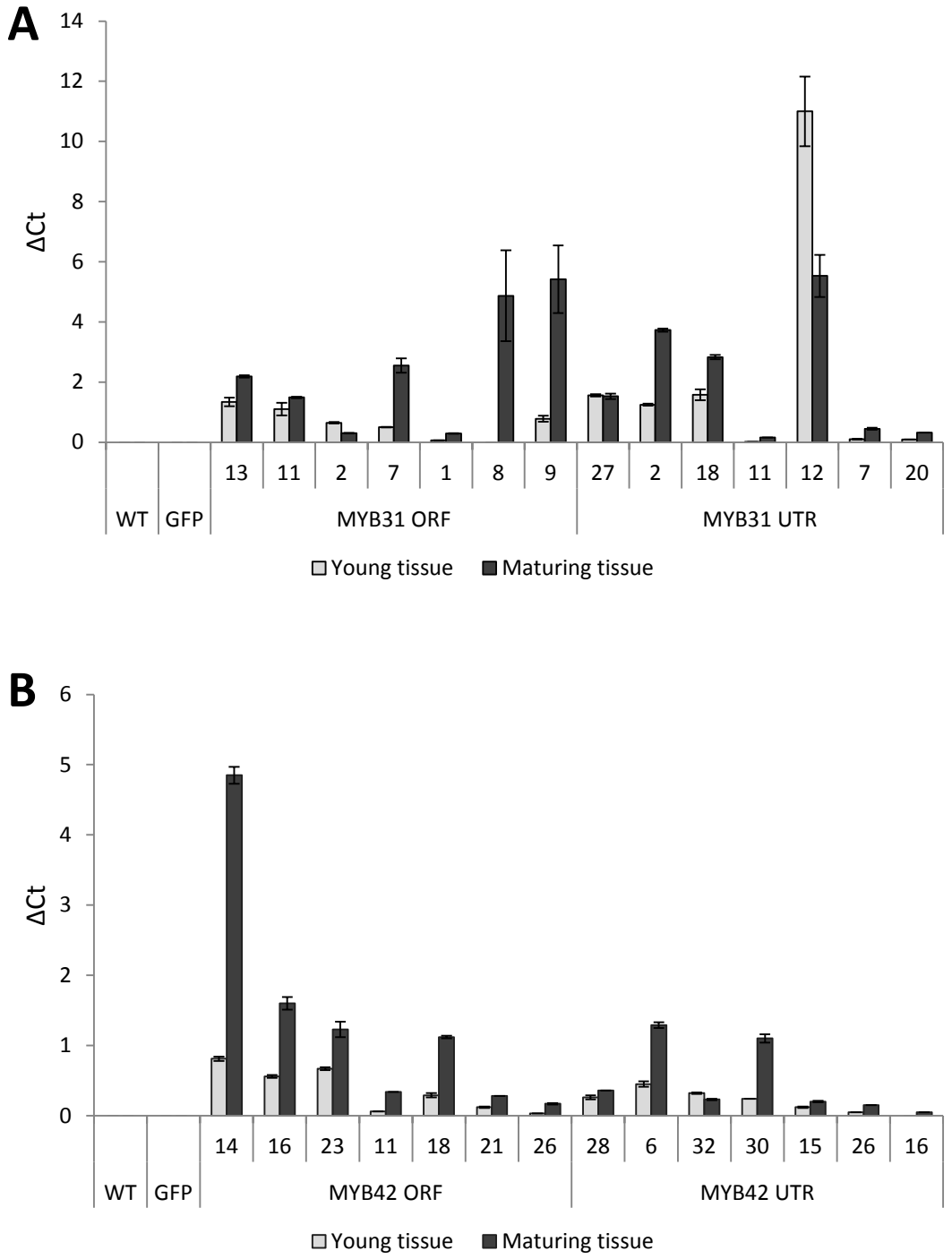


FIGURE 5.1: ΔCt values showing standard error of the mean of (A) *Zm*MYB31 and (B) *Zm*MYB42 expression in the MYB transformed sugarcane plants after qPCR analysis of young and maturing internode tissues post-harvest. Each sample underwent qPCR in triplicate. Plants are listed in ascending total lignin content for each line. WT n = 9; GFP n = 3; n = individually analysed plants. Statistics were not performed on this data as there was no MYB expression detected in controls.

TABLE 5.5: Summary of the general trends and regulatory effects[1] of *Zm*MYB31 and *Zm*MYB42 over lignin biosynthesis gene expression levels after expression in Arabidopsis (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009) and the current results in sugarcane[2].

Lignin gene	**Published *Zm*MYB31**	***Zm*MYB31 sugarcane**			**Published *Zm*MYB42**	***Zm*MYB42 sugarcane**		
		leaf	**y**	**m**		**leaf**	**y**	**m**
PAL		↓	-	↓	↓	↓	↑	~↓
C4H	-	↓	↓	↓	↓	↓	↓	↓
C3H	↓	↓	↓	↓	-	↓	↓	↓
4CL	↓	↓	↓	~↓	↓	↑	~↓	~↓
HCT					↓			
CCoAOMT	-	↓	~↓	↓	-	~↓	~↑	~↑
F5H	↓	~↓	↓	↓	↓	~↓	~↓	~↓
CAD	~↑	↓	~↑	~↑	↓	↓	↑	↑
CCR		↓	↓	~↓	-	~↓	~↑	~↑
COMT	↓	↓	↓	~↓	↓	↓	~↑	~↓

[1] Arrows represent influence on gene expression levels (up or down regulated) compared against wild type and GFP control plants. A dash represents no change in expression levels. A blank cell indicates transcription levels for that gene were not quantified or reported. y = young stem tissue, m = maturing stem tissue. ~ represents a slight but not obvious trend of increased/decreased expression.

[2] The findings provide an overall impression of the up or down-regulation of gene expression for each MYB gene. For each construct (when compared to controls) there are MYB transgenics with decreased expression, no change or increased expression for each lignin gene. The results for each individual plant can be seen in Appendix G.

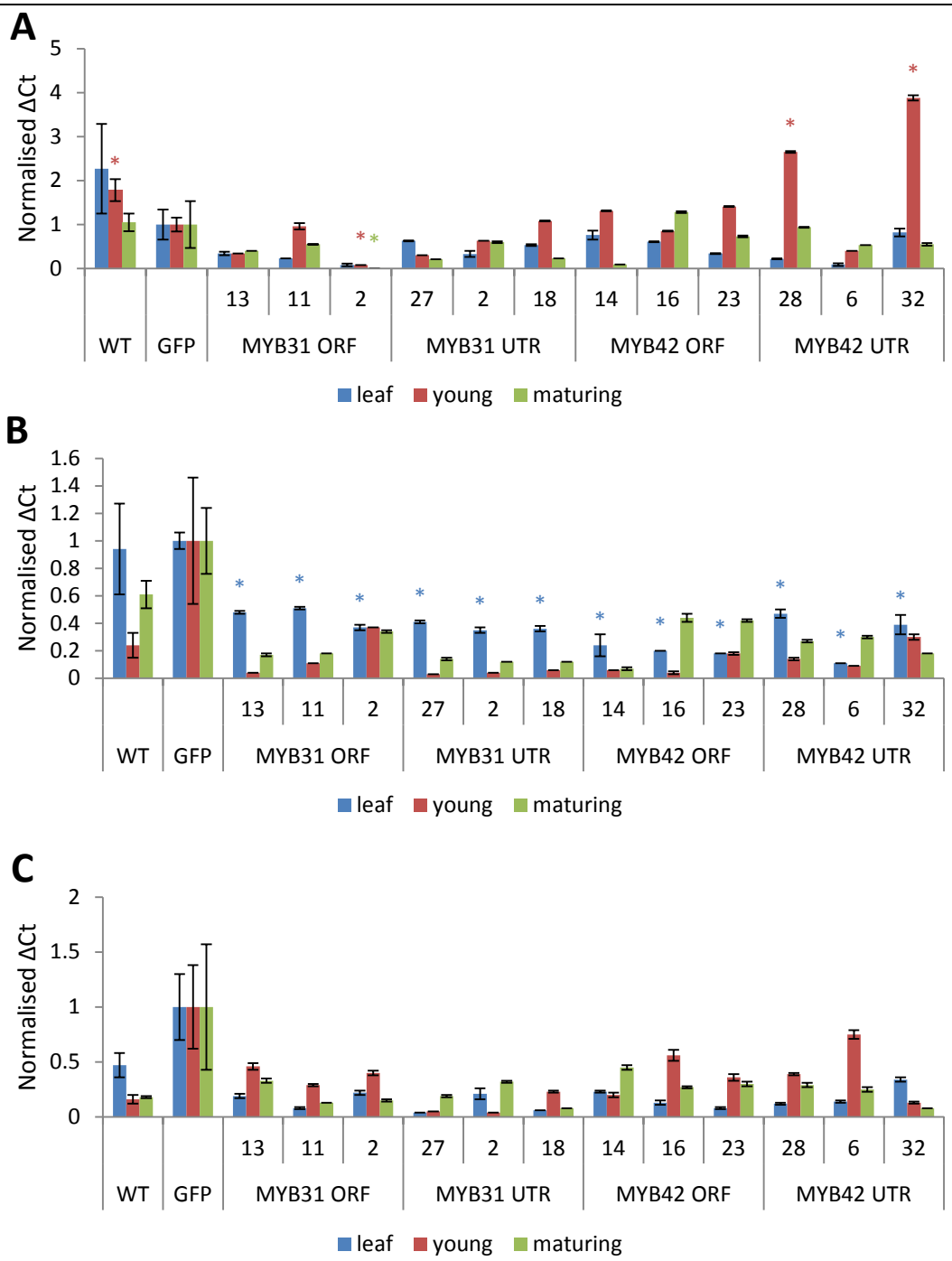


FIGURE 5.2: ΔCt expression of lignin biosynthesis genes (showing standard error of the mean) in sugarcane expressing *Zm*MYB31 or *Zm*MYB42 normalised against GFP controls after qPCR analysis for PAL (A); C4H (B); and C3H (C). Leaf (blue), young (red) and maturing (green) refers to gene expression levels in leaf tissue, young internode tissue and maturing internode tissue respectively. MYB plants statistically different to GFP controls after a two-tailed t-test assuming unequal variance, p = 0.05, are shown by a coloured asterisk respective to tissue type. WT n = 9; GFP n = 3; n = individually analysed plants.

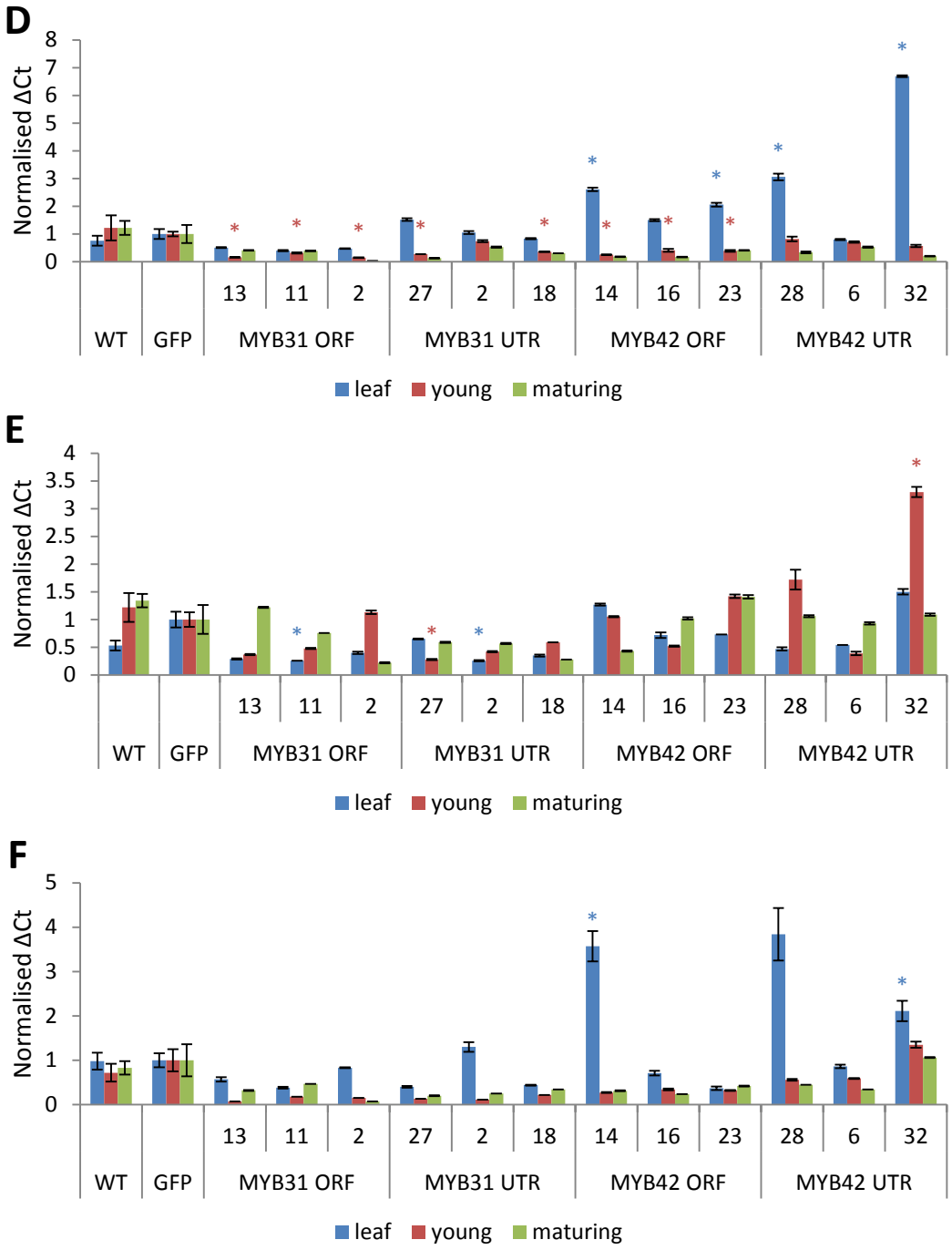


FIGURE 5.2 (continued): ΔCt expression of lignin biosynthesis genes (showing standard error of the mean) in sugarcane expressing *Zm*MYB31 or *Zm*MYB42 normalised against GFP controls after qPCR analysis for 4CL (D); CCoAOMT (E); and F5H (F). Leaf (blue), young (red) and maturing (green) refers to gene expression levels in leaf tissue, young internode tissue and maturing internode tissue respectively. MYB plants statistically different to GFP controls after a two-tailed t-test assuming unequal variance, p = 0.05, are shown by a coloured asterisk respective to tissue type. WT n = 9; GFP n = 3; n = individually analysed plants.

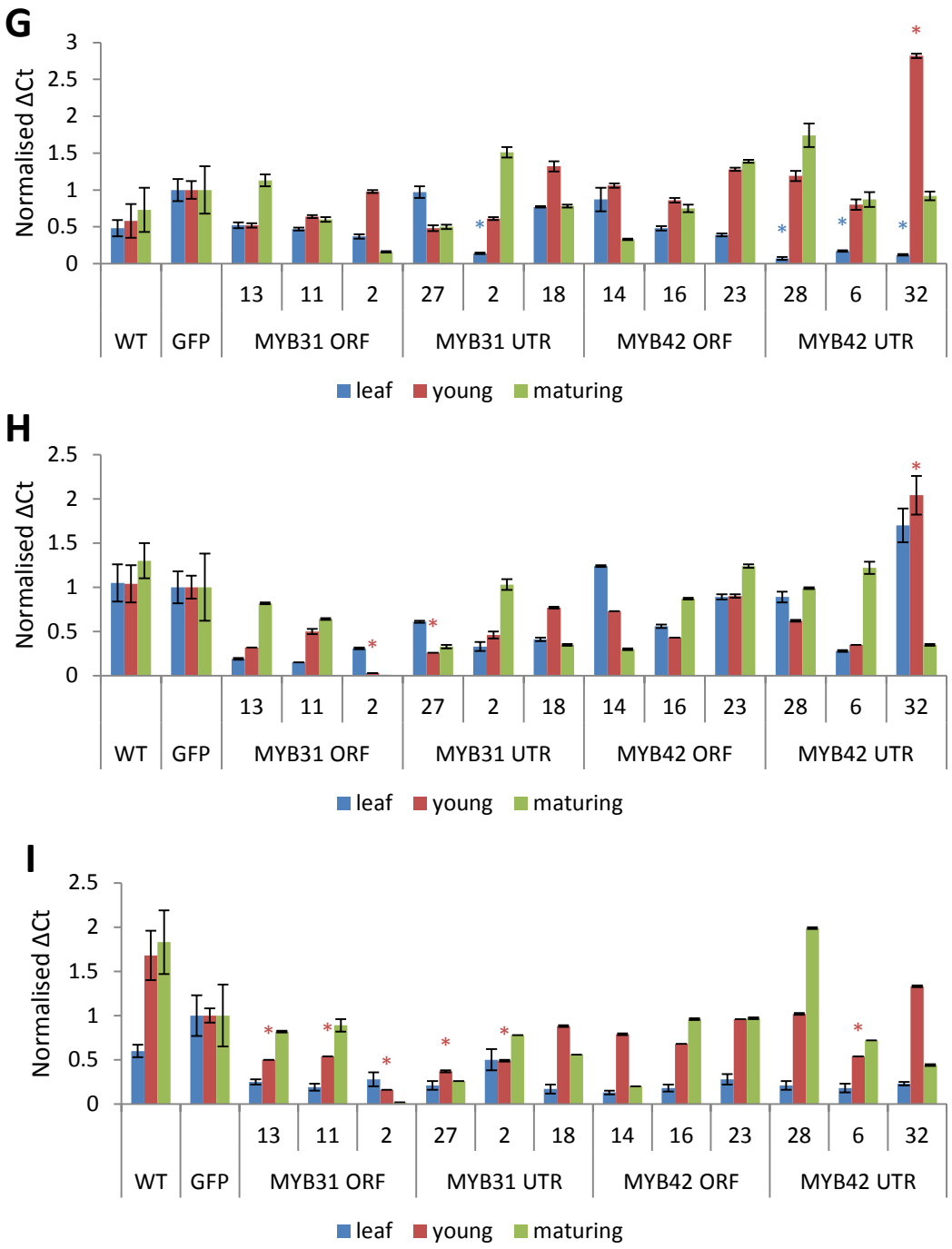


FIGURE 5.2 (continued): ΔCt expression of lignin biosynthesis genes (showing standard error of the mean) in sugarcane expressing *Zm*MYB31 or *Zm*MYB42 normalised against GFP controls after qPCR analysis for CAD (G); CCR (H); and COMT (I). Leaf (blue), young (red) and maturing (green) refers to gene expression levels in leaf tissue, young internode tissue and maturing internode tissue respectively. MYB plants statistically different to GFP controls after a two-tailed t-test assuming unequal variance, p = 0.05, are shown by a coloured asterisk respective to tissue type. WT n = 9; GFP n = 3; n = individually analysed plants.

The lignin gene expression results for MYB31 aligned with the published results more closely than the MYB42 expression results (Table 5.5).

5.2.4 Cell wall compositional analysis of MYB transgenic sugarcane

Cell wall compositional analysis was performed on the same plants that underwent stem qPCR analysis (seven MYB plants per line and nine wild type and three GFP controls, Table 5.4). This involved determination of acid soluble and acid insoluble lignin content as well as cellulose (glucose) and hemicellulose (xylose, galactose and arabinose) content. The percentage of lignin, cellulose and hemicellulose in the cell wall for MYB31 and MYB42 plants are seen in Table 5.6 and Table 5.7, respectively. For each MYB line the plants are listed in ascending order based on total lignin content. For all components there were no significant differences seen between wild type and GFP controls. After analysis all mass balances were within 10% of starting material.

5.2.4.1 Cell wall composition of MYB31 expressing sugarcane

Of the MYB31 expressing sugarcane only MYB31 UTR 27 showed a significant decrease in total lignin content (Table 5.6). This same plant also had significantly decreased acid soluble lignin whilst having significantly increased glucose and xylose. Two other MYB31 UTR plants showed a significant decrease in acid insoluble lignin and a significant increase in xylose (Table 5.6). Of the MYB31 ORF plants, none showed significant decreases in total lignin or acid insoluble lignin, with only MYB31 ORF 2 showing a decrease in acid soluble lignin. Approximately half the MYB31 plants (ORF and UTR) had significantly increased xylose and galactose (Table 5.6). There were no plants with a significant decrease in any structural carbohydrates. Overall, MYB31 expression appeared to have little impact on lignin and glucose levels but increased the synthesis of structural hemicelluloses.

5.2.4.2 Cell wall composition of MYB42 expressing sugarcane

Three MYB42 ORF plants analysed (14, 16 and 23) showed a significant decrease in total lignin content which was the result of significant decreases in acid insoluble lignin, as there were no changes to acid soluble lignin in these three plants (Table 5.7). MYB42 ORF 21, 16 and 11 were the only plants to show significant changes to acid soluble lignin, glucose and xylose respectively (Table 5.7). Three MYB42 UTR plants analysed (28, 6 and 32) also

TABLE 5.6: MYB31 sugarcane cell wall composition. The percentage of each component of the total composition is shown with the standard error of the mean. Samples significantly different to the GFP controls after a two-tailed t-test assuming unequal variance, $p = 0.05$, are shown in bold. WT n = 9; GFP n = 3; n = individually analysed plants. Plants are listed in ascending total lignin content for each line.

		Total lignin		Acid insoluble lignin		Acid soluble lignin		Glucose		Xylose		Galactose		Arabinose	
		%	+/-	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-
Wild type	n=9	23.61	0.24	18.18	0.26	5.44	0.07	46.95	0.34	20.30	0.28	0.33	0.16	1.94	0.11
GFP	n=3	23.37	0.39	17.92	0.38	5.45	0.10	47.17	0.45	20.23	0.28	0.10	0.08	1.69	0.22
MYB31 ORF	13	21.74	0.15	16.49	0.09	5.25	0.18	48.77	0.33	20.27	0.17	0.00	0.00	1.49	0.06
	11	22.23	0.15	16.60	0.17	5.64	0.11	45.45	0.41	**22.05**	**0.15**	**1.24**	**0.02**	**2.95**	**0.03**
	2	22.50	0.04	17.61	0.09	**4.89**	**0.07**	47.79	0.30	20.83	0.13	**1.05**	**0.02**	2.21	0.04
	7	23.07	0.30	17.61	0.22	5.45	0.12	46.74	0.61	**22.11**	**0.17**	0.00	0.00	2.24	0.06
	1	23.08	0.29	17.86	0.24	5.22	0.14	45.08	0.18	**23.14**	**0.18**	**0.98**	**0.01**	2.24	0.03
	8	23.88	0.40	18.47	0.40	5.41	0.04	46.01	0.82	21.34	0.37	**1.24**	**0.05**	2.71	0.07
	9	24.97	0.47	19.72	0.40	5.25	0.11	46.88	0.56	20.88	0.39	**1.07**	**0.05**	2.28	0.03
MYB31 UTR	27	**21.03**	**0.19**	16.29	0.14	**4.74**	**0.05**	**51.76**	**0.51**	**24.27**	**0.23**	0.00	0.00	2.19	0.05
	2	21.18	0.18	**15.63**	**0.12**	5.55	0.12	45.63	0.41	**22.03**	**0.32**	**1.33**	**0.02**	**3.02**	**0.05**
	18	21.48	0.19	**15.57**	**0.18**	5.91	0.08	47.40	0.36	**22.11**	**0.26**	**1.28**	**0.00**	2.71	0.06
	11	22.25	0.07	17.10	0.09	5.16	0.14	46.05	0.14	**22.44**	**0.08**	**1.06**	**0.05**	2.58	0.03
	12	22.40	0.16	17.01	0.04	5.39	0.15	47.12	0.31	20.56	0.14	0.00	0.00	2.06	0.02
	7	22.79	0.11	17.13	0.06	5.66	0.15	45.38	0.17	20.15	0.15	0.32	0.26	1.92	0.05
	20	22.86	0.07	17.47	0.05	5.40	0.12	47.60	0.15	21.38	0.13	0.00	0.00	1.84	0.08

TABLE 5.7: MYB42 sugarcane cell wall composition. The percentage of each component of the total composition is shown with the standard error of the mean. Samples significantly different to the GFP controls after a two-tailed t-test assuming unequal variance, $p = 0.05$, are shown in bold. WT n = 9; GFP n = 3; n = individually analysed plants. Plants are listed in ascending total lignin content for each line.

		Total lignin		Acid insoluble lignin		Acid soluble lignin		Glucose		Xylose		Galactose		Arabinose	
		%	+/-	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-
Wild type	n=9	23.61	0.24	18.18	0.26	5.44	0.07	46.95	0.34	20.30	0.28	0.33	0.16	1.94	0.11
GFP	n=3	23.37	0.39	17.92	0.38	5.45	0.10	47.17	0.45	20.23	0.28	0.10	0.08	1.69	0.22
MYB42 ORF	14	**18.51**	**0.13**	**12.84**	**0.07**	5.67	0.20	49.24	0.14	21.28	0.22	**1.08**	**0.02**	2.63	0.04
	16	**20.86**	**0.05**	**15.25**	**0.01**	5.61	0.04	**49.62**	**0.38**	20.81	0.08	0.32	0.26	2.14	0.03
	23	**21.58**	**0.22**	**15.85**	**0.25**	5.73	0.06	47.86	0.18	19.60	0.34	0.30	0.24	2.17	0.05
	11	22.05	0.09	16.94	0.12	5.11	0.07	47.92	0.38	**22.67**	**0.24**	**0.98**	**0.02**	2.08	0.02
	18	22.38	0.10	16.76	0.09	5.62	0.01	48.69	0.55	19.90	0.09	0.00	0.00	1.96	0.02
	21	22.47	0.14	17.74	0.07	**4.73**	**0.15**	48.85	1.12	21.82	0.38	0.00	0.00	1.21	0.08
	26	22.99	0.09	17.68	0.01	5.30	0.10	46.61	0.33	21.61	0.22	**1.10**	**0.06**	2.50	0.06
MYB42 UTR	28	**19.42**	**0.29**	**13.78**	**0.34**	5.64	0.05	**50.55**	**0.13**	21.78	0.48	0.63	0.26	2.60	0.04
	6	**20.21**	**0.03**	**14.40**	**0.06**	5.81	0.08	46.41	0.16	**23.28**	**0.14**	**1.15**	**0.04**	**3.03**	**0.01**
	32	**20.91**	**0.24**	**15.46**	**0.14**	5.45	0.14	47.83	0.29	21.53	0.11	0.28	0.23	1.96	0.05
	30	21.51	0.30	**15.29**	**0.19**	6.21	0.20	49.12	0.28	**23.94**	**0.23**	**1.21**	**0.02**	2.63	0.03
	15	22.57	0.10	16.73	0.08	5.84	0.06	47.10	0.31	21.59	0.22	0.00	0.00	1.72	0.07
	26	23.48	0.11	18.36	0.07	5.12	0.08	47.41	0.45	**23.58**	**0.28**	0.00	0.00	1.67	0.03
	16	24.15	0.21	18.98	0.14	5.17	0.08	47.06	0.31	20.89	0.18	0.00	0.00	1.41	0.04

showed a significant decrease in total lignin content as a result of significant decreases in acid insoluble lignin, as there were no changes to acid soluble lignin in these three plants (Table 5.7). Additionally, MYB42 UTR 30 also showed a decrease to acid insoluble lignin, but not to total lignin (Table 5.7). MYB42 UTR 28 was the only plant with increased glucose content, with three additional plants having increased xylose content (Table 5.7). There were MYB42 ORF and UTR plants with increases to galactose and arabinose content, though overall, the changes to structural carbohydrates were not specific to the plants with changes to lignin content.

5.2.5 Enzymatic hydrolysis of MYB and control bagasse samples

The three MYB plants with the largest decrease in lignin composition were selected to undergo enzymatic hydrolysis. Enzymatic hydrolysis was performed on three plants per MYB line, four wild type and three GFP transgenic control plants (Table 5.4). Hydrolysis was carried out for 72 hours with sampling at six different time points. A random selection of both MYB and control plants had non-pretreated bagasse hydrolysed alongside pretreated bagasse with both samples originating from the same stock. In all cases, pretreated samples showed a significant improvement in glucose released over untreated samples (Figure 5.3) which supported the effectiveness of the pretreatment method used. All figures and tables list MYB plants in ascending order of original total lignin content for each line. There were no significant differences in glucose release between the wild type and GFP control plants across the six time points.

When compared to GFP controls, only two of the six MYB31 plants showed a significant increase in glucose release after 72 hours whereas all six MYB42 expressing plants released significantly more glucose than the GFP controls (Figure 5.4). It also appeared that the MYB42 plants outperformed the MYB31 plants as all the MYB42 plants released more glucose than four of the six MYB31 plants, and two MYB42 plants released significantly more glucose than all the MYB31 plants (Figure 5.4). No statistical differences between wild type and GFP controls were seen.

The final amount of glucose released by each plant was also plotted against the compositional data from the pretreated and non-pretreated plants used for enzymatic hydrolysis. Strong correlations were not observed between glucose released after 72 hours of enzymatic hydrolysis and the amount of residual components (total lignin, acid insoluble

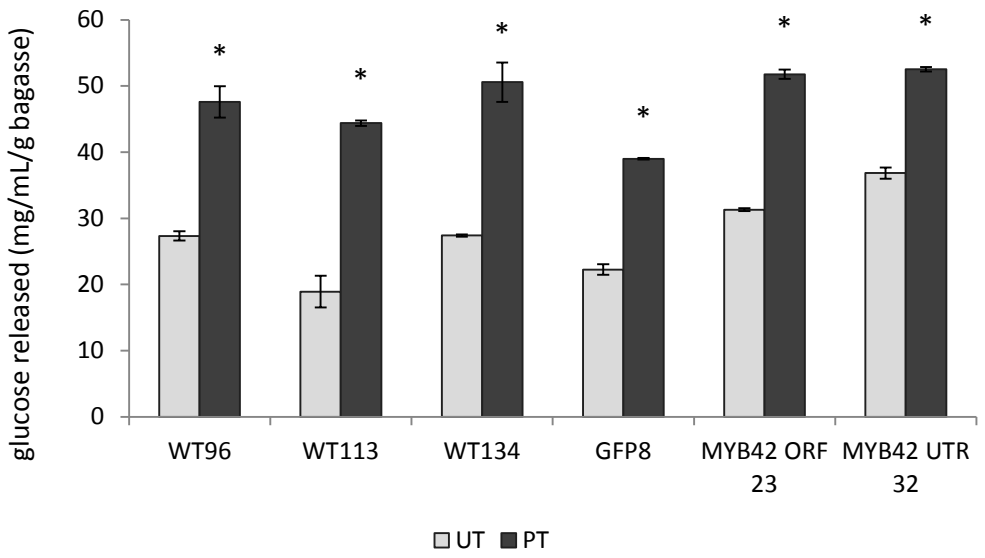


FIGURE 5.3: Total glucose concentration (mg/mL) in enzymatic hydrolysis solution per gram of untreated (UT) and pretreated (PT) bagasse (g bagasse) after 72 hours incubation showing standard error of the mean. The UT and PT results were compared for each plant individually using a two-tailed t-test assuming unequal variance, p = 0.05 with an asterisk indicating a statistical difference.

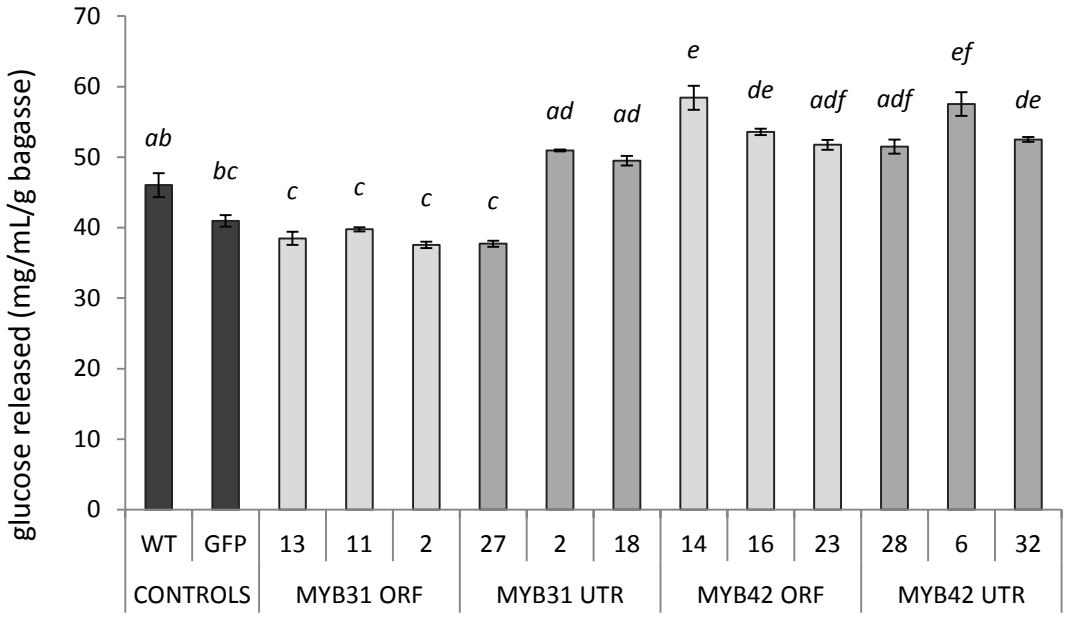


FIGURE 5.4: Total glucose concentration (mg/mL) in enzymatic hydrolysis solution per gram of bagasse (g bagasse) after 72 hours incubation showing standard error of the mean. Different letters above bars (*a-f*) indicate significant differences after a one-way ANOVA with Tukey post-hoc analysis, p = 0.05. WT n = 3; GFP n = 3, n = individually analysed plants.

lignin, acid soluble lignin, total carbohydrates, glucose or xylose) remaining after the pretreatment of the bagasse samples (data not shown). When the glucose released after 72 hours of enzymatic hydrolysis was plotted against the cell wall components for untreated bagasse samples negative correlations were observed for total lignin (R^2 = -0.675) and acid insoluble lignin (R^2 = -0.718) (Figure 5.5), but strong correlations were not observed for acid soluble lignin, total carbohydrates, glucose or xylose (data not shown).

As previously described, the glucose concentration was measured at six time points over the 72 hour incubation period (Table 5.8). There was a trend for the plants that had significantly higher release of glucose at the 72 hour time point also having significantly higher glucose released at earlier time points. Significantly higher glucose released at earlier time points by the MYB sugarcane is indicative of these plants showing an increased rate of cellulose conversion to glucose (Figure 5.6 and Figure 5.7). Four MYB42 plants showed significantly higher rates of glucose conversion as early as the six hour time point and continued to release significantly more glucose at each of the later time points (Table 5.8). The plants with significantly more glucose released after 72 hours all show signs of increased rate of glucose conversion by the 12 hour time point (Table 5.8).

5.2.6 Cellulose crystallinity index of MYB bagasse

The tissue requirements of pretreatment for enzymatic hydrolysis left enough bagasse only for six control plants and five MYB plants to undergo determination of cellulose crystallinity index (Appendix H). Statistical analysis was not performed as the limited number of samples would not provide reliable results. The ranges of crystallinity were 46.96% - 51.81% for control plants and 45.56% - 47.50% for MYB expressing plants (Appendix H).

5.2.7 Sucrose content of juice extracted from MYB plants and controls

Plants that underwent enzymatic hydrolysis were also assessed for sucrose content of extracted juice (Table 5.4) to determine if changes in lignin content or structure had affected juice composition and quantity as carbon partitioned for cell wall synthesis may affect the flux of carbon directed to sucrose formation (Papini-Terzi *et al.*, 2009). This is a key determinant of plant quality in the sugarcane industry. Overall only MYB31 UTR 2 showed a significant decrease in sucrose content when compared against the GFP control plants (Figure 5.8). All other plants were consistent with controls (Figure 5.8). Glucose and

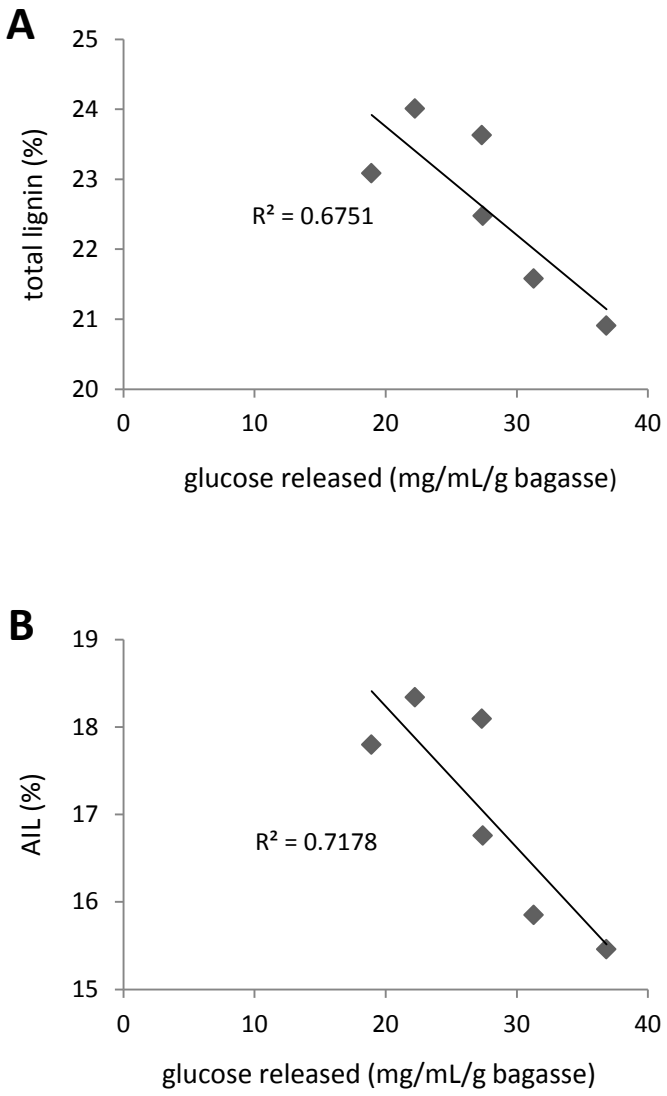


FIGURE 5.5: R^2 values between total lignin (A); and acid insoluble lignin (B) of non-treated bagasse and glucose released into enzymatic hydrolysis solution (mg/mL) per gram (g) bagasse after 72 hours of enzymatic hydrolysis.

TABLE 5.8: Glucose released into enzymatic hydrolysis solution (mg/mL) per gram (g) of bagasse measured at six time points over 72 hours. The glucose released is shown with standard error of the mean. Samples significantly different to the GFP control group after a two-tailed t-test assuming unequal variance, $p = 0.05$, are shown in bold. WT n = 3; GFP n = 3, n = individually analysed plants.

		0 hrs		6 hrs		12 hrs		24 hrs		48 hrs		72 hrs	
		mg/mL	+/-	mg/mL	+/-	mg/mL	+/-	mg/mL	+/-	mg/mL	+/-	mg/mL	+/-
CONTROLS	WT	0.92	0.05	26.17	0.40	32.89	0.86	39.77	0.55	43.55	0.76	46.02	1.70
	GFP	1.02	0.09	24.61	1.31	29.79	0.69	36.18	1.15	40.65	0.85	40.97	0.83
MYB31 ORF	13	1.07	0.07	25.72	0.20	31.38	0.12	34.94	0.65	38.06	0.72	38.47	0.93
	11	1.09	0.00	26.39	0.22	31.51	0.21	37.63	1.04	36.99	0.31	39.76	0.29
	2	0.97	0.07	23.85	0.43	27.83	0.95	**30.80**	**0.70**	37.91	1.77	37.56	0.45
MYB31 UTR	27	0.87	0.07	23.03	0.13	28.44	0.70	32.83	0.38	**35.38**	**0.67**	37.71	0.42
	2	1.07	0.11	29.84	0.26	**38.52**	**0.62**	**43.01**	**0.85**	**48.86**	**0.82**	**50.96**	**0.13**
	18	1.22	0.08	29.23	0.10	**35.86**	**0.42**	**42.57**	**0.16**	**46.90**	**0.77**	**49.48**	**0.69**
MYB42 ORF	14	**1.57**	**0.06**	**34.29**	**0.15**	**41.06**	**0.41**	**47.41**	**0.35**	**53.89**	**1.49**	**58.43**	**1.71**
	16	1.38	0.02	**34.80**	**0.10**	**43.31**	**0.36**	**47.79**	**0.22**	**51.92**	**0.89**	**53.58**	**0.46**
	23	1.42	0.02	**33.18**	**0.69**	**43.03**	**1.23**	**44.61**	**0.59**	52.14	3.23	**51.76**	**0.70**
MYB42 UTR	28	0.77	0.05	31.44	0.19	**39.43**	**0.62**	**44.71**	**0.51**	**48.65**	**0.28**	**51.49**	**1.00**
	6	1.15	0.02	**33.01**	**0.55**	**42.89**	**0.45**	**49.95**	**1.70**	**54.11**	**0.90**	**57.53**	**1.70**
	32	0.79	0.09	27.49	0.09	**36.54**	**0.13**	42.03	0.28	**48.03**	**0.34**	**52.51**	**0.33**

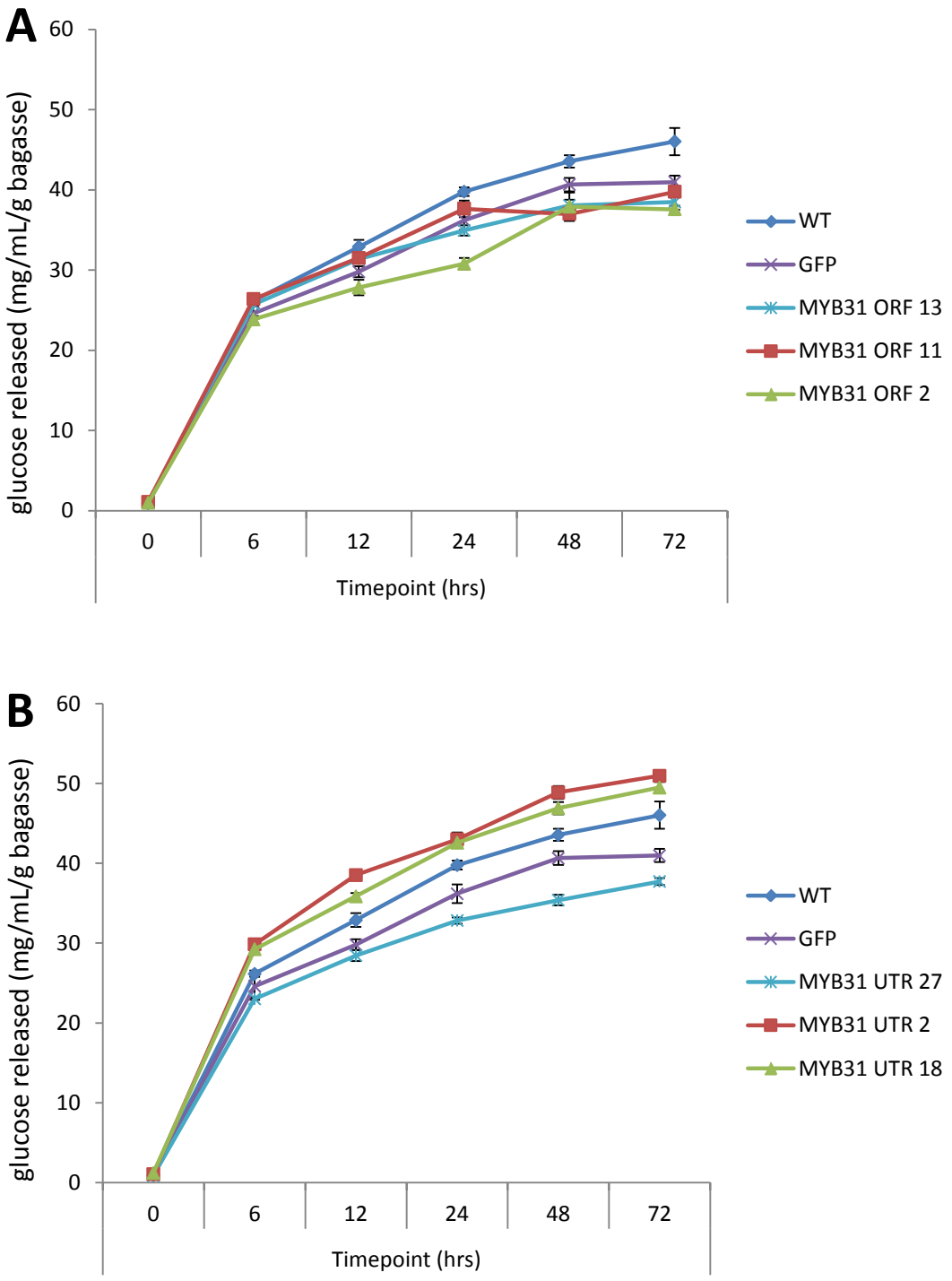


FIGURE 5.6: Total glucose concentration in enzymatic hydrolysis solution (mg/mL) per gram (g) of bagasse showing standard error of the mean measured at six time points over a period of 72 hours incubation for MYB31 ORF plants (A); and MYB31 UTR plants (B) using values from Table 5.8. WT n = 3; GFP n = 3, n = individually analysed plants.

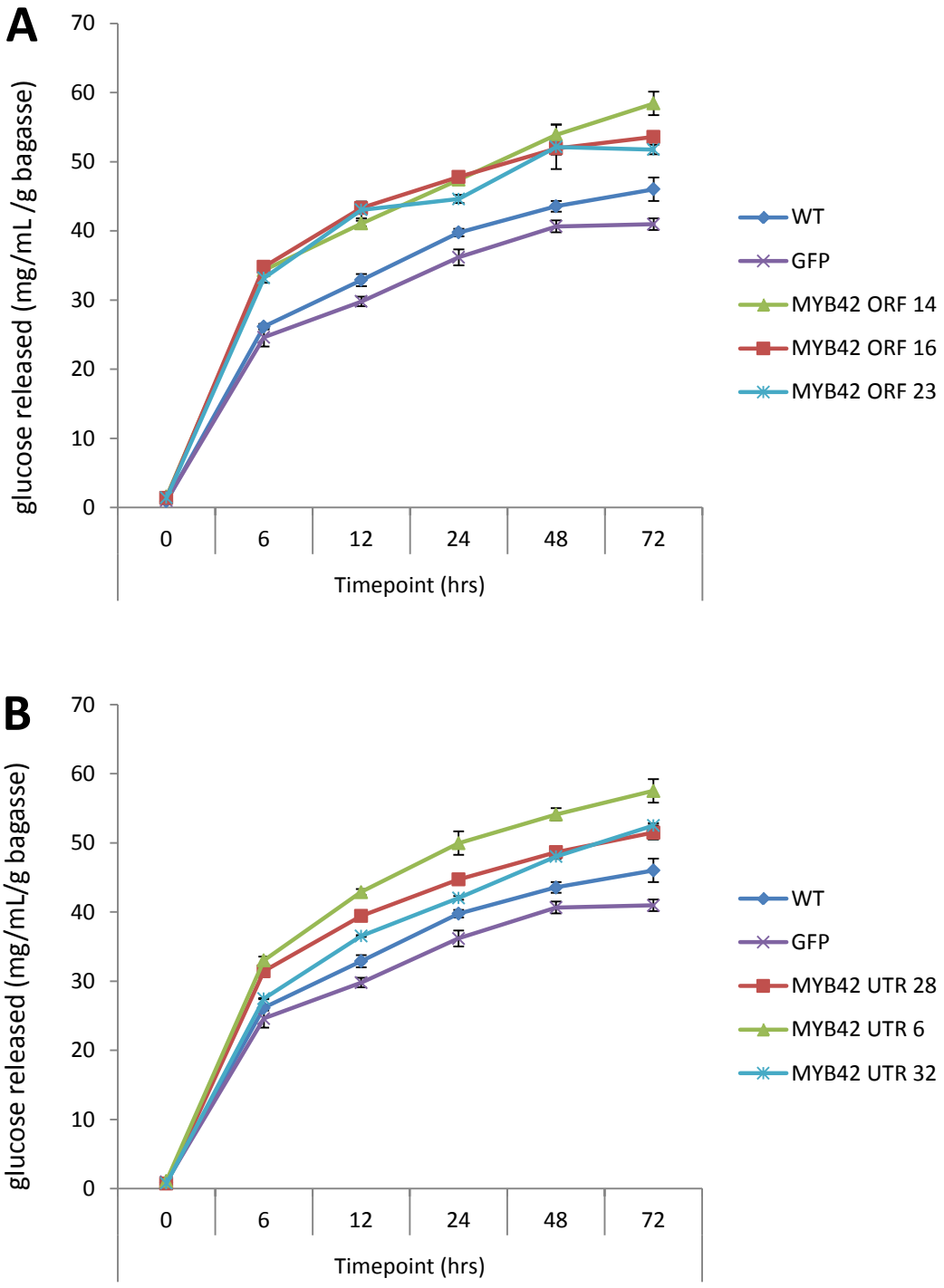


FIGURE 5.7: Total glucose concentration in enzymatic hydrolysis solution (mg/mL) per gram (g) of bagasse showing standard error of the mean measured at six time points over a period of 72 hours incubation for MYB42 ORF plants (A); and MYB42 UTR plants (B) using values from Table 5.8. WT n = 3; GFP n = 3, n = individually analysed plants.

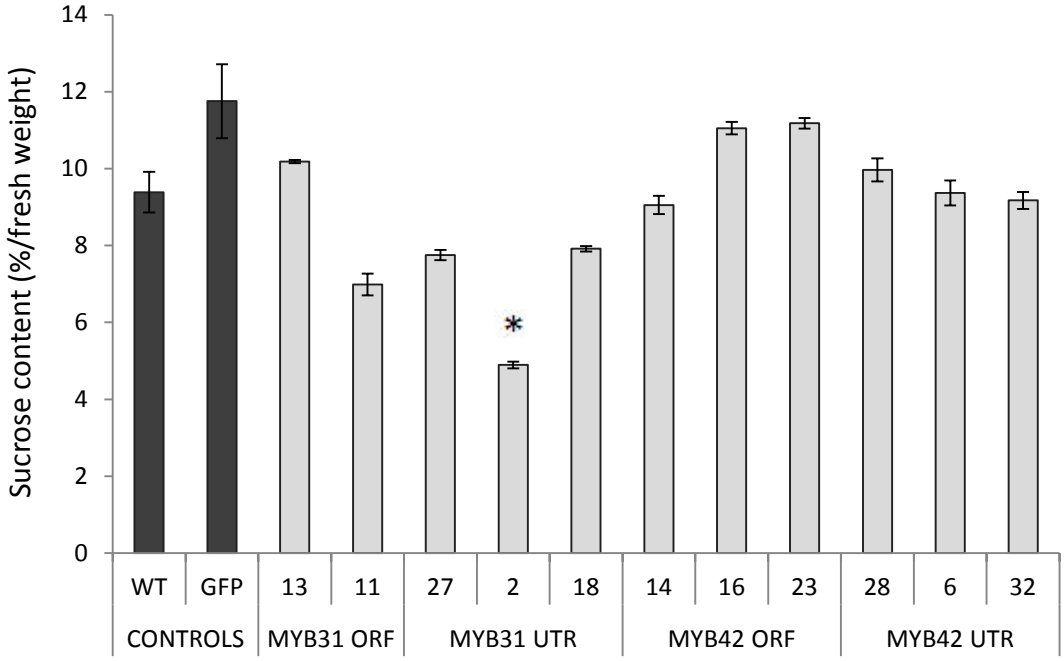


FIGURE 5.8: Sucrose content (%/fresh weight) of extracted juice of MYB sugarcane plants selected for enzymatic hydrolysis (showing standard error of the mean). An asterisk indicates a significant difference to GFP controls after a two-tailed t-test assuming unequal variance, $p = 0.05$. WT n = 4; GFP n = 3, n = individually analysed plants.

fructose were assessed, but not detected using this method. MYB31 ORF 2 did not have sufficient tissue for juice analysis.

5.2.8 Phenotypic measurements of MYB plants

Phenotypic measurements were taken at the time of harvest. Plant height, number of internodes, internode diameter and average internode length were recorded and calculated (Appendix I). A *z* score was calculated based on the average results of the GFP transgenic control plants and any MYB plant with a *z* score greater than 2 or -2, indicating the result was more than two standard deviations from the control group result, were considered different to controls. Overall, there were very few MYB31 or MYB42 plants with different heights or total internode numbers. MYB31 ORF 7 was taller than controls and MYB42 UTR 30 was shorter, and MYB31 ORF 2 was the only plant with a different number of internodes (Appendix I). Approximately half of MYB31 ORF and UTR plants had smaller average internode diameters, and the majority of MYB31 plants had greater average length

of internodes (Appendix I). Of the MYB42 plants, two had increased internode diameters and four had decreased diameters. Similar results were seen for internode length as four MYB42 plants had increased internode length and two had decreased length (Appendix I). No differences were seen between wild type and GFP controls for height, total number of internodes or average internode diameter, however, wild type plants did have longer average internode length than GFP controls (Appendix I).

5.3 DISCUSSION

*Zm*MYB31 and *Zm*MYB42 were successfully cloned from germinated maize embryos and inserted in the genome of regenerated sugarcane plants (Table 5.4). There was 100% alignment of the amino acid sequences of the R2 and R3 domains, necessary for functionality, between the cloned MYB31 and MYB42 genes (Appendix F) and the published sequences (Fornalé *et al.*, 2006). The R2R3 domains of MYB transcription factors bind to conserved AC elements within the promoters of lignin biosynthetic genes, allowing for regulation of gene expression levels (Fornalé *et al.*, 2010; Lauvergeat *et al.*, 2002; Patzlaff *et al.*, 2003a; Patzlaff *et al.*, 2003b). There were some differences in amino acid sequence within the C-terminal end of the translated MYB31 and MYB42 proteins and also in the nucleotide sequences of the 5' and 3' UTR sequences between cloned and published MYB31 sequences (Appendix F).

These differences were considered to be due to varietal differences in corn used for cloning purposes and did not appear to negatively affect the expression of the MYB genes (Figure 5.1) nor in turn their overall ability to down-regulate the expression of genes within the lignin biosynthesis pathway (Table 5.5; Figure 5.2; Appendix G). As the C-terminal region is responsible for the regulatory function of the translated protein (Bedon *et al.*, 2007; Jiang *et al.*, 2004; Legay *et al.*, 2007; Tamagnone *et al.*, 1998), it is possible that the amino acid differences may have had a more subtle influence on the functionality of the translated MYB proteins. This may have contributed to the differences seen in lignin biosynthesis gene down-regulation in this study and published studies (Table 5.5; Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009). Furthermore, very little transgene silencing was observed for MYB31 or MYB42 (Figure 5.1), which is a known occurrence in transgenic sugarcane

(Mudge *et al.*, 2009). This may be in part due to the functionality of the maize Ubiquitin promoter in transgenic sugarcane (Mudge *et al.*, 2009; Osabe *et al.*, 2009).

As described, the ORFs of MYB31 and MYB42 were cloned with and without adjacent sections of the 5' and 3' UTR. Evidence exists that the inclusion of UTR sequences can improve gene expression (Hughes, 2006; Mignone *et al.*, 2002; Pesole *et al.*, 2001; Sharma and Sharma, 2009; Wilkie *et al.*, 2003), and it was found that the constructs containing UTR sequences for both MYB31 and MYB42 had more PCR and qPCR positive plants than their ORF counterpart (Table 5.4); however, these differences in plant numbers were only slight. Retaining or omitting the UTR sequences did not appear to affect the level of down-regulation of lignin biosynthesis genes (Appendix G), deposition of cell wall components (Table 5.6 and Table 5.7), juice sucrose levels (Figure 5.8), or plant phenotype (Appendix I) in plants expressing MYB31 or MYB42. As the combination of *Zm*Ubi-iUbi promoter sequences results in high transgene expression in monocot species (Christensen and Quail, 1996), it is possible that any impact the UTR sequences may have had was masked. The utilisation of alternative promoter sequences with varying strength for transgene expression may highlight subtle differences between constructs that have the UTR sequences retained and omitted.

All MYB42 ORF and MYB42 UTR plants analysed released significantly more glucose, indicating the UTR sequences did not contribute any additional benefits to the enzymatic hydrolysis of bagasse (Table 5.8). Including the UTR sequences did result in an increase in MYB31 expressing plants which released significantly more glucose after enzymatic hydrolysis. Of the three MYB31 ORF plants analysed, none released significantly more glucose whereas two of the three MYB31 UTR plants analysed did (Table 5.8). The increase in glucose released by MYB31 UTR plants over MYB31 ORF plants suggests that it may be beneficial to include these sequences when transforming sugarcane for second generation bioethanol production. If this avenue of research were pursued, it would also be optimal to clone the UTR sequences in their entirety to confirm if their addition has any obvious benefits. It is also possible that the UTR sequences affected the translation levels of the MYB genes, which could be determined if additional research also included measuring protein content.

Both MYB31 and MYB42 have previously been found to down-regulate multiple genes within the lignin biosynthesis pathway (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et*

al., 2009), and this down-regulatory control was also observed in this study. When compared with published data, the current results for MYB31 match more closely than for MYB42 (Table 5.5; Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009). MYB31 appears to down-regulate all genes analysed except CAD whereas MYB42's regulatory effects are more subtle and while approximately half of the genes analysed appear to be down-regulated, interestingly, half appear to be up-regulated (Table 5.5). The genes down-regulated by MYB42 generally appear earlier in the biosynthesis pathway and the genes up-regulated function later in the pathway. Previous research has found different MYB genes target different genes within the lignin biosynthesis pathway with a general consensus that at least one gene early in the biosynthesis pathway is targeted (Table 2.2; Fornalé *et al.*, 2006; Jin *et al.*, 2000; Omer *et al.*, 2013; Shen *et al.*, 2012; Tamagnone *et al.*, 1998; Zhu *et al.*, 2013). The current results for MYB31 and MYB42 fit within this general pattern. The down-regulatory effects of MYB31 also appear more consistent across the three tissue types analysed than MYB42 (Table 5.5). These results suggested that MYB31 would reduce lignin content in sugarcane to a greater extent than MYB42. There was no correlation between the expression levels of MYB31 (Figure 5.1) and MYB42 (Figure 5.1) with the degree of down-regulation (Appendix G) as higher levels of MYB expression did not necessarily lead to greater reductions in lignin biosynthesis gene expression levels.

Once it was determined that both MYB genes were exerting down-regulatory control over the lignin biosynthesis gene pathway the next step was to determine if this translated into a reduction of lignin in the composition of the secondary cell wall. Previous research shows a strong correlation between reduction of lignin gene expression by MYB genes and reduction in lignin content (Table 2.2; Fornalé *et al.*, 2010; Omer *et al.*, 2013; Shen *et al.*, 2012; Sonbol *et al.*, 2009; Zhu *et al.*, 2013).

Of the 14 MYB31 expressing sugarcane analysed (Table 5.6) only MYB31 UTR 27 had a significant decrease in total lignin by 10%. Within this line, MYB31 UTR 2 and MYB31 UTR 18 had 13% reductions in acid insoluble lignin but not total lignin. This is far from the 8-fold (Fornalé *et al.*, 2006) and 70% (Fornalé *et al.*, 2010) decreases in lignin content previously reported in Arabidopsis after the transgenic expression of MYB31. This difference may be due the expression of this transcription factor in different plant species using different promoters and terminators. Alternatively, as discussed in Section 4.3.3, sugarcane has a large and complex polyploid genome (Grivet and Arruda, 2001) which increases the likelyhood of gene homologues, including of lignin biosynthesis pathway genes. If by chance

MYB31 was only able to down-regulate a number of these potential homologues, the decreased expression may have been compensated for by unregulated homologues, therefore reducing the overall effect of MYB31 expression on lignin biosynthesis. Although MYB31 plants showed down-regulated expression of many lignin biosynthesis genes (Figure 5.2, Appendix G), it was not known if qPCR primers were amplifying all potential lignin gene homologues (Section 4.3.3), therefore a possibility exists that the compensatory expression of unregulated homologues went undetected. Approximately half of the MYB31 expressing sugarcane plants (ORF and UTR) had significantly increased xylose and galactose (Table 5.6) whereas Fornalé *et al.* (2010) reported no changes in the structural carbohydrate composition of lignin reduced Arabidopsis. Within sugarcane, MYB31 expression appears to have little impact on lignin and glucose levels but increases the synthesis of hemicelluloses.

Six sugarcane plants expressing MYB42 (three ORF plants and three UTR plants) showed a significant decrease in total lignin content between 8% – 21% (Table 5.7) whereas previous research found Arabidopsis plants expressing MYB42 had greater reductions in lignin content of 2.1 fold (Fornalé *et al.*, 2006) and by 60% (Sonbol *et al.*, 2009). Two sugarcane plants had significant increases in glucose content whereas no changes in cellulose content were found in MYB42 expressing Arabidopsis (Sonbol *et al.*, 2009). Hemicelluloses were increased in Arabidopsis (Sonbol *et al.*, 2009) as they were in several MYB42 expressing sugarcane plants (Table 5.7). These changes in carbohydrate content did not correlate with changes to lignin content within the MYB42 sugarcane (Table 5.7).

The cell wall compositional analysis was unexpected after the expression data of the lignin biosynthesis pathway was examined. This data strongly suggested that the MYB31 expressing lines would show greater reductions in lignin content than the MYB42 lines based on the number of lignin biosynthesis pathway genes down-regulated (Table 5.5; Appendix G). This hypothesis is also supported by previous research which found MYB31 expressing Arabidopsis had less lignin than their MYB42 counterparts (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006). The current results do not support this hypothesis as only one MYB31 expressing plant had a reduction in lignin of 10% whereas six MYB42 expressing plants had lignin reductions of 8%-21%. These results suggest that within sugarcane, MYB42 affects the deposition rates of lignin whereas MYB31 increases the hemicellulose content within secondary cell walls but has little effect on lignin production.

The findings also suggest that transcription levels may not necessarily reflect translation levels of lignin biosynthesis genes. Additionally, the rates of protein turnover by the translated lignin biosynthesis enzymes in the MYB31 and MYB42 expressing plants are unknown. It is possible that the genes with higher expression levels in young internode tissue (Table 5.5, Appendix G) have gene products that may persist in maturing tissue, allowing for continued lignin biosynthesis without the need for continued high expression of the lignin biosynthesis gene. Although the genes of the lignin biosynthesis pathway have been well characterised (Section 1.3), a full understanding of metabolic flux through the pathway remains to be established, making it difficult to predict the outcomes of modifying the expression levels of lignin biosynthesis pathway genes. Assaying the enzyme activity of translated lignin biosynthesis genes, or the concentration of enzyme product throughout the stem, would provide a clearer picture of the down-regulatory effects of MYB expression over these genes, but was beyond the scope of the current study.

Reduced lignin content in sugarcane has previously led to improved saccharification (Benjamin *et al.*, 2013; Benjamin *et al.*, 2014; Jung *et al.*, 2012; Jung *et al.*, 2013; Masarin *et al.*, 2011) as has the expression of MYB31 and MYB42 in Arabidopsis (Fornalé *et al.*, 2010; Sonbol *et al.*, 2009). The expression of *Pv*MYB4 in transgenic switchgrass is the only report of a monocot species having improved saccharification resulting from the expression of a MYB gene (Shen *et al.*, 2012; Shen *et al.*, 2013). This is the first report of MYB31 and MYB42 being expressed in a monocot species, and the first report of MYB transcription factors being expressed in sugarcane to alter lignin biosynthesis for improved biofuel production. In this study the three plants with the least amount of total lignin from each line were assessed for enzymatic hydrolysis performance against controls over a period of 72 hours.

When the final glucose released after 72 hours of enzymatic hydrolysis is compared (Figure 5.4) it is clear that the MYB42 lines performed better than the MYB31 lines and the controls. All six MYB42 plants hydrolysed released significantly more glucose than the transgenic controls after 72 hours (Figure 5.4). These plants also had a faster rate of glucose conversion than controls (Figure 5.7). The composition analysis found all six of these plants had significant reductions in total lignin as a result of reductions in acid insoluble lignin (Table 5.7). In contrast to these results, no MYB31 ORF plants and two MYB31 UTR plants had higher levels of glucose released than controls.

It has been observed that GFP expression can influence the expression of endogenous genes in transgenic plants (*pers. comm*. M. Bateson) and therefore the genes involved in the synthesis and deposition of cell wall components could potentially have been affected in the GFP controls plants. It was decided to compare MYB plants to transgenic control plants, as unlike wild type plants, they have undergone the same regeneration stresses and conditions as MYB transgenic plants. Therefore differences detected in MYB plants when compared with transgenic controls can be attributed to MYB expression and not the process of generating transgenic plants. Future research may include an additional control line expressing only the *Zm*Ubi-nptII-nos cassette (Section 5.1.2.1) to remove any influence the expression of GFP may have. If MYB plants were compared against wild type controls instead, then this would have yielded a different set of results. For example, if the glucose released after 72 hours of enzymatic hydrolysis by MYB and wild type plants are compared, the number of MYB plants that released significantly more glucose is reduced (Figure 5.4). All three MYB31 ORF plants and MYB31 UTR 27 released significantly less glucose than wild types, and MYB31 UTR 2 and 18 were comparable to wild types (they released higher levels of glucose than GFP controls) (Figure 5.4). MYB42 ORF 14 and 16, and MYB42 UTR 6 and 32 released significantly more glucose than wild type controls. While the total number of MYB plants that released significantly higher levels of glucose is reduced when comparing against wild type controls, the trend of MYB42 plants performing better than MYB31 plants is still clear (Figure 5.4).

The MYB31 ORF results are not surprising as no MYB31 ORF plants had reductions in lignin (Table 5.6). Interestingly MYB31 UTR27 had significantly less total lignin but did not release significantly more glucose whereas MYB31 UTR 2 and MYB31 UTR 18 did not have significantly less total lignin but did release significantly more glucose (Table 5.6, Figure 5.4), which is likely due to the significant reductions in acid insoluble lignin in these two plants (Table 5.6). The rate of glucose conversion in the MYB31 plants reflects the results at 72 hours, with only MYB31 UTR 2 and MYB31 UTR 18 outperforming the transgenic controls (Figure 5.6). Similar to the enzymatically hydrolysed MYB42 plants (Table 5.8), MYB31 UTR 2 and MYB31 UTR 18 did have significantly less acid insoluble lignin (Table 5.6). The finding that all plants (MYB31 and MYB42) with significant reductions in acid insoluble lignin content performed significantly better than transgenic controls after enzymatic hydrolysis suggests that it may be this cell wall component, and not total lignin or acid soluble lignin content that is an influencing factor on enzymatic hydrolysis performance. Decreases in acid insoluble lignin has also been reported alongside improvements in saccharification in

switchgrass (Xu *et al*., 2011), maize (Fornalé *et al.*, 2012) and Arabidopsis (Fornalé *et al.*, 2010; Sonbol *et al*., 2009).

Previous research with MYB31 and MYB42 found that the expression of MYB42 in Arabidopsis improved enzymatic hydrolysis by 68% (Sonbol *et al.*, 2009) compared with an increase of 14% after the expression of MYB31 (Fornalé *et al.*, 2010). The current results found that MYB42 lines outperformed MYB31 lines which support the previous findings. A direct comparison of current and published results is not possible as the published papers used a different enzyme cocktail mix for hydrolysis and did not report the FPU or protein levels as well as not reporting all the conditions used during hydrolysis.

Due to limited bagasse available for enzymatic hydrolysis analysis, the enzymatic hydrolysis section of this study utilised one pretreatment condition (1% (w/w) sulfuric acid) and subsequently a single digestive enzyme cocktail mix at one concentration (6 FPU Accellerase 1500 with the addition of β-glucosidase) as described in Section 3.7.3. Of the different pretreatments available, acid pretreatment was selected as this would not change the lignin content of bagasse as it disrupts the hemicellulose content of lignocellulosic biomass and exposes the cellulose to hydrolytic enzymes (Schell *et al*., 2003; Silverstein *et al*., 2007), thus better elucidating any digestive differences between bagasse samples with differences in lignin content. Accellerase 1500 was selected as it was available in the CTCB laboratories, and had previously been used to successfully hydrolyse bagasse (*pers. comm*. Z. Zhang). A low FPU concentration was used to again better elucidate any differences in bagasse digestibility. It would be expected that the use of different pretreatment conditions combined with varying cocktail mixes and concentrations would produce different results, which is an aspect future research may explore through the use of larger sample numbers, thus increasing the amount of available bagase. For example, if an alkaline pretreatment was used, which removes lignin as opposed to hemicellulose (Silverstein *et al*., 2007), it would be expected that transgenics and wild types would release similar amounts of glucose as the glucose contents between controls and transgenics do not overly differ (Table 5.6 and Table 5.7). Additionally, if an increased concentration of Accellerase 1500 was used, it would be expected that total glucose yields would be higher, but the rates of glucose conversion would still remain similar to those found in the current research (Figure 5.6 and Figure 5.7).

For transgenic sugarcane expressing MYB genes to be beneficial to and accepted by the global sugarcane industry it is important that any change to cell wall composition is not detrimental to the phenotype of the sugarcane nor to the juice sucrose levels which is the cornerstone of the sugarcane industry. Overall, the majority of phenotypic differences were in internode diameter and length for both MYB31 and MYB42 plants, with few differences seen in plant height or total number of internodes (Appendix I). The differences in phenotypic measurements were spread throughout the MYB plants which underwent compositional analysis (Table 5.6 and Table 5.7) and were no more or less prevalent in the plants with significantly reduced lignin contents or altered polysaccharide contents (Appendix I). Previous research has reported decreased growth rates and dwarfed phenotypes after MYB directed lignin reductions in poplar and Arabidopsis (Legay *et al.*, 2010), tobacco (Omer *et al.*, 2013; Shen *et al.*, 2012; Tamagnone *et al.*, 1998) and switchgrass (Shen *et al.*, 2012; Shen *et al.*, 2013). Reduced height was also observed in Arabidopsis expressing *Zm*MYB31 and *Zm*MYB42 (Fornalé *et al.*, 2010; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009). These findings are not supported overall by the current findings as no MYB plant with significant reductions in lignin (Table 5.6 and Table 5.7) showed a difference in plant height (Appendix I).

The MYB31 (Fornalé *et al.*, 2010) and MYB42 (Sonbol *et al.*, 2009) expressing Arabidopsis with reduced height had severe lignin reductions of up to 70% and 60% respectively. Vanholme *et al.* (2010) suggested that more modest lignin reductions may not result in these detrimental phenotypes which may help explain the current findings. Of the six MYB31 plants that underwent enzymatic hydrolysis (Table 5.8) only one plant had a decrease in total lignin and two plants in acid insoluble lignin of 10% and 13% respectively (Table 5.6). This is better highlighed by the MYB42 plants in this study. All the MYB42 plants analysed via enzymatic hydrolysis (Table 5.8) had significant reductions in lignin content (Table 5.7) but were not different in height to controls (Appendix I). As the greatest lignin reduction of the MYB42 plants was 21% in MYB42 ORF 14, these reductions in lignin content are modest when compared with the 60% - 70% previously reported (Fornalé *et al.*, 2010; Sonbol *et al.*, 2009) and may help explain why the MYB31 and MYB42 lignin reduced plants in this research did not show any height differences. Some of these lignin reduced plants did show differences in average internode diameter and length which may be attributed to the changes in cell wall composition.

5.3.1 Concluding remarks

Transcription factors are increasingly gaining research attention due to their wide-ranging regulatory effects and they are predicted to play a key role in the future of GM crops (Ambawat *et al.*, 2013; Century *et al.*, 2008). This prediction is supported by the positive finding that MYB31 and MYB42 expression can improve the saccharification potential of Arabidopsis, a model dicot plant species, and is further supported by current findings. The research presented in this thesis is the first known report of MYB31 and MYB42 being expressed in sugarcane, a commercially relevant monocot species important to the Australian economy. Previous research found MYB31 and MYB42 were able to reduce lignin deposition and subsequently increase saccharification in transgenic Arabidopsis through the down-regulation of genes within the lignin biosynthesis pathway (Fornalé *et al.*, 2010; Sonbol *et al.*, 2009,) and these findings were replicated in the current study. Although there are no published reports of MYB31 or MYB42 being expressed in a monocot species, it was hypothesised that similar results would be achievable in sugarcane as the MYB genes are cloned from maize, which is genetically closely related to sugarcane (Selman-Housein *et al.*, 1999). Improved saccharification of lignin reduced switchgrass expressing *Pv*MYB4 (Shen *et al*., 2012; Shen *et al*., 2013) also provided support for the replication of this result in sugarcane expressing MYB31 and MYB42. Within the MYB subgroup G4, MYB genes with down-regulatory control over lignin biosynthesis (Du *et al.*, 2012a; Fornalé *et al.*, 2010), *Pv*MYB4 is most closely related to *Zm*MYB31 and *Zm*MYB42 (Shen *et al*., 2012; Shen *et al*., 2013). Additionally, both MYB31 and MYB42 have previously been shown to suppress the maize COMT gene promoter (Fornalé *et al.*, 2006), a gene within the lignin biosynthesis pathway.

Effects on gene expression within the lignin biosynthesis pathway of MYB42 expressing plants was limited and unexpectedly led to sugarcane plants with greater reductions in lignin than any plant within the MYB31 expressing lines. Although the expression levels of the lignin biosynthesis genes suggest otherwise, it may be that MYB42 expressing plants have lower enzyme activity levels of lignin biosynthesis genes than MYB31 expressing plants, hence having lower levels of lignin deposition. Determination of enzyme activity levels of the lignin biosynthesis pathway genes may provide a better understanding of the effects of MYB31 and MYB42 on translation levels. Additionally, the inclusion of additional stem regions in qPCR analysis would provide a more detailed analysis of lignin gene

expression levels as the two stem internodes analysed (internodes 1 and 7) may not be representitive of the effects of MYB31 and MYB42 expression.

Other limitations include that only one plant for each transgenic event was analysed. While multiplying and analysing multiple clones of each transgenic event would provide more robust data, glasshouse space and time were limited. The current research did provide an overall impression on the functionality of MYB31 and MYB42 in sugarcane which was the general aim of this research project. Further research involving the expression of MYB31 and MYB42 in sugarcane would require analysing an increased number of plants to confirm the findings of this study, and field trials to confirm no negative effects on growth or sucrose production outside of glasshouse conditions, as well as the stability of MYB gene expression in transgenic sugarcane.

Although MYB31 seemed promising as a means to reduce lignin content as it was able to down-regulate many genes within the lignin biosynthesis pathway, the results found that this down-regulation did not carry over to cell wall synthesis with only a limited number of MYB31 sugarcane having reductions in lignin content. In turn this led to a limited number of MYB31 plants having improved saccharification. Alternatively, MYB42 expression in sugarcane better met the aims of this research with six MYB42 plants analysed having decreased lignin content, which led to increased glucose release by enzymatic hydrolysis with no reduction to juice sucrose levels and minimal phenotypic effects. This research highlights MYB42 as a transcription factor of interest for improving the production of second generation bioethanol from sugarcane bagasse.

Chapter 6

RNAi targeting of sugarcane lignin biosynthesis genes

The down-regulation of genes within the lignin biosynthesis pathway has been shown to be an effective way of reducing lignin content and improving the digestibility of plants for biofuel production (Hisano *et al.*, 2009; Li *et al.*, 2008). Targeting multiple genes within the pathway, for example by the expression of transcription factors, could potentially lead to indiscriminate down-regulation of multiple genes within this pathway (Table 2.2). While this may reduce lignin content, targeting specific genes for down-regulation may have a more controlled effect with additional benefits. For example, there are genes within the later stages of the lignin biosynthesis pathway with roles in the synthesis of specific monomers (Figure 1.1). By down-regulating these genes it may be possible to specifically alter the synthesis of specific lignin monomers, namely G and S (Figure 1.1). Reducing the ratio of these monomers may reduce the number of different bond types thus simplifying the deposited lignin polymer and potentially increase the release of fermentable sugars during enzymatic hydrolysis (Buanafina, 2009; Grabber, 2005; Li *et al.*, 2008). Published research has found correlations between reduced S monomers and increased saccharification (Baxter *et al.*, 2014; Chen and Dixon, 2007; Fornalé *et al.*, 2012; Fu *et al.*, 2011a; Fu *et al.*, 2011b; Jung *et al.*, 2012; Jung *et al.*, 2013; Sonbol *et al.*, 2009) which supports the purposeful alteration of the lignin monomer ratio.

This research focused on the RNAi targeting and down-regulation of caffeoyl CoA 3-O-methyltransferase (CCoAOMT), ferulate 5-hydroxylase (F5H) and caffeic acid 3-O-methyltransferase (COMT) as their positions in the lignin biosynthesis pathway would potentially reduce the G and S monomers available for lignin polymer formation (Figure 1.1). As they are positioned later in the biosynthesis pathway this may help reduce the chance of detrimental phenotypes (Chen and Dixon, 2007; Reddy *et al.*, 2005) which can occur when genes early in the pathway are down-regulated (Baucher *et al.*, 2003; Besseau *et al.*, 2007; Coleman *et al.*, 2008; Wagner *et al.*, 2007). RNAi has previously been shown as an effective method of gene down-regulation in sugarcane (Osabe *et al.*, 2009), including

the down-regulation of COMT, one of the genes focused upon in this current research (Jung *et al.*, 2012; Jung *et al.*, 2013).

CCoAOMT is a key enzyme in the synthesis of G and S monomers (Figure 1.1) (Barrière *et al.*, 2004). CCoAOMT has previously been down-regulated by 5% - 22% after RNAi expression in alfalfa (Chen and Dixon, 2007; Chen *et al.*, 2006), *Pinus radiata* (Wagner *et al.*, 2011) and maize (Li *et al.*, 2013). In all reports there was a decrease in G monomers in the lignin polymer. Sense and antisense suppression of CCoAOMT expression has produced similar results in alfalfa (Guo *et al.*, 2001a; Marita *et al.*, 2003; Nakashima *et al.*, 2008), flax (Day *et al.*, 2009), poplar (Meyermans *et al.*, 2000; Zhong *et al.*, 2000) and tobacco (Pinçon *et al.*, 2001; Zhao *et al.*, 2002). Reduced CCoAOMT expression increased the efficiency of enzymatic hydrolysis in alfalfa (Chen and Dixon, 2007) and Arabidopsis (Van Acker *et al.*, 2013), as well as improving rumen digestibility of alfalfa (Guo *et al.*, 2001a; Guo *et al.*, 2001b; Marita *et al.*, 2003). CCoAOMT may also be involved in the production of ferulate residues which aid in cross-linking cell wall components contributing to structural stability and therefore the recalcitrant nature of cell walls to enzymatic hydrolysis (Barrière *et al.*, 2004; Grabber, 2005). A reduction in CCoAOMT activity may in turn reduce the production of ferulates, resulting in reduced cross-linkages in the cell wall, increased susceptibility of the cell wall to enzymatic degradation, and improved saccharification (Buanafina, 2009; Grabber, 2005; Li *et al.*, 2008).

F5H performs a necessary step in the production of the S lignin monomer (Figure 1.1) (Weng and Chapple, 2010). A decrease or increase in F5H expression has been found to reduce or increase, respectively, the S monomer presence accordingly in alfalfa (Chen and Dixon, 2007; Chen *et al.*, 2006; Nakashima *et al.*, 2008; Reddy *et al.*, 2005), poplar (Stewart *et al.*, 2009) and Arabidopsis (Van Acker *et al.*, 2013). A decrease in F5H expression did not reduce lignin content in alfalfa (Chen and Dixon, 2007; Chen *et al.*, 2006; Nakashima *et al.*, 2008; Reddy *et al.*, 2005) or Arabidopsis (Van Acker *et al.*, 2013). These findings suggests that the down-regulation of F5H may simplify the lignin polymer by reducing S monomers which may improve the effectiveness of enzymatic hydrolysis without reducing lignin content, potentially avoiding detrimental phenotypes which can result from reductions in lignin content.

Alongside F5H, COMT is the other key gene in the production of S lignin monomers (Figure 1.1) (Bonawitz and Chapple, 2010; Weng and Chapple, 2010). The RNAi approach has been

used to down-regulate COMT expression in canola (Bhinu *et al.*, 2009), alfalfa (Chen and Dixon, 2007; Chen *et al.*, 2006), switchgrass (Baxter *et al.*, 2014; Fu *et al.*, 2011a; Samuel *et al.*, 2014) and sugarcane (Jung *et al.*, 2012; Jung *et al.*, 2013). A decrease in S monomers as well as an overall decrease in total lignin content was reported in all studies. Similar results have been found after sense or antisense down-regulation of COMT in alfalfa (Guo *et al.*, 2001a; Guo *et al.*, 2001b; Marita *et al.*, 2003; Nakashima *et al.*, 2008), poplar (Jouanin *et al.*, 2000), tobacco (Pinçon *et al.*, 2001; Zhao *et al.*, 2002) and maize (He *et al.*, 2003; Piquemal *et al.*, 2002). Down-regulation of COMT expression improved enzymatic hydrolysis in alfalfa (Chen and Dixon, 2007), Arabidopsis (Van Acker *et al.*, 2013), switchgrass (Baxter *et al.*, 2014; Fu *et al.*, 2011a) and sugarcane (Jung *et al.*, 2012; Jung *et al.*, 2013) as well as increasing rumen digestibility of alfalfa (Guo *et al.*, 2001b), maize (He *et al.*, 2003; Piquemal *et al.*, 2002) and tall fescue (Chen *et al.*, 2004).

There are currently no published reports of enzymatic hydrolysis being performed on a monocot species reduced in CCoAOMT or F5H expression. This research will provide insight into this area. If the down-regulation of CCoAOMT or F5H expression improves the yield of fermentable monosaccharides from sugarcane bagasse then this opens new opportunities for the sugarcane industry to diversify income sources by increasing the value of bagasse. The finding that RNAi reduction of COMT expression in sugarcane improves saccharification (Jung *et al.*, 2012; Jung *et al.*, 2013) is strong support that the current research will find a similar result in an important Australian commercial cultivar. The production of second generation bioethanol from this currently underutilised source of lignocellulosic sugars will also have long-term environmental and economic benefits. This research consisted of specific aims to determine if the down-regulation of these genes can improve saccharification of bagasse:

1) Clone sequences of CCoAOMT, F5H and COMT genes from sugarcane and construct RNAi plasmids for the transformation and regeneration of RNAi expressing sugarcane plants
2) Assess the expression levels of CCoAOMT, F5H and COMT to determine if the expression of these genes is reduced in their respective transgenic lines
3) Quantify secondary cell wall components (lignin, cellulose and hemicellulose) in sugarcane plants showing down-regulation of their RNAi-targeted gene
4) Perform enzymatic hydrolysis on plants with reduced lignin contents to determine if this increases the release of fermentable glucose monomers

5) Determine if the RNAi down-regulation of CCoAOMT, F5H or COMT affects juice sucrose levels and therefore the production of edible sugar from these plants.

6.1 SPECIFIC METHODS

6.1.1 Generation of RNAi constructs targeting CCoAOMT, F5H and COMT

6.1.1.1 Identification of sugarcane CCoAOMT, F5H and COMT coding sequences

Caffeoyl CoA 3-O-methyltransferase (CCoAOMT) and ferulate 5-hydroxylase (F5H) sugarcane sequences were not available as GenBank entries. Consensus sequences were developed by using available maize accessions and BLAST searching the sugarcane EST database for sequences with very high homology. These sugarcane EST sequences were aligned to create consensus sugarcane coding sequences for CCoAOMT and F5H of 774 and 888bp in length respectively. A GenBank search identified a sugarcane caffeic acid 3-O-methyltransferase (COMT) sequence with its 1089bp coding region annotated. These sequences are seen in Appendix B.

6.1.1.2 PCR amplification of sugarcane lignin biosynthetic gene fragments

Primers were designed to amplify a sequence of approximately 400bp from CCoAOMT, F5H and COMT for use as the RNAi target sequence. Each forward primer included a *SmaI* site at the 5' end for use in subsequent cloning steps (Table 6.1). The PCR reactions utilised Kapa HiFi DNA polymerase (Kapa Biosystems, MA, USA) for product amplification. CCoAOMT and COMT fragments were amplified from a cDNA template and the F5H fragment was amplified from genomic DNA. The gDNA and cDNA templates both originated from the KQ228 sugarcane cultivar. The sequence amplified for each gene is underlined in Appendix B. The PCR products were ligated into pGEM-T Easy (Promega, WI, USA) and colonies determined to contain each required amplicon were sequenced (BigDye Terminator 3.1, Life Technologies, CA, USA).

TABLE 6.1: Primers used for PCR fragment amplification of sugarcane CCoAOMT, F5H and COMT gene sequences for RNAi vector construction. Primers introduce a *SmaI* restriction site at 5' end of PCR fragment (underlined). Sequences amplified are underlined in Appendix B.

Amplicon	Forward (5'-3')	Reverse (5'-3')	Size (bp)
CCoAOMT	CCCGGGGACCTCTACCAGTACATCCTGGAC	CGTCCACGAAGACGAAGTCGAAC	415
F5H	CCCGGGCTCAAGTGCGTCATCAAGGAGAC	AAGATGTCGCCCATGTCCAGCTC	401
COMT	CCCGGGCACGGACCCGCGCTTCAACCGC	CAGCACGCACTCGACGACGATC	406

6.1.1.3 Generation of RNAi constructs

The pGEM-T Easy plasmids with sense fragments were digested with *SpeI* and *PstI* to linearise the plasmid downstream of the 3' end of the insert. Antisense fragments were digested out of the pGEM-T Easy backbone with *SphI* and *PstI*. A synthetic intron (syntron; *pers.comm*. B. Dugdale, Appendix J) was isolated from a pGEM-T Easy plasmid (previously generated, *pers. comm*. B. Williams) with *SpeI* and *SphI*. Using a functional intron, such as the syntron, as a spacer can increase the silencing effects of RNAi constructs (Wesley *et al.*, 2004). Previous research incorporating the syntron into transgene constructs has found an enhancement of transgene expression and efficient intron splicing (*pers. comm.* B. Dugdale) and was thus considered suitable as an RNAi spacer sequence.

A tri-ligation of the three fragments using T4 DNA ligase (Roche, Basel, CHE) resulted in a sense-syntron-antisense cassette flanked by *SmaI* sites in pGEM-T Easy. Integrity of this cassette was confirmed by multiple restriction digests. The cassette was isolated from pGEM-T Easy after digestion with *SmaI* and ligated into the *SmaI* site of an existing pBluescript (Agilent Technologies, CA, USA) entry vector locating the cassette between a maize Ubiquitin (Ubi) promoter and intron (iUbi) (Christensen and Quail, 1996; Christensen *et al.*, 1992) and the nopaline synthase (nos) terminator (Bevan *et al.*, 1983a) (Appendix D). The three *Zm*Ubi-iUbi-sense/syntron/antisense-nos/pBS RNAi vectors (Appendix J) were confirmed by restriction digest followed by sequencing of the backbone/insert cloning junctions.

6.1.2 Transformation of callus with *Zm*Ubi-iUbi-RNAi-nos/pBS constructs

Callus was co-bombarded with individual RNAi/pBS constructs and *Zm*Ubi-iUbi-nptII-nos/pUC19. The neomycin phosphotransferase II (nptII) gene confers G418 disulfate salt (Roche, Basel, CHE) resistance to callus for selection (Bevan *et al.*, 1983b; Colbére-Garapin *et al.*, 1981). Transformation control callus was bombarded with *Zm*Ubi-nptII-nos/pUC19 only and hereon the regenerated plants are referred to as UKN controls. Although MYB transgenic controls were co-bombarded with GFP and UKN plasmids (Section 5.1.2.1), it was decided not to use the GFP plasmid for the microprojectile bombardment of RNAi transgenic control callus as the GFP controls in the MYB experiment (Chapter 5) had very poor survival rates during glasshouse acclimatisation, with only three plants out of eleven acclimatised surviving (Table 5.4). Untransformed control (wild type) callus did not undergo microprojectile bombardment and would allow for further analysis and comparisons of the effects of RNAi gene down-regulation over the lignin biosynthesis pathway. After plant regeneration from callus, individual events were transferred to growth chambers for continued development. Each plant regenerated was considered an independent transgenic event as only one plant per callus piece was subcultured during tissue culture.

6.1.3 Screening regenerated events for RNAi/pBS construct expression

Regenerated sugarcane plants were screened by qPCR as endpoint PCR screening for the RNAi cassette from gDNA provided unreliable results due to the inability of numerous Taq polymerases to reliably amplify any section of the hairpin sequence (data not shown). The RNAi cassette was considered to be stably integrated and functioning in all regenerated plants with reduced mRNA expression of the RNAi targeted gene when compared to wild type and transgenic controls after qPCR quantification of expression levels. *Zm*Ubi-nptII-nos/pUC19 cassette was confirmed by endpoint PCR to be present in gDNA of transgenic controls.

For initial qPCR screening RNA was extracted from leaf tissue of three week old acclimatised sugarcane plants regenerated from callus. cDNA was synthesised and used in qPCR analysis. The qPCR primers (Table 6.2) were designed to amplify regions of the gene not included in the RNAi targeted sequence (underlined in Appendix B) to avoid any amplification of the expressed RNAi construct sequence. Transgenic plants representing a

TABLE 6.2: Primers designed for screening of CCoAOMT, F5H and COMT RNAi/pBS regenerated events and qPCR quantification of targeted gene expression levels.

Amplicon	Forward (5'-3')	Reverse (5'-3')	Size (bp)
CCoAOMT	ACCTCATCGCAGACGAGAAGAAC	AGCCGCTCGTGGTAGTTGAGGTAG	91
F5H	GGTTCATCGACAAGATCATCGAC	GTCGGGGCTCTTCCCGCGCTTCAC	53
COMT	TACGGGATGACGGCGTTCGAGTAC	GTGATGATGACCGAGTGGTTCTT	92

spectrum of expression levels of the RNAi targeted gene and control plants were transferred to the glasshouse.

6.1.4 Glasshouse growth and harvesting of RNAi sugarcane

Transgenic plants and controls were grown in the glasshouse for nine months before being destructively harvested for analysis. Growing tissue culture generated sugarcane under glasshouse conditions allowed for plant generation and growth conditions to be controlled throughout plant development, and therefore, any changes in lignin biosynthesis could be attributed to the RNAi constructs and not to the growing environment, which would not be possible in field conditions. All plants were watered to saturation two days before being destructively harvested to avoid results being affected by any potential drought related stress response. Harvesting occurred between 10am and 5pm over three consecutive days. Before measuring and cutting, all leaf tissue and sheaths were removed, and the internodes were counted as per van Dillewijn (1952). The length of the stalk was measured (internode 1 to the final internode); and the number of internodes and the diameter of internodes 2, 4, 6, 8, 10, 12 and 14 were recorded. Average internode length was calculated by dividing height by total number of internodes. For all analyses only internode tissue was used. The internodes collected for each assay can been seen in Table 6.3, alongside the General Methods Chapter sections with the associated protocols.

TABLE 6.3: Internodes harvested from glasshouse grown RNAi sugarcane for analysis

Internode number	Analysis conducted	Method section with protocol
1	qPCR gene expression analysis in young tissue	3.6.2
7-10	Cell wall and enzymatic hydrolysis analysis	3.7
11	qPCR gene expression analysis in maturing tissue	3.6.2
12-13	Juice analysis	3.8

6.1.5 Selection of plants and order of post-harvest analyses

The analyses conducted on the harvested sugarcane were performed in a predetermined order. As discussed in Section 5.1.3, initial regenerated sugarcane were analysed due to the timeframe and space constraints. Initially cDNA was extracted from the internodes harvested for young and maturing stem for qPCR analysis. RNAi plants were selected based on the lowest expression levels of the targeted lignin biosynthesis gene after initial qPCR results performed on leaf tissue. To determine the specificity of each RNAi construct, the expression levels of a closely related gene was also quantified: COMT in CCoAOMT RNAi plants; C3H in F5H RNAi plants; and CCoAOMT in COMT RNAi plants. CCoAOMT and COMT are both *O*-methyltransferases (Zhang *et al.*, 2012) and F5H and C3H are both hydroxylases within the P450 gene family (Liu, 2012). All primer sequences are available in Table 3.1.

After post-harvest qPCR analysis, four plants per RNAi line were selected for compositional analysis to quantify cell wall lignin and carbohydrate content. These plants were selected based on having the lowest expression levels of their targeted lignin biosynthesis gene in maturing internode tissue. Following compositional analysis, three plants per RNAi line along with randomly selected wild type and UKN controls were pretreated before undergoing enzymatic hydrolysis for 72 hours with six sampling time points. The effectiveness of the pretreatment was assessed by enzymatic hydrolysis of untreated bagasse for a random selection of plants. These selected plants had their initial bagasse stock divided with one portion being pretreated and the other portion not undergoing any pretreatment. Both bagasse portions were then enzymatically hydrolysed concurrently. Quantification of juice sugar components was also performed on the plants selected for

enzymatic hydrolysis. Plant numbers analysed at each stage are outlined in Table 6.4. All RNAi plants were compared statistically against the UKN control plants.

Statistical analysis involved either a two-tailed t-test assuming unequal variance, p = 0.05, or a one-way ANOVA with Tukey post-hoc analysis, p = 0.05, as appropriate, comparing transgenic plants to transgenic controls. As phenotypic measurements could only be made once per transgenic plant, the number of standard deviations (z scores) for each RNAi plant measurement were calculated against the UKN transgenic controls. Measurements were considered different to controls if a z score greater than 2 or -2 was calculated.

TABLE 6.4: Total number of RNAi sugarcane plants analysed at each stage of analysis

Plant	Initial leaf qPCR	Harvested plants qPCR	Cell wall composition	Enzymatic hydrolysis	Juice analysis
WT	5	7	7	3	3
UKN	5	6	6	3	3
CCoAOMT RNAi	12	9	4	3	3
F5H RNAi	15	9	4	3	3
COMT RNAi	13	8	4	3	3

6.2 RESULTS

6.2.1 Amplification and cloning

Using GenBank, sugarcane COMT was already available and consensus sequences were identified and assembled for sugarcane CCoAOMT and F5H. BLAST searching with these consensus sequences showed homology between the sugarcane gene sequences and sequences of these genes in other plant species, which supported the identity of the selected sugarcane sequences. Primers were designed to amplify approximately 400bp fragments from these consensus sequences for use in RNAi construct design as previously described. There was a minimum of 98% homology between the nucleotide sequence of the cloned fragments and the original sugarcane consensus sequences for CCoAOMT, F5H and COMT. The differences in nucleotide sequence were attributed to natural variation

between sugarcane cultivars. As the RNAi target sequences were approximately 400bp long, a 2% difference in base pair sequence was not expected to affect the silencing specificity of the RNAi constructs.

6.2.2 Regeneration, acclimatisation and qPCR screening of sugarcane plants

Plants were successfully regenerated for each RNAi construct and for the control lines after microprojectile bombardment. However, for unclear reasons, the rate of plant regeneration was lower than expected. MYB transformations (Section 5.1.2) had a regeneration rate of 0.96 plants per bombarded plate (130 plants from 135 plates, including transgenic controls) whereas the RNAi transformations had a regeneration rate of 0.40 plants per bombarded plate (49 plants from 123 plates, including transgenic controls). There was no apparent difference in wild type regeneration between MYB and RNAi transformations. mRNA transcript level was quantified in three week old leaf samples after qPCR with results suggesting that some plants were showing reduced expression levels of their respective RNAi targeted gene (Figure 6.1, Appendix K). Due to the low regeneration rate, all qPCR screened plants acclimatised were transferred to the glasshouse, and subsequently harvested.

6.2.3 qPCR analysis of harvested RNAi sugarcane

As lignin biosynthesis occurs most prevalently in the plant stem, qPCR was performed on cDNA synthesised from RNA extracted from young and maturing internode tissue. The plants from each RNAi line selected for analysis (Table 6.4) had the lowest expression levels of their respective targeted lignin biosynthetic gene in leaf tissue after initial qPCR analysis of leaf tissue (Appendix K). Expression levels of RNAi targeted genes in plants that underwent compositional analysis are seen in Figure 6.1 and all results are seen in Appendix K. The results in Figure 6.1 are representative of the overall trends seen in Appendix K. CCoAOMT and F5H RNAi plants both showed the greatest reduction of targeted gene expression in maturing tissue with little to no down-regulation observed in leaf or young internode tissue. Although CCoAOMT and F5H appear to be down-regulated in maturing tissue in the respective RNAi plants, statistical analysis did not find any significant reductions in maturing tissue (Figure 6.1). This was considered to be a result of the large variations in levels of expression in the UKN control plants. COMT RNAi plants showed little down-regulation of COMT expression in leaf, young or maturing tissues

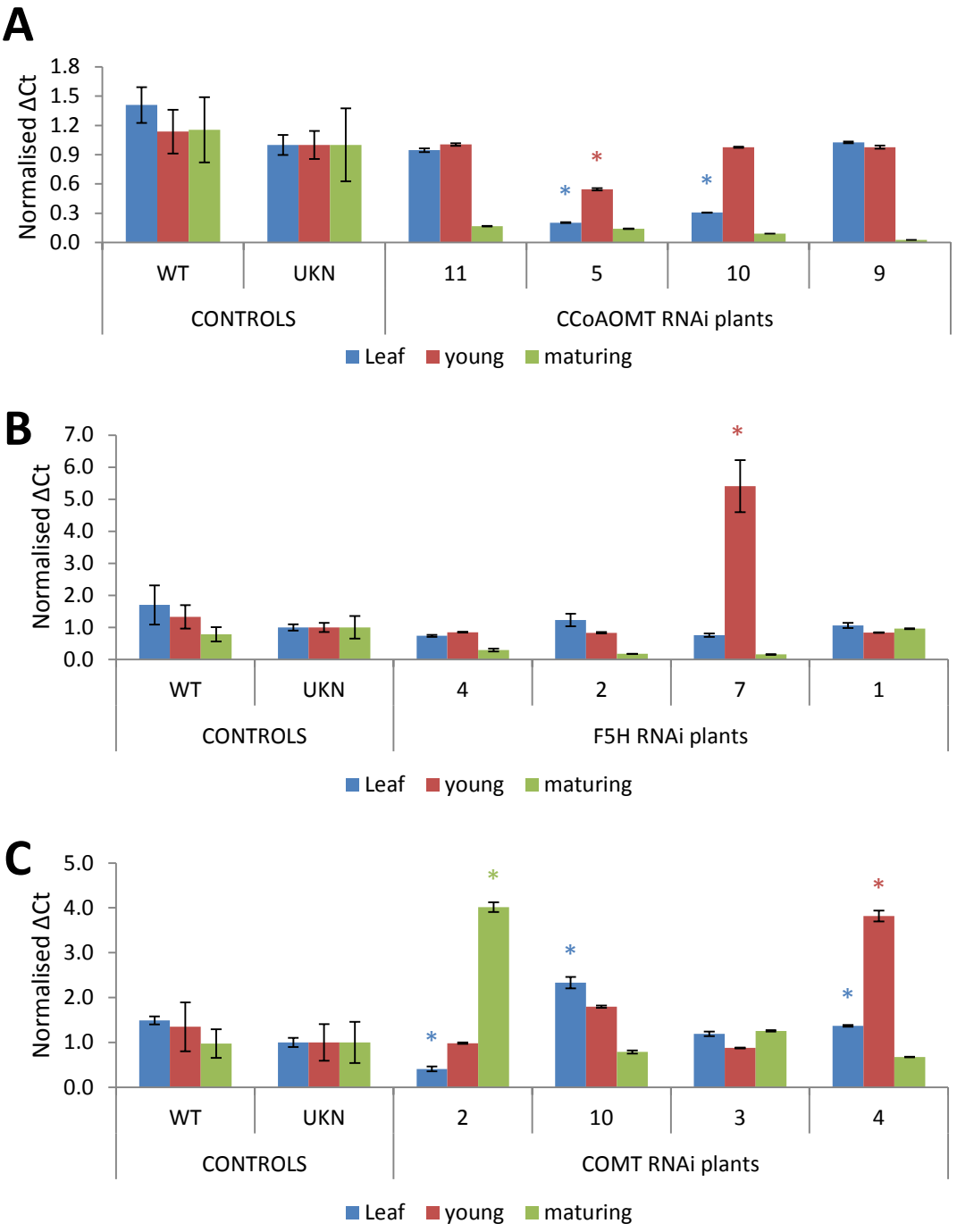


FIGURE 6.1: qPCR quantified expression (showing standard error of the mean) of RNAi targeted genes in leaf (blue), young (red) and maturing (green) internode tissue for plants that underwent cell wall compositional analysis normalised against UKN transgenic controls. RNAi plants statistically different to UKN controls after a two-tailed t-test assuming unequal variance, p = 0.05, are shown by a coloured asterisk respective to tissue type. WT n = 7; UKN n = 6; with n = individual plants analysed. Plants are listed in ascending order of total lignin content. (A) CCoAOMT expression in CCoAOMT RNAi targeted plants; (B) F5H expression in F5H RNAi targeted plants; (C) COMT expression in COMT RNAi targeted plants.

(Figure 6.1, Appendix K). In maturing tissue only two COMT RNAi plants (COMT-4 and COMT -10) showed a reduction in COMT expression of 21% - 32%. The remaining COMT RNAi plants had expression levels higher than controls including COMT-2 with a 4-fold increase in COMT expression in maturing tissue (Figure 6.1). In contrast the CCoAOMT RNAi plants that underwent compositional analysis (Figure 6.1) had reductions to CCoAOMT expression of 83% - 97%. F5H RNAi plants had mixed reductions in F5H expression of 4% - 84% (Figure 6.1).

The specificity of each RNAi construct was assessed by qPCR quantification in young and maturing internode tissue of another lignin biosynthesis gene of the same gene family: COMT expression for CCoAOMT RNAi targeted plants; C3H for F5H RNAi plants; and CCoAOMT for COMT RNAi plants. Overall the CCoAOMT, F5H and COMT RNAi sugarcane plants did not show reduced expression of COMT, C3H and CCoAOMT, respectively (Appendix L). This result supported the level of RNAi construct specificity aimed for when primers were originally designed for cloning of target sequences.

6.2.4 Cell wall compositional analysis of RNAi transgenic sugarcane

Four plants per line which showed the greatest reduction in the expression of the RNAi targeted gene in the maturing internode tissue were selected for cell wall compositional analysis (Table 6.4, Figure 6.1, Appendix K). The majority of the RNAi plants across the three construct lines had total lignin contents similar to that of controls (Table 6.5). The only exceptions were CCoAOMT-9 and F5H-1 having significant increases to total lignin and COMT-2 having significantly reduced lignin (Table 6.5). In all cases this was due to changes in the acid insoluble content of these plants. There were no CCoAOMT or F5H plants with reduced lignin and no COMT plants with increased lignin content.

There were also a number of changes to structural carbohydrate content found in RNAi lines (Table 6.5). CCoAOMT-9, -10 and -5 had significantly reduced arabinose, glucose and xylose contents respectively. F5H-1 and -4 had reduced arabinose and F5H-2 showed reductions in glucose, xylose and arabinose. The only COMT plant with a change to carbohydrate content was COMT-10 with reduced arabinose. Overall there was little change to glucose and xylose contents with nearly half of all RNAi plants showing reduced arabinose content. Galactose was not detected in any samples (controls and transgenics) in this study. There were no differences found when comparing wild type plants with UKN

TABLE 6.5: Cell wall composition of CCoAOMT, F5H and COMT RNAi sugarcane plants. The percentage of each component of the total composition is shown with the standard error of the mean. Samples significantly different to the UKN transgenic controls after a two-tailed t-test assuming unequal variance, p = 0.05, are shown in bold. Plants are listed in ascending order of total lignin content. WT n = 7; UKN n = 6; with n = individual plants analysed.

		Total lignin		Acid insoluble lignin		Acid soluble lignin		Glucose		Xylose		Galactose		Arabinose	
		%	+/-	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-	%	+/-
WT	n=7	22.09	0.27	17.50	0.26	4.59	0.03	50.55	0.46	22.72	0.26	0.0	0.0	1.81	0.06
UKN	n=6	21.65	0.38	17.13	0.38	4.53	0.09	49.83	0.55	23.53	0.25	0.0	0.0	2.03	0.08
CCoAOMT RNAi	11	21.40	0.20	16.83	0.14	4.57	0.07	52.04	0.91	23.43	0.50	0.0	0.0	1.92	0.04
	5	22.51	0.16	17.97	0.20	4.54	0.06	50.04	1.01	**20.37**	**0.42**	0.0	0.0	1.65	0.13
	10	22.61	0.17	18.07	0.13	4.55	0.08	**47.90**	**0.35**	23.37	0.17	0.0	0.0	2.07	0.01
	9	**23.05**	**0.08**	**18.48**	**0.02**	4.58	0.09	49.14	0.78	22.76	0.44	0.0	0.0	**1.68**	**0.06**
F5H RNAi	4	22.02	0.10	17.56	0.07	4.45	0.11	49.38	0.96	21.81	0.49	0.0	0.0	**1.60**	**0.01**
	2	22.61	0.25	17.97	0.19	4.65	0.07	**47.31**	**0.27**	**22.37**	**0.16**	0.0	0.0	**1.73**	**0.03**
	7	22.72	0.07	18.07	0.08	4.65	0.02	48.92	1.14	21.96	0.45	0.0	0.0	1.76	0.07
	1	**24.74**	**0.25**	**20.58**	**0.16**	4.16	0.09	49.55	0.59	23.00	0.28	0.0	0.0	**1.54**	**0.02**
COMT RNAi	2	**19.59**	**0.06**	**15.01**	**0.06**	4.58	0.09	51.07	0.92	23.10	0.45	0.0	0.0	1.87	0.04
	10	21.06	0.15	16.62	0.02	4.45	0.14	51.56	1.19	23.04	0.51	0.0	0.0	**1.63**	**0.02**
	3	21.85	0.04	17.13	0.17	4.72	0.14	49.64	0.99	24.10	0.46	0.0	0.0	1.82	0.01
	4	22.08	0.08	17.41	0.06	4.67	0.06	48.82	0.20	23.31	0.18	0.0	0.0	1.96	0.09

controls for any cell wall component analysed. Throughout the cell wall compositional analysis all mass balances were within 10% of starting material.

6.2.5 Enzymatic hydrolysis of RNAi and control bagasse samples

Enzymatic hydrolysis was performed on the three CCoAOMT and F5H plants with the highest lignin contents. As no plant had significantly less lignin (Table 6.5), it was decided to assess the plants with highest lignin content as any change in the monomer ratio that benefits saccharification may be more pronounced using these plants. As COMT-2 showed significantly reduced lignin (Table 6.5) it was decided to analyse the three COMT RNAi plants with the least amount of lignin. Three wild type and three UKN plants were selected at random to act as controls. The enzymatic hydrolysis data is presented as glucose released into the hydrolysis solution (mg/mL) per gram (g) of bagasse. Untreated and pretreated bagasse from randomly selected samples was hydrolysed to confirm the effectiveness of the pretreatment step. In all cases, pretreated samples showed a significant increase in glucose released (Figure 6.2) which supported the effectiveness of the pretreatment method used. All figures and tables list RNAi plants in ascending order of original total lignin content for each line.

The final amount of glucose released by each plant after 72 hours of enzymatic hydrolysis is seen in Figure 6.3. There was a significant difference (p = 0.044) between controls at this time point with the wild type controls releasing significantly more glucose than the UKN transgenic controls. When compared with transgenic controls each RNAi line had at least one plant release significantly more glucose than controls. CCoAOMT-5, F5H-2, COMT-2 and COMT-3 all released significantly more glucose and interestingly CCoAOMT-9 released significantly less (Figure 6.3). The remaining plants in each line were comparable with UKN control plants. There was little positive correlation between glucose in the cell wall of pretreated bagasse and glucose released after 72 hours enzymatic hydrolysis (R^2 = 0.384).

There was no correlation between the remaining cell wall components and glucose released after 72 hours enzymatic hydrolysis for pretreated bagasse (data not shown). For untreated bagasse, weak negative correlations were seen between total lignin (R^2 = -0.375) and acid insoluble lignin (R^2 = -0.5714) with glucose released after 72 hours of hydrolysis. No correlations were seen for the remaining cell wall components of untreated bagasse (data not shown).

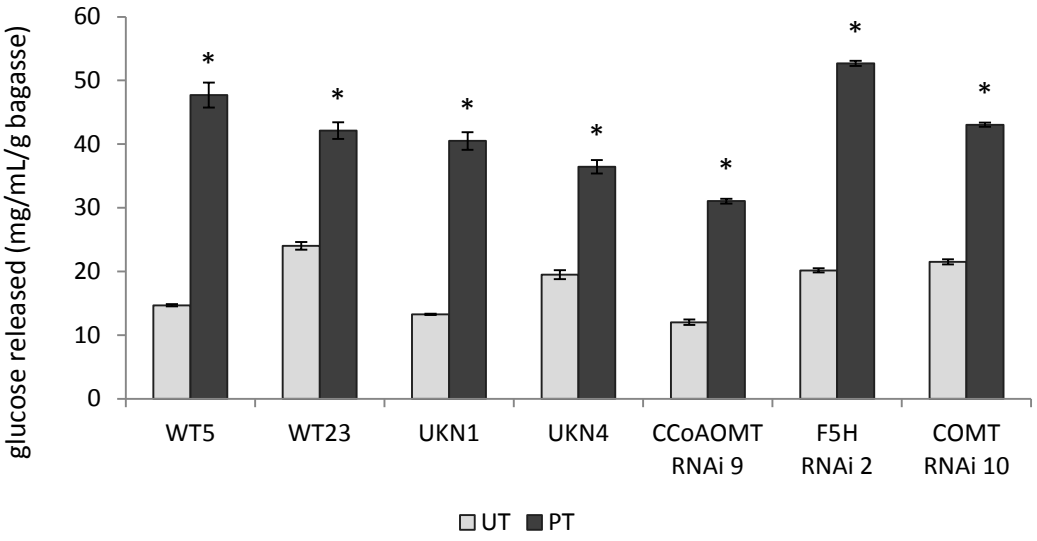


FIGURE 6.2: Total glucose concentration (mg/mL) in enzymatic hydrolysis solution per gram of untreated (UT) and pretreated (PT) bagasse (g bagasse) after 72 hours incubation showing standard error of the mean. The UT and PT results were compared for each plant individually using a two-tailed t-test assuming unequal variance, p = 0.05 with an asterisk indicating a statistical difference.

Over the course of 72 hours, enzymatic hydrolysis reactions were sampled at 0, 6, 12, 24, 48, and 72 hour time points (Table 6.6 and Figure 6.4). Although no significant differences were seen between the WT and UKN controls over the initial time points the WT controls released significantly more glucose than the UKN controls at the 24 hour and 72 hour time point. There were no clear reasons why this occurred. All four RNAi plants (CCoAOMT-5, F5H-2, COMT-2 and COMT-3) that released significantly more glucose after 72 hours of enzymatic hydrolysis (Figure 6.3) showed significant levels of glucose being released for earlier time points (Table 6.6) indicating an increased rate of glucose conversion (Figure 6.4). Furthermore, after the 48 hour time point the glucose released by the WT and UKN controls plateaus whereas the glucose released by these four RNAi plants does not appear to have started plateauing (Figure 6.4). One plant (CCoAOMT-9) released significantly less glucose than the UKN controls at each time point (Table 6.6) thus significantly reducing its rate of glucose conversion (Figure 6.4). F5H-7 was the only plant that had a significant difference in glucose release (6 hour time point) which did not differ significantly to UKN controls after 72 hours (Table 6.6).

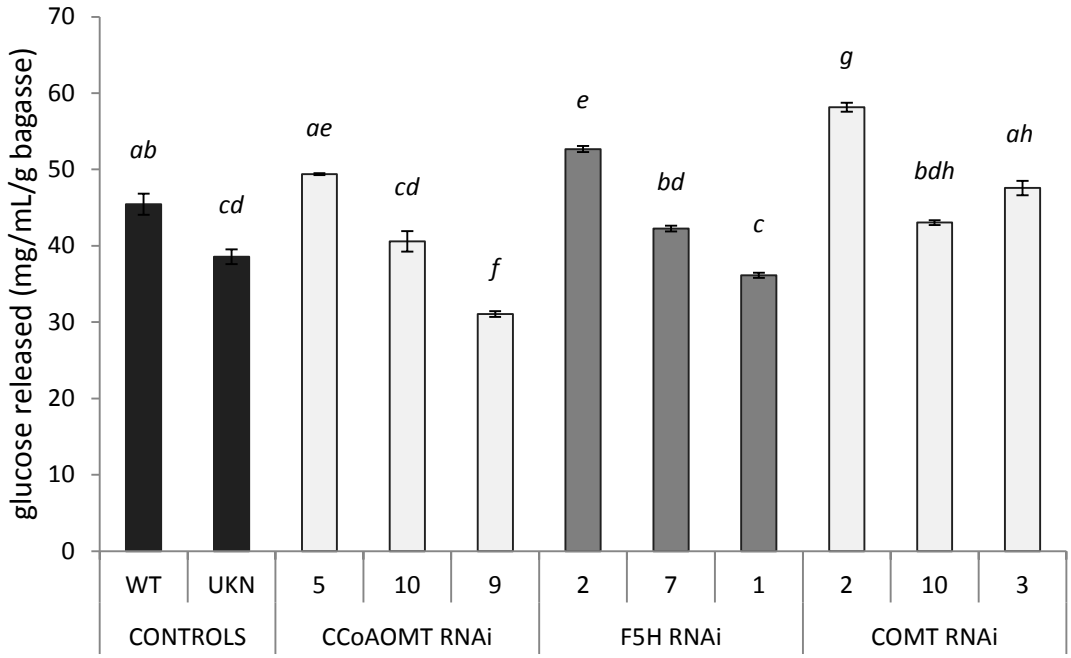


FIGURE 6.3: Total glucose concentration (mg/mL) in enzymatic hydrolysis solution per gram of bagasse (g bagasse) after 72 hours incubation showing standard error of the mean. Different letters above bars (*a*-*h*) indicate significant differences after one-way ANOVA with Tukey post-hoc analysis (p = 0.05). WT n = 3; UKN n = 3, with n = individual plants analysed. Plants for each RNAi line are listed in ascending order of total lignin content.

6.2.6 Cellulose crystallinity index of RNAi bagasse

The tissue requirements of pretreatment for enzymatic hydrolysis left enough bagasse only for some samples to undergo determination of cellulose crystallinity index (Appendix M). Statistical analysis was not performed as the limited number of samples (two control plants and four RNAi plants) would not provide reliable results. The ranges of cellulose crystallinity were 48.21% - 53.51% for control plants and 53.89% - 57.86% for the three RNAi lines (Appendix M).

TABLE 6.6: Glucose released in enzymatic hydrolysis solution (mg/mL) per gram of bagasse measured at six time points for CCoAOMT, F5H and COMT RNAi plants. The glucose released is shown with the standard error of the mean. Samples significantly different to UKN controls after a two-tailed t-test assuming unequal variance, p = 0.05 are shown in bold. WT n = 3; UKN n = 3, with n = individual plants analysed. Plants for each line are listed in ascending order of total lignin content.

		0 hrs		6 hrs		12 hrs		24 hrs		48 hrs		72 hrs	
		mg/mL	+/-	mg/mL	+/-	mg/mL	+/-	mg/mL	+/-	mg/mL	+/-	mg/mL	+/-
CONTROLS	WT	0.0	0.0	30.31	1.35	35.03	1.18	**39.78**	**0.86**	46.33	1.81	**45.46**	**1.38**
	UKN	0.0	0.0	24.69	0.56	31.38	1.03	35.50	0.68	38.86	1.04	38.58	0.96
CCoAOMT-RNAi	5	0.0	0.0	**31.46**	**0.19**	38.94	1.79	**43.42**	**0.18**	**48.21**	**0.26**	**49.39**	**0.13**
	10	0.0	0.0	25.67	1.18	28.45	0.32	32.01	0.63	37.82	1.13	40.59	1.34
	9	0.0	0.0	**18.57**	**0.05**	**23.90**	1.42	**26.51**	**0.16**	**29.60**	**0.36**	**31.06**	**0.39**
F5H-RNAi	2	0.0	0.0	**32.38**	**0.29**	**38.06**	0.35	**45.39**	**0.51**	**50.92**	**0.13**	**52.68**	**0.40**
	7	0.0	0.0	**27.53**	**0.36**	36.06	2.89	37.30	0.85	41.83	0.18	42.27	0.38
	1	0.0	0.0	22.36	0.24	31.68	4.29	32.07	0.67	35.32	0.37	36.14	0.34
COMT-RNAi	2	0.0	0.0	**32.52**	**0.33**	44.16	3.11	**47.31**	**0.85**	**54.79**	**0.79**	**58.16**	**0.59**
	10	0.0	0.0	24.87	0.23	30.75	0.30	37.04	0.56	41.27	0.05	43.05	0.32
	3	0.0	0.0	26.75	0.58	33.69	0.28	**40.80**	**0.62**	**45.50**	**0.37**	**47.57**	**0.94**

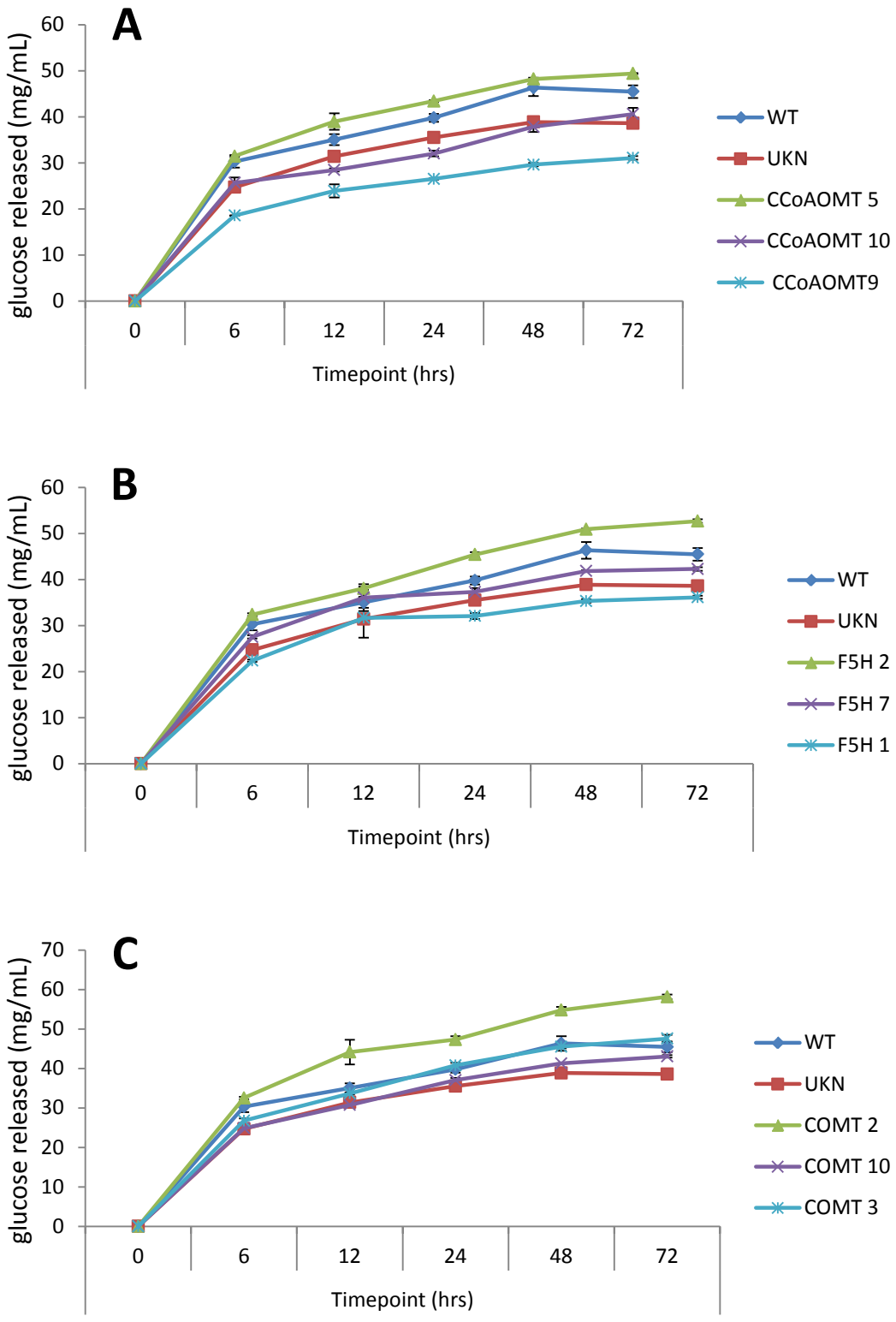


FIGURE 6.4: Total glucose concentration in enzymatic hydrolysis solution (showing standard error of the mean) measured at six time points over a period of 72 hours incubation based on weight of bagasse hydrolysed (mg/mL/g bagasse) for CCoAOMT RNAi (A); F5H RNAi (B); and COMT RNAi (C). WT n = 3; UKN n = 3, with n = individual plants analysed. Values with statistical analysis are seen in Table 6.6.

6.2.7 Sucrose content of juice extracted from RNAi plants and controls

The control and RNAi plants assessed for enzymatic hydrolysis were also assessed for sucrose content of extracted juice (Table 6.4) to determine if the changes in lignin content or structure had affected juice composition and quantity as carbon partitioned for cell wall synthesis may affect the flux of carbon directed to sucrose formation (Papini-Terzi *et al.*, 2009). Two F5H RNAi plants (1 and 7) showed a significant increase in sucrose levels when compared to UKN controls (Figure 6.5). All other plants were consistent with controls Figure 6.5). Although CCoAOMT-9 and COMT-2 have elevated levels of sucrose, they are not significant (*p* values of 0.054 and 0.275 respectively). Statistical analysis was performed twice to confirm these results. Glucose and fructose were assessed, but not detected using this method.

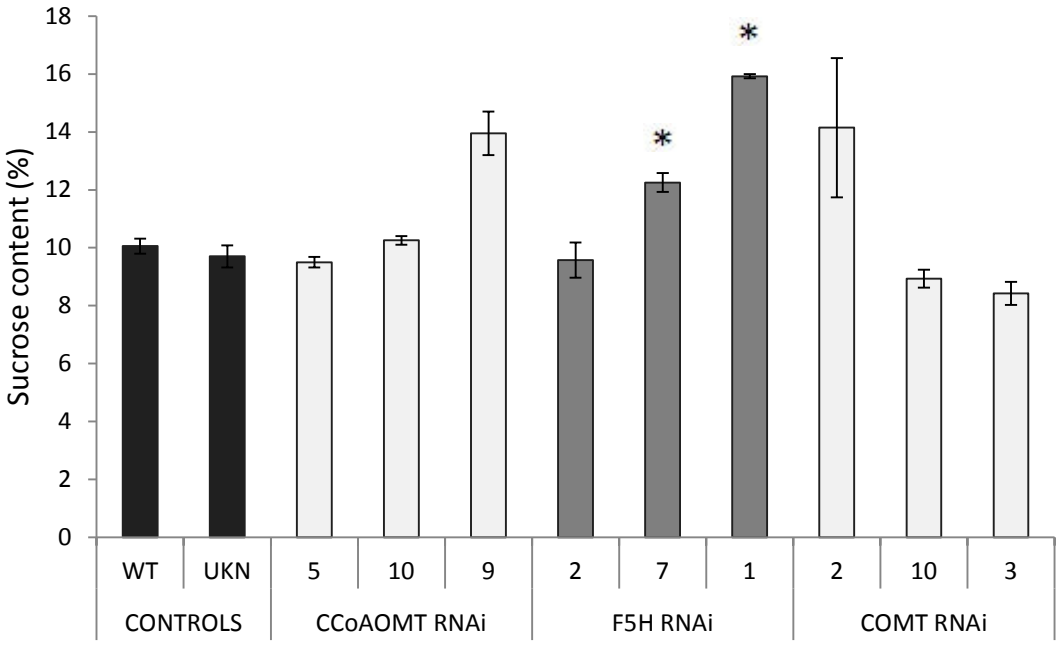


FIGURE 6.5: Sucrose content (%/fresh weight) of extracted juice of RNAi sugarcane plants selected for enzymatic hydrolysis showing standard error of the mean. RNAi plants statistically different to UKN controls after a two-tailed t-test assuming unequal variance, *p* = 0.05, are shown by an asterisk WT n = 3; UKN n = 3, with n = individual plants analysed.

6.2.8 Phenotypic measurements of RNAi plants

Phenotypic measurements were taken at the time of harvest. The height of the stalk, number of internodes, internode diameter and average internode length were recorded and calculated (Appendix N). A *z* score was calculated based on the average results of the UKN transgenic control plants and any RNAi plant with a *z* score greater than 2 or -2, indicating the result was more than two standard deviations from the control group result, were considered different to controls. Overall, there were few phenotypic differences detected between RNAi plants and controls. CCoAOMT-5 was the only plant shorter than controls with no plants have heights greater than controls (Appendix N). F5H-4 had a greater number of internodes and COMT-4 and COMT-3 had increased internode diameters (Appendix N). One plant from each RNAi line, CCoAOMT-5, F5H-4 and COMT-2, showed decreases in average internode length (Appendix N). No correlations were apparent between changes in cell wall composition (Table 6.5) and phenotypes (Appendix N).

6.3 DISCUSSION

The overall aim of this research was to improve the enzymatic digestibility of bagasse from a commercial Australian sugarcane cultivar by altering lignin deposition. These alterations were achieved by employing RNAi to specifically target and reduce the expression of three lignin biosynthesis genes: CCoAOMT, F5H and COMT. Previous research has found the down-regulation of these genes can alter the lignin polymer as well as reducing the overall deposition of lignin which has led to improved saccharification and rumen digestibility in both dicot and monocot species, including sugarcane.

6.3.1 Analysis of CCoAOMT-RNAi sugarcane

Within the lignin biosynthesis pathway, CCoAOMT is involved in the synthesis of G and S monomers (Figure 1.1) (Barrière *et al.*, 2004). CCoAOMT RNAi plants selected for compositional analysis showed reduced levels of CCoAOMT by 83% - 97% in maturing tissue (Figure 6.1). However, despite this down-regulation, no plants were found to have decreased lignin content (Table 6.5), and alternatively, CCoAOMT-9 showed significant increases in lignin content (Table 6.5). In young stem tissue, CCoAOMT expression was less reduced (0% - 46%) than in maturing stem (Figure 6.1) which may provide a reason for this

unexpected result. Lignin polymers may have been synthesised and deposited in younger tissue during higher CCoAOMT expression, and therefore, the reduction in CCoAOMT expression in maturing tissue may come too late to influence lignin content in the stem.

Additionally, the rates of protein turnover by the translated CCoAOMT enzymes in the CCoAOMT-RNAi plants are unknown. It is possible that the expression of CCoAOMT enzymes in young internode tissue (Figure 6.1) may have resulted in gene products persisting in maturing tissue, thus allowing for continued lignin biosynthesis without the need for continued high expression of CCoAOMT. Although the genes of the lignin biosynthesis pathway have been well characterised (Section 1.3), a full understanding of metabolic flux through the pathway remains to be established, making it difficult to predict the outcomes of targeting the expression levels of individual lignin biosynthesis pathway genes. Another factor that may have influenced the lignin content of the CCoAOMT plants is that any down-regulation of CCoAOMT may have been compensated for by an increase in expression levels of other genes within the lignin biosynthesis pathway, and thus the overall level of lignin biosynthesis may have not been reduced. COMT expression was quantified in the CCoAOMT-RNAi plants to determine RNAi construct specificity (Sections 6.1.5 and 6.2.3). The results suggest that COMT expression may have been increased in CCoAOMT-RNAi plants, especially in plants CCoAOMT-5 and CCoAOMT-10 (Appendix L). Quantification of all genes of the lignin biosynthesis pathway would provide a clearer picture as to the compensatory effect of non-targeted genes, however, this represented a significant amount of work and time which was not feasible given the timeframe and scope of the current research.

A reduction in lignin content after down-regulation of CCoAOMT is well documented, and the finding that CCoAOMT-9 had a significant increase in lignin content with the remaining three plants having lignin contents comparable with controls is contrary to our hypothesis. RNAi targeting of CCoAOMT reduced lignin content by 5% - 22% in alfalfa (Chen and Dixon, 2007; Chen *et al.*, 2006), *Pinus radiata* (Wagner *et al.*, 2011), and in maize (Li *et al.*, 2013). Similarly, lignin reductions of 12% - 40% were found after sense and antisense down-regulation of CCoAOMT in alfalfa (Marita *et al.*, 2003), flax (Day *et al.*, 2009), poplar (Meyermans *et al.*, 2000; Zhong *et al.*, 2000) and tobacco (Pinçon *et al.*, 2001; Zhao *et al.*, 2002) as well as in mutant *ccoaomt* Arabidopsis lines (Van Acker *et al.*, 2013). No literature was found reporting a significant increase in lignin content after RNAi targeting of lignin biosynthesis pathway gene CCoAOMT.

There were reported increases in cellulose content after CCoAOMT down-regulation in maize (Li *et al.*, 2013) and alfalfa (Marita *et al.*, 2003) which was not seen in this research with CCoAOMT-10 showing a decrease in glucose levels and the remaining three plants showing no changes (Table 6.5). Except for CCoAOMT-5 and CCoaOMT-9 having reduced xylose and arabinose respectively (Table 6.5) there were no other changes in hemicellulose content of CCoAOMT-RNAi plants which is supported by findings by Li *et al.* (2013) in maize. After 72 hours of enzymatic hydrolysis, CCoAOMT-9 was the only plant of the three RNAi lines to release significantly less glucose than controls (Table 6.6), which may be a result of a significant increase in lignin content in this plant (Table 6.5). Previous research found reduced expression of CCoAOMT improves saccharification in Arabidopsis (Van Acker *et al.*, 2013) and alfalfa (Chen and Dixon, 2007). Although no differences in lignin content were detected in CCoAOMT-5 (Table 6.5), this plant released 28% more glucose than the UKN control after 72 hours (Table 6.6) supporting the published research.

CCoAOMT is involved in the synthesis of G and S monomers (Figure 1.1) (Barrière *et al.*, 2004), and previous RNAi targeting of CCoAOMT has effected the G:S ratio in the deposited lignin polymer (Chen and Dixon, 2007; Chen *et al.*, 2006; Li *et al.*, 2013; Wagner *et al.*, 2011). Changing this monomer ratio may reduce lignin recalcitrance during saccharification by reducing the number of different monomer bond types within the deposited lignin polymer (Buanafina, 2009; Grabber, 2005; Li *et al.*, 2008). As discussed, lignin content was not reduced in any CCoAOMT-RNAi plant (Table 6.5) perhaps as a result of CCoAOMT activity in young tissue during lignin synthesis. Of the CCoAOMT-RNAi plants analysed, only CCoAOMT-5 showed a reduction in CCoAOMT expression in young stem tissue (46%) (Figure 6.1). While this reduction may not have been great enough to reduce lignin content, it may have been enough to effect the G:S monomer ratio. To accept this hypothesis, first the monomer ratio of the lignin polymer in CCoAOMT-5 would need to be determined (see Section 6.3.6).

An unexpected finding occurred when the enzymatic hydrolysis results of the wild type and transgenic UKN controls were compared. Although there appeared to be little difference in expression levels of CCoAOMT, F5H, C3H and COMT (Figure 6.1, Appendix K, Appendix L) and no significant differences in cell wall composition (Table 6.5) between controls, the wild type plants released more glucose than the UKN controls at each time point and significantly more glucose at the 24 hour and 72 hour time points (Table 6.6). The reason for these differences was not obvious and no published literature was found which

reported a similar finding, though cell wall compositional analysis of poplar found wild types differed significantly from transgenic controls expressing the *β*-glucuronidase reporter gene (Jefferson, 1987; Jefferson *et al.*, 1987) (H. Coleman, unpublished data). Performing analysis with a higher number of controls, or analysing plants of a subsequent generation may help determine if this finding is repeatable. Further research to determine if the *Zm*Ubi-nptII-nos/pUC19 (UKN control plasmid) affects saccharification would also potentially provide further insight into this finding.

6.3.2 Analysis of F5H-RNAi sugarcane

F5H is a later gene in the lignin biosynthesis pathway (Figure 1.1) and is a key enzyme in the synthesis of the S monomer (Weng and Chapple, 2010). The F5H-RNAi sugarcane analysed had minimal to no reductions of F5H expression in leaf and young internode tissue yet three of the four plants had reductions of 71% - 84% in F5H expression in the maturing stem (Figure 6.1). Despite this down-regulation of F5H in maturing tissue, no plants were found to have decreased lignin content (Table 6.5). This is similar to the CCoAOMT-RNAi findings, and as previously discussed (Section 6.3.1), F5H was not down-regulated in young stem tissue (Figure 6.1) which may have allowed lignin to be synthesised and deposited before F5H expression is reduced in more mature stem tissue. Not knowing if the F5H enzyme product synthesised in young tissue persists in maturing tissue, as well as the metabolic flux being difficult to predict after targeted gene down-regulation may also contribute to this finding (discussed in Section 6.3.1). As previously discussed (Section 6.3.1), the decreases in F5H expression in the F5H-RNAi plants may have been compensated for by other genes within the lignin biosynthesis pathway. In maturing internode tissue, C3H expression is approximately double that of controls for the plants analysed (Appendix L).

As for CCoAOMT-9, F5H-1 also showed a significant increase in lignin content (Table 6.5). No literature was found reporting a significant increase in lignin content after RNAi targeting of F5H, however Van Acker *et al.* (2013) did report an increase in lignin content in an *f5h* mutant Arabidopsis. The remaining F5H-RNAi plants with no changes to lignin content support published findings that a reduction in F5H expression in alfalfa (Chen and Dixon, 2007; Chen *et al.*, 2006; Nakashima *et al.*, 2008; Reddy *et al.*, 2005) and Arabidopsis (Van Acker *et al.*, 2013) did not lead to plants with decreased lignin content.

Additionally, Van Acker *et al.* (2013) found no detectable changes in hemicellulose polysaccharides in the mutant *f5h* Arabidopsis. Three of the four F5H-RNAi plants had decreases in arabinose content (Table 6.5), which although significant accounted for less than 0.5% difference in cell wall composition. Only F5H-2 had reduced levels of glucose and xylose (Table 6.5) with the remaining F5H-RNAi plants showing no change to glucose or xylose levels. Although having no change in lignin content (Table 6.5), F5H-2 released 36.5% more glucose than the UKN control after 72 hours (Table 6.6), which as previously discussed (Section 6.3.1), may be due to changes in G:S monomer ratio due to the role of F5H in S monomer synthesis (Figure 1.1). This improvement in glucose release supports previous findings that reduced expression of F5H improves saccharification in Arabidopsis (Van Acker *et al.*, 2013), although given that only one plant in the F5H-RNAi line released significantly more glucose, the finding by Chen and Dixon (2007) that F5H down-regulation did not improve enzymatic hydrolysis is also supported.

6.3.3 Analysis of COMT-RNAi sugarcane

Alongside F5H, COMT is another key enzyme involved in the synthesis of S monomers (Figure 1.1) (Bonawitz and Chapple, 2010; Weng and Chapple, 2010). There is no overall trend in the expression levels of COMT in the COMT-RNAi plants. COMT-2 had a 59% decrease in COMT expression in leaf tissue, little change in young internode tissue and a 4-fold increase in COMT expression in maturing tissue (Figure 6.1). COMT-10 had a 2.3 and 1.8-fold increase in COMT expression in leaf and young internode tissue respectively, but then showed a 21% decrease in expression in maturing tissue (Figure 6.1). Expression in COMT-4 showed a similar trend to COMT-10 in that COMT expression in young internode tissue was 3.8-fold higher than controls before decreasing 32% below control expression levels in maturing tissue. COMT-3 did not show any change to expression levels across the three tissue types (Figure 6.1).

In this study no COMT RNAi plants showed a decrease in COMT expression in young internode tissue (internode 1). Jung *et al.* (2012) and Jung *et al.* (2013) found decreases of 67% - 97% and 80% - 92% of COMT expression in internode three of COMT-RNAi sugarcane respectively. Although decreases in COMT expression were found in maturing internodes (internode 11) of 21% (COMT-10) and 32% (COMT-4) these decreases are minimal when compared with published results (Jung *et al.*, 2012; Jung *et al.*, 2013). Both the published studies and this study used the same COMT accession for primer design (AJ231133) and

there was approximately 250bp overlap between the RNAi target sequence used in this study and in the published studies (Jung *et al.*, 2012; Jung *et al.*, 2013). The sequence targeted by Jung *et al.* (2012) and Jung *et al.* (2013) is further upstream than the sequence targeted in this research, resulting in different sequences of the SAM-binding pocket (Jung *et al.*, 2012; Louie *et al*., 2010) being targeted by the non-overlapping regions, which may have increased the effectiveness of the Jung *et al*. (2012) COMT-RNAi construct. It has been found that secondary structure of the siRNA target site of the mRNA transcript can strongly influence the silencing efficiency of the siRNA (Bohula *et al*., 2003; Luo and Chang, 2004) and may explain the differences between this study and the results of Jung *et al*. (2012). Additionally, the differences in COMT down-regulation may also be attributed to Jung *et al.* (2012) and Jung *et al.* (2013) using the *Os*C4H promoter, *Pn*4CL spacer intron and the CaMV 35S terminator as opposed to the *Zm*Ubi promoter, syntron spacer intron and *nos* terminator used in this current research.

COMT-2 was the only plant across the three RNAi lines with a significant reduction in total lignin content (Table 6.5). Although COMT-10 and COMT-4 had reduced COMT expression in maturing stem, they had expression levels higher than controls in young tissue (Figure 6.1), which, as previously discussed in Section 6.3.1, may have allowed for lignin polymer synthesis and deposition to occur before the reduction of COMT expression. The 9.5% decrease in lignin content in COMT-2 is within the range of lignin reductions previously reported by RNAi targeting of COMT. Jung *et al.* (2012) found lignin reductions of 3.9% - 13.7% in glasshouse grown sugarcane and 5.5% - 12% reductions in field grown sugarcane (Jung *et al.*, 2013). RNAi targeting of COMT reduced lignin content by 6.4% - 14.7% in switchgrass (Baxter *et al.*, 2014; Fu *et al.*, 2011a; Samuel *et al.*, 2014). Other research has reported greater reductions of lignin content of 20% in alfalfa (Chen *et al.*, 2006), 35% in *Brassica napus* (Oraby and Ramadan, 2014) and 40% in canola (Bhinu *et al.*, 2009).

The finding that COMT-2 had significantly reduced lignin was unexpected given that COMT-2 had a 4-fold increase of COMT expression in maturing tissue and COMT expression in young stem was equivalent to that of controls (Figure 6.1). No clear reason was found that explained this finding and no literature was found that reported an increase in expression of an RNAi targeted gene. Fu *et al.* (2011a) found an approximate 90% decrease in COMT expression led to a 70% reduction in COMT enzyme activity level and a maximum reduction in lignin content of 14.7%. Similarly, Chen *et al.* (2006) found an approximate 40% - 80% reduction of COMT enzyme activity led to lignin reductions of 20%. These studies highlight

that a reduction in COMT transcription does not lead to an equal reduction in translation which in turn does not reduce lignin deposition by an equivalent percentage. Additionally, Jung *et al.* (2012) reported 65% - 95% (approximate) reductions of COMT expression in COMT-RNAi sugarcane with 3.9% - 13.7% reductions in lignin and Jung *et al.* (2013) found 5.5% - 12% less lignin in COMT-RNAi sugarcane with COMT expression reduced by 80% - 92%. Although these studies did not report an increase in COMT expression, they do support the present research that COMT expression levels may not clearly reflect changes to lignin deposition. Additionally, not knowing if the COMT enzyme product synthesised in young tissue persists in maturing tissue, as well as the metabolic flux being difficult to predict after targeted gene down-regulation may also contribute to perplexing nature of this finding (discussed in Section 6.3.1). As previously discussed (Section 6.3.1), any changes in COMT expression in the COMT-RNAi plants may have been compensated for by other genes within the lignin biosynthesis pathway. In all plants analysed, CCoAOMT expression was higher in both young and maturing internode tissue when compared with controls (Appendix L).

There were no changes in glucose levels in the cell wall of COMT-RNAi plants which supports the findings in sugarcane (Jung *et al.*, 2013) and switchgrass (Baxter *et al.*, 2014; Fu *et al.*, 2011a). Samuel *et al.* (2014) and Jung *et al.* (2013) reported higher xylose content in COMT reduced RNAi switchgrass and sugarcane respectively whereas the only change in hemicellulose content in this study was a decrease in arabinose in COMT-10 (Table 6.5). Jung *et al.* (2013) reported no changes in arabinose, though given that arabinose only accounts for a very small percentage of the cell wall it was considered unlikely that the 0.4% reduction in COMT-10 (Table 6.5) would have any noticeable effect on plant phenotype or saccharification potential. During enzymatic hydrolysis, both COMT-2 and COMT-3 released significantly more glucose after 72 hours (Figure 6.3) This is supported by the findings of published research that alfalfa (Chen and Dixon, 2007), switchgrass (Baxter *et al.*, 2014; Fu *et al.*, 2011a) and sugarcane (Jung *et al.*, 2012; Jung *et al.*, 2013) with reduced lignin content due to COMT down-regulation have increased glucose yields after enzymatic hydrolysis. After pretreatment, Jung *et al.* (2012) found glucose released by the glasshouse-grown COMT-RNAi sugarcane improved by up to 34% and field-grown RNAi-COMT sugarcane released up to 32% more glucose (Jung *et al.*, 2013). This research found COMT-2 and COMT-3 released 51% and 23% more glucose respectively than UKN controls after 72 hours (Figure 6.3, Table 6.6). COMT-2 was the only plant with a significant reduction in total lignin content (Table 6.5) which may explain the 51% increase in glucose

release. The improved glucose release by COMT-3, which had a lignin content equivalent to that of the controls (Table 6.5), may again be explained by a potential change to the G:S monomer ratio (Section 6.3.1) due to the role of COMT in the lignin biosynthesis pathway (Figure 1.1). The improvement in saccharification after COMT reduction is an encouraging result for the application of RNAi to the KQ228 sugarcane cultivar.

6.3.4 Juice sucrose levels of RNAi sugarcane

For lignin reduced sugarcane to remain commercially viable it is important that the alterations to cell wall composition do not produce detrimental phenotypes nor affect the juice sucrose content of the sugarcane as the carbon flux directed for sucrose synthesis may be affected by the partitioning of carbon for cell wall synthesis (Papini-Terzi *et al.*, 2009). The plants that underwent enzymatic hydrolysis were also assessed for juice sugar content (Figure 6.5). All CCoAOMT and COMT RNAi plants had sucrose levels comparable with controls (Figure 6.5). This partially supports the findings of Jung *et al.* (2013) in which two COMT-RNAi sugarcane plants had soluble solids (Brix) levels comparable with controls and two plants had significant reductions. Brix is a measurement of soluble solids in extracted juice and an estimate of sucrose levels (Jung *et al.*, 2013; Papini-Terzi *et al.*, 2009). Two of the F5H-RNAi plants in this study, F5H-7 and F5H-1, had significant increases in sucrose levels (Figure 6.5). This is the opposite finding of Papini-Terzi *et al.* (2009) in which F5H expression was reduced in high brix sugarcane when compared with low brix sugarcane. The finding that F5H-1 and F5H-7 had significantly increased sucrose levels is of great commercial interest as this would add significant monetary value to these sugarcane plants (*pers. comm.* I. O'Hara). An increased number of plants analysed and field trials would be necessary to confirm if this a potential avenue to increase economic value of sugarcane, or an anomaly. Papini-Terzi *et al.* (2009) did find COMT and PAL were up-regulated in high brix plants and additionally that PAL expression was inducible by sucrose. This led Papini-Terzi *et al.* (2009) to suggest increased sucrose may induce lignin biosynthesis which could explain the significant increase in lignin in F5H-1 (Table 6.5) due to the significant increase in sucrose (Figure 6.5). The research by Papini-Terzi *et al*. (2009) suggests a link between lignin biosynthesis and sucrose content, which may also be an avenue of future research if manipulations to lignin biosynthesis can also influence sucrose content.

6.3.5 Phenotypic effects of RNAi targeting of CCoAOMT, F5H and COMT

After nine months of growth in glasshouse conditions there were few phenotypic differences in the CCoAOMT, F5H and COMT-RNAi lines and the control plants (Appendix N). This is generally supported by published literature. No phenotypic differences were observed in CCoAOMT down-regulated plants (Day *et al.*, 2009; Meyermans *et al.*, 2000; Van Acker *et al.*, 2013; Zhao *et al.*, 2002), F5H down-regulated plants (Van Acker *et al.*, 2013) or in COMT down-regulated plants (Bhinu *et al.*, 2009; Fu *et al.*, 2011a; Jouanin *et al.*, 2000; Jung *et al.*, 2012; Pinçon *et al.*, 2001; Van Acker *et al.*, 2013; Zhao *et al.*, 2002). Jung *et al.* (2013) found phenotypic differences in COMT-RNAi sugarcane with lignin reductions of 8% - 12% but not in plants with lignin reductions of 6% or less. The only plant in the current study with a significant reduction in lignin content (COMT-2) had a lignin reduction of 9.5% (Table 6.5). While the height of COMT-2 was less than that of the controls, this difference was found to be less than two standard deviations, so was not considered different to controls (Appendix N).

6.3.6 NMR analysis of H:G:S monomer ratio in RNAi sugarcane bagasse

Interestingly, although CCoAOMT-5, F5H-2 and COMT-3 released significantly more glucose after enzymatic hydrolysis (Table 6.6), compositional analysis of these three plants did not show any significant differences in their lignin content (Table 6.5). Additionally, F5H-2 had significant reductions in all structural carbohydrates (Table 6.5). Improved glucose release without change to lignin content supports the hypothesis that targeting these genes for down-regulation may also change the G and S monomer ratio of the lignin polymer. Previous research has suggested changes in the ratio of these monomers in plants may simplify the bond-types in the lignin polymer and subsequently decreasing its recalcitrance to enzymatic digestion (Buanafina, 2009; Grabber, 2005; Jackson *et al.*, 2008; Li *et al.*, 2008; Sonbol *et al.*, 2009). Due to limited tissue available, Nuclear Magnetic Resonance (NMR) was employed to assess the monomer ratio in the bagasse used in this research.

After tissue had been ground in the McCrone micronising mill in preparation for enzymatic hydrolysis, a sample was set aside for NMR analysis. This sample had not undergone any pretreatment. Ground tissue was mixed with either DMSO-*d6* (Kim *et al.*, 2008) or 4:1 mix of DMSO-*d6* /pyridine-*d5* (Kim and Ralph, 2010) to measure the H:G:S ratio using gel-state NMR. After multiple attempts with optimisation to tissue preparation and NMR settings, the spectra obtained were not comparable with that of the published results (Kim and

Ralph, 2010; Kim *et al.*, 2008). The most likely reason for this is the NMR machine and probe were not suitable for gel-state samples (*pers. comm.* M. Wellard). Gel-state NMR was selected as it requires less tissue and preparation than solid-state NMR of extracted lignin polymers. As no gel-state NMR machine with a suitable probe was available at Queensland University of Technology, it was decided to send the samples to Syracuse University where a suitable NMR machine is available. At the time of this thesis submission tissue preparation and NMR conditions were being optimised and the monomer ratios were not yet available.

6.3.7 Concluding remarks

This research employed RNAi to specifically down-regulate the expression of CCoAOMT, F5H and COMT, three genes within the lignin biosynthesis pathway, with the aim of altering lignin deposition and improving the release of glucose after enzymatic hydrolysis for second generation bioethanol production. Recent publications involving RNAi down-regulation of COMT in sugarcane found increases in glucose yields after enzymatic hydrolysis (Jung *et al.*, 2012; Jung *et al.*, 2013) positively supports the replication of these results in an important Australian sugarcane cultivar. Currently, there is no published data involving the specific down-regulation of CCoAOMT or F5H in sugarcane.

The overall results of this research support the hypothesis that the down-regulation of these genes can improve saccharification potential of the sugarcane bagasse. One plant from each of the CCoAOMT and F5H-RNAi lines and two plants from the COMT-RNAi line released significantly more glucose after enzymatic hydrolysis. The reasons for this improved glucose release was not clear. Of the four plants, only COMT-2 had a significant reduction in lignin, which has previously been shown to improve saccharification (Jung *et al.*, 2012; Jung *et al.*, 2013), although, this plant did not show a decrease in COMT expression. CCoAOMT-5 and F5H-2 did show reduced expression of their respective targeted gene but no changes to lignin content. The final plant with improved saccharification, COMT-3, did not show any down-regulation to COMT expression nor any changes to lignin content. Altering the lignin monomer composition can reduce recalcitrance to enzymatic degradation (Chen and Dixon, 2007; Grabber, 2005; Huntley *et al*., 2003; Simmons *et al.,* 2010), and as CCoAOMT-5, F5H-2 and COMT-3 did not have reduced lignin content but released significantly more glucose, this supports the hypothesis that targeting CCoAOMT, F5H and COMT could influence the monomer ratio of the deposited lignin polymer based

on their positions in the lignin biosynthesis pathway. NMR is currently underway to provide data that can confirm if the monomer ratio has changed in these plants.

As previously mentioned (Section 6.2.2), there was a low rate of RNAi transgenic plant regeneration (0.40 plants per microbombarded plate) when compared with MYB regeneration rates (0.96 plants per microbombarded plate) and as discussed (Sections 6.3.1, 6.3.2 and 6.3.3), the RNAi constructs in this research do not appear as efficient at gene down-regulation or reducing lignin content when compared with published studies. It is possible that the RNAi constructs developed in this study were efficient, but reduced targeted gene expression levels to a point that the plants were not able to regenerate. If this was the reason for low regeneration rates, it would have been expected that the UKN transgenic controls would have regenerated in greater numbers, however the rate was the same for the RNAi transgenics, which does not support this hypothesis.

Finding no decreases in sucrose juice levels and few detrimental phenotypic traits is further support for the introduction of RNAi technology into the Australian sugarcane industry. Although the findings that RNAi targeting of CCoAOMT, F5H and COMT can improve saccharification is a positive result, further research should involve larger sample sizes of transgenic plants to confirm these findings are replicable as well as field testing to ensure the changes to cell wall composition are stable. Jung *et al.* (2013) found COMT-RNAi plants had stable reductions of lignin in field settings which is encouraging for the current research. The production of sugarcane bagasse more amenable to enzymatic hydrolysis through the down-regulated in CCoAOMT, F5H and COMT has the potential to increase product diversification of the global sugarcane industry. The production of bioethanol from lignin-altered bagasse will increase the value of sugarcane bagasse and benefit farmers and the wider global economy as well as having positive environmental impacts by reducing the use of non-renewable energy sources.

Chapter 7

General Discussion

As oil use is increasing at a rate unsustainable for the environment and unmatchable by current levels of oil production, a major shift away from oil as a source of energy is unavoidable (Murray and King, 2012). Increased research into renewable energy has highlighted second generation bioethanol, which can be produced from waste lignocellulosic biomass, as an alternative energy option to oil (Borrion *et al*., 2012; Limayem and Ricke, 2012). The hurdle presented by the recalcitrant nature of lignin polymers, one of the three main constituents of lignocellulosic biomass, to the cost-competitive production of second generation bioethanol is well established (Benjamin *et al*., 2013; Chen and Dixon, 2007; Yang *et al*., 2013). The findings presented in this thesis, through the completion of three specific aims, contribute to and expand current knowledge of lignin biosynthesis and manipulation in sugarcane towards the production of economically viable second generation bioethanol which will have global, as well as local, economic and environmental benefits.

The first aim of this thesis was to develop a profile of lignin biosynthesis in sugarcane stem internodes of increasing maturity by comparing expression patterns of lignin biosynthetic pathway genes with the deposition of lignin. Structural carbohydrate deposition was also examined to further knowledge of the development of sugarcane secondary cell walls. Currently there is little information available concerning lignin biosynthesis and deposition in sugarcane and no known published reports in any plant species which correlate lignin biosynthesis pathway gene expression levels with the deposition rates of lignin within the secondary cell wall. Given the importance of lignin to the cost-competitive production of second generation bioethanol, knowledge surrounding the gene expression patterns during lignin biosynthesis and deposition gained from this research project will aid in biotechnological approaches to lignin manipulation.

The expression of lignin biosynthesis genes showed two different patterns in maturing tissue. Genes either had highest expression in young tissue which decreased as tissue

matured, or expression levels remained similar throughout stem maturation. In terms of quantification of secondary cell wall components in internodes of increasing maturity, the youngest internode tissue had less lignin than the more mature sections, before a significant increase in lignin deposition resulted in a lignin content that was stable throughout remaining stem sections. Alternatively, the carbohydrate content did not change throughout stem maturation. These findings provide insight into this process and suggest that the synthesis and deposition of cell wall lignin is complete early in internode development and is preceded by high levels of lignin biosynthesis pathway gene expression levels in younger tissue. If lignin biosynthesis genes are being targeted to reduce lignin content in sugarcane, it may be most effective to use a promoter which is functional in very young tissue when the lignin biosynthesis pathway is transcriptionally most active. The high expression levels of C4H in root tissue in comparison with stem tissue expression was an unexpected finding that may also have biotechnological applications in root preferred expression in sugarcane.

As discussed in Section 4.3, there were some differences in results between published studies and the current research, with one main difference being the use of glasshouse versus field grown sugarcane plants. All results presented in this thesis are on greenhouse grown plants, and results would need to be confirmed using field grown plants. Further work into lignin biosynthesis and deposition in sugarcane should involve a comparison between these two growing environments as this may be an influencing factor on plant development. This information would be of interest for the wider research community as it would provide insight into the relatability glasshouse-based studies have with field grown plants. The gene expression patterns and cell wall deposition rates found in the first aim of this thesis will aid with targeted manipulation of the lignin biosynthesis pathway for improved bioethanol production. Although changes to lignin biosynthesis gene expression may be possible through traditional breeding, the genetic complexity of sugarcane, and the time required by traditional breeding to development and improve sugarcane cultivars, makes the application of biotechnology to sugarcane an attractive alternative (Canilha *et al.*, 2012; de Souza *et al*., 2013; Suprasanna *et al*., 2011). This concept was explored in the second and third aims of this thesis through the genetic manipulation of sugarcane lignin biosynthesis using biotechnological approaches.

The second aim of this thesis was to improve saccharification of sugarcane bagasse by reducing the lignin content in the bagasse through the generation of transgenic sugarcane

expressing maize MYB transcription factors. The involvement of MYB transcription factors in the future of GM crops has been predicted (Ambawat, *et al*., 2013; Century *et al*., 2008) and the increasing number of MYB genes being identified as having various levels of regulatory control over the lignin biosynthesis pathway (Table 2.1 and Table 2.2) increases the likelihood of the involvement of MYBs in the development of lignin-altered crops for second generation biofuel production. Previous research expressing MYB transcription factors found improved saccharification in switchgrass (Shen *et al*., 2012; Shen *et al*., 2013) and Arabidopsis (Fornalé *et al*., 2010; Sonbol *et al*., 2009) and the results of this thesis in sugarcane contributes to these positive results.

The current findings represent the first known report of MYB31 and MYB42 being expressed in a monocot species and the first report of MYB transcription factors being over-expressed in sugarcane to improve biofuel production through the down-regulation of lignin biosynthetic genes. Increased production and use of biofuels will have environmental benefits, in addition to increasing the economic value of bagasse for the sugarcane industry by adding a secondary product in the production of sugar. For this aim, MYB31 and MYB42 were cloned and transformed into sugarcane. Transgenic plants were assessed for any regulatory effects these transcription factors may have over genes within the lignin biosynthesis pathway and subsequently, it was determined if this down-regulation resulted in alterations in the deposition of secondary cell wall components. Enzymatic hydrolysis was performed on a selection of MYB expressing sugarcane to determine if changes in secondary cell wall composition increased the release of glucose monomers. Finally, the plants assessed by enzymatic hydrolysis also had juice sucrose levels quantified to determine that changes in lignin biosynthesis did not negatively affect this important aspect of sugarcane growth.

*Zm*MYB31 and *Zm*MYB42 were successfully incorporated into the genome of sugarcane. Analysis of the transgenic plants found an overall trend of MYB31 down-regulating more genes within the lignin biosynthesis pathway than MYB42. Surprisingly, this down-regulation did not correlate well with decreases in lignin content. Despite MYB31 showing the greater tendency for gene down-regulation, only one MYB31 expressing plant had a decrease in total lignin, with an additional two plants showing decreases to acid insoluble lignin contents. In contrast, six MYB42 expressing sugarcane showed significant decreases to total lignin content. A small number of MYB42 expressing plants had increased structural carbohydrate content, whereas approximately half of MYB31 plants had increased

hemicellulose contents. Further analysis on a selection of plants found that the changes in lignin content improved saccharification as two of the six MYB31 plants, and all six MYB42 plants analysed had significantly increased glucose yields after enzymatic hydrolysis. Overall the MYB42 plants outperformed the MYB31 plants not just in the number of plants with improved saccharification, but also in the amount of glucose released. For sugarcane plants with improved digestibility to be accepted by the sugarcane industry it is important that any changes to cell wall composition do not negatively affect the juice sucrose content of the sugarcane plants. The finding that only a single MYB31 plant had a reduction to sucrose content, with the remaining MYB31 plants, and all MYB42 plants, having sucrose levels equivalent to controls is a positive finding.

The overall objective of the second aim was to assess if the over-expression of MYB31 or MYB42 in sugarcane would improve saccharification through the down regulation of lignin biosynthesis. The objective was successfully achieved with the overall results suggesting that MYB42 was more effective than MYB31 at reducing lignin content of sugarcane bagasse, and subsequently improving the glucose yield after enzymatic hydrolysis. Importantly, that there was little detrimental effect on phenotype or juice sucrose levels further supports MYB42 as a potential candidate in the future of sugarcane biotechnology and the production of second generation bioethanol. Further research would involve field trials to determine the improvements found in the current research are replicable in field grown plants.

The third research aim of this thesis was to improve the digestibility of sugarcane bagasse through the expression of RNAi constructs targeting genes involved in lignin biosynthesis. This alternative biotechnological approach to lignin alteration in sugarcane was presented in this thesis as the blanket down-regulation of genes within the lignin biosynthesis pathway by transcription factors, such as MYBs, may have complex outcomes difficult to predict or control. Specifically targeting genes for down-regulation via transgenic expression of RNAi cassettes may allow for a more controlled outcome to the lignin polymer and its deposition rate. The application of this technology to sugarcane is supported by recent publications in which an RNAi construct targeting COMT, a gene within the lignin biosynthesis pathway, was expressed in sugarcane and resulted in a reduction to lignin content and improved glucose release after enzymatic hydrolysis (Jung *et al*., 2012; Jung *et al*., 2013)., This aim focused specifically on the RNAi down-regulation of CCoAOMT, F5H and COMT genes of the lignin biosynthesis pathway as the literature suggests that

targeting these genes may change the H:G:S monomer ratio, as well as the deposition rate, of lignin polymers which may further improve the bioethanol yield from this biomass. There are currently no published reports of enzymatic hydrolysis being performed on a monocot species reduced in CCoAOMT or F5H expression. This research will provide insight into this area, as well as confirming if the published results in COMT-RNAi sugarcane can be replicated in an important commercial Australian sugarcane cultivar.

This research objective required the completion of five specific aims. Initially sequences of CCoAOMT, F5H and COMT were cloned and used in the construction of RNAi vectors, which were subsequently used in the transformation and regeneration of transgenic sugarcane. The expression levels of CCoAOMT, F5H and COMT were assessed in transgenic sugarcane to confirm that these genes were being down-regulated by their respective RNAi vector prior to secondary cell wall composition being quantified. A selection of these plants underwent enzymatic hydrolysis to determine any improvements to glucose release resulting from the targeting of specific lignin biosynthesis genes. The final aim was quantifying juice sucrose levels to confirm the down-regulation of CCoAOMT, F5H or COMT did not negatively affect potential sucrose production from these plants.

qPCR analysis of regenerated plants after transformation with RNAi vectors yielded mixed results. It appeared that CCoAOMT and F5H expression was being reduced in maturing internodes, but not in younger internode tissue, whereas COMT expression levels were either comparable to controls or unexpectedly increased in young and maturing stem tissue. These changes in expression levels were not reflected in the cell wall compositions as only a single COMT-RNAi plant showed a reduction in lignin content and two plants (one CCoAOMT-RNAi and one F5H-RNAi) had increased lignin contents. Overall there was little change to cellulose and hemicellulose levels across the CCoAOMT, F5H or COMT RNAi plants. As expected, the COMT-RNAi plant with reduced lignin content released more glucose after enzymatic hydrolysis, though less expectedly, three additional plants with no change to lignin content also had higher glucose yields than controls. One possible reason for this finding is that these plants, while not having a reduction to total lignin content, had a different H:G:S monomer ratio due to the roles of CCoAOMT, F5H and COMT in the lignin biosynthesis pathway. NMR is currently being optimised to assess this possibility, and if results can confirm a change in the H:G:S monomer ratio of the lignin polymers in RNAi plants with improved saccharification but no change to lignin content, then this may contribute to a new avenue of research. Additionally, two F5H-RNAi plants had increased

levels of juice sucrose, with the remaining RNAi plants analysed having levels comparable to controls.

The results in this study are not as conclusive as those of Jung *et al*. (2012) and Jung *et al*. (2013) but do provide some positive outcomes in terms of improved saccharification of bagasse, as well as representing the first report of CCoAOMT and F5H being targeted by RNAi in sugarcane, and COMT being targeted in the Australian cultivar KQ228. The current results also suggest a number of avenues for further research. Firstly, a larger sample size would be required to assess the repeatability that RNAi targeting of CCoAOMT and F5H can improve saccharification of bagasse without reducing lignin content, and thus potentially avoiding any detrimental phenotypes associated with decreased lignin. Additionally, NMR results will also provide information on the H:G:S monomer ratios, and if they are found to be different in the transgenic RNAi plants with improved digestibility, this is another possible option to reduce lignin polymer recalcitrance in sugarcane. The finding that field grown COMT-RNAi sugarcane released higher levels of glucose after enzymatic hydrolysis (Jung *et al*., 2013) is strong support that RNAi technology can be applied to field-grown Australian cultivars with similar results, but remains to be assessed.

The overall aim of these research projects was to explore lignin biosynthesis in sugarcane with a view to alter the deposition of lignin polymers for improved enzymatic digestion. This was successfully achieved as the expression patterns of nine of the lignin biosynthesis pathway genes were determined, and this is the first known study to align these expression patterns with lignin deposition. Additionally, it was confirmed that biotechnology, in the form of MYB transcription factors and RNAi vectors, could be applied to a commercial Australian sugarcane cultivar to improve the digestion of bagasse, an important advancement towards decreasing the production costs of second generation bioethanol. In common, future research for the three research projects presented in this thesis includes analysis of field grown plants. This represents a large, but necessary step towards the realisation of GM sugarcane being grown for the dual purposes of sucrose and biofuel production. Assessing the repeatability of the results presented in this thesis through the analysis of a larger sample size of glasshouse grown plants beforehand would be a recommended step before taking this research to field trials.

The results presented in this thesis increase the understanding of lignin biosynthesis and deposition in sugarcane and will be of value when deciding the most appropriate

approaches to facilitate the development of commercial sugarcane cultivars with increased saccharification potential of their bagasse. Sugarcane bagasse is a currently underutilised source of fermentable lignocellulosic sugars, the use of which for bioethanol production can economically benefit sugarcane farmers, and increase the sustainability of the sugarcane industry, by adding value to a waste biomass product (Canilha *et al.*, 2012; Lakshmanan *et* al., 2005). Additionally, any research and development into biofuel technology conducted in Australia will potentially have a very large global appeal due to the majority of sugarcane being grown in other countries, especially Brazil, India, China and Thailand (Botha and Moore, 2014) and the majority of bioethanol currently being produced in the USA and Brazil (Chong and O'Shea, 2012). The commercialisation of renewable energy will also have environmental benefits by reducing oil use and reliance (Murray and King, 2012; Yang *et al*., 2013). For lignin reduced GM crops to become commercially viable and accepted by the farming community, they require levels of fitness equivalent to that of current wild type plants (Baxter and Stewart, 2013; Eudes *et al.*, 2014), since it is established that reducing lignin content can have detrimental effects on plant growth and development (Bonawitz and Chapple, 2013). The results presented in this thesis found increases in saccharification in both MYB and RNAi expressing plants with phenotypes generally comparable to controls. This is a positive finding and supports Jung *et al*. (2012) and Jung *et al*. (2013) that the application of biotechnology to produce lignin reduced sugarcane is a step forward for the production of cost competitive second generation bioethanol from bagasse.

Appendices

APPENDIX A: Media recipes

Lysozyme Broth (LB) liquid: 1% (w/v) bacto-tryptone, 0.5% (w/v) bacto yeast extract, 0.5% (w/v) NaCl, pH 7 – 7.5 (NaOH) (Bertani, 1951; Miller, 1972).

Lysozyme Broth (LB) solid: LB Liquid with 15g/L agar added.

Murashige and Skoog basal medium (MS) liquid: MS medium with vitamins (PhytoTechnology Laboratories, KS, USA), 2% sucrose (w/v), pH 5.7 (KOH) (Murashige and Skoog, 1962).

Murashige and Skoog basal medium (MS) semi-solid: MS liquid with 8g/L agar added.

MSC_3 medium semi-solid: MS liquid, 2% sucrose (w/v), 0.05% casein hydrolysate (w/v), 10% coconut water (v/v), 3mg/L 2,4-dichlorophenoxy acetic acid (2,4-D) (w/v), 8g/L agar, pH 5.7 (KOH) (Bower *et al.*, 1996).

MSC_3 osmotic medium semi-solid: MSC_3 medium semi-solid, 0.2M mannitol, 0.2M sorbitol (Bower *et al.*, 1996).

SOB medium liquid: 2% tryptone (w/v), 0.5% yeast extract (w/v), 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$, pH 6.7 – 7.0 (Hanahan, 1983).

Transformation buffer (TB) liquid: 10mM Pipes, 15mM $CaCl_2$, 250mM KCl, pH 6.7 (KOH), 55mM $MnCl_2$ (Inoue *et al.*, 1990).

MS regeneration media liquid: MS liquid, 0.2% myo-inositol (w/v).

APPENDIX B: Sugarcane accession sequences and constructed EST consensus sequences (5' - 3') for qPCR primer design (Table 3.1)

The CCoAOMT, F5H and COMT sequences amplified for RNAi construct design are underlined.

Phenylalanine ammonia lyase (PAL)

ATGGCGGGCAACGGCGCCATCGTGGAGAGCGACCCGCTGAACTGGGGCGCGGCGGCAGCGGAGC
TGGCGGGGAGCCACCTGGACGAGGTGAAGCGCATGGTGGCGCAGGCCCGGCAGCCCGTGGTGAA
GATCGAGGGCTCCACGCTCCGCGTCGGCCAGGTGGCCGCCGTCGCCGCCGCCAAGGACGCGTCGG
GCGTCGCCGTCGAGCTCGACGAGGAGGCCCGCCCCCGCGTCAAGGCCAGCAGCGAGTGGATCCTC
GACTGCATCGCCCACGGCGGCGACATCTACGGCGTCACCACCGGCTTCGGCGGCACCTCCCACCGC
CGCACCAAGGACGGGCCCGCTCTCCAGGTCGAGCTGCTCAGGCATCTCAACGCCGGAATCTTCGGC
ACCGGCAGCGATGGCCACACGCTGCCGTCGGAGGTCGTCCGCGCGGCGATGCTGGTGCGCATCAA
CACCCTCCTCCAGGGCTACTCGGGCATCCGCTTCGAGATCCTGGAGGCCATCACCAAGCTGCTCAAC
ACCGGGGTCAGCCCGTGCCTGCCGCTCCGGGGCACCATCACCGCGTCGGGCGACCTCGTCCCGCTC
TCCTACATCGCCGGCCTCATCACGGGCCGCCCCAACGCGCAGGCCACCACCGTCGACGGGAGGAAG
GTGGACGCCGCCGAGGCGTTCAAGATCGCCGGCATCGAGGGCGGCTTCTTCAAGCTCAACCCCAA
GGAAGGTCTCGCCATCGTCAACGGCACCTCCGTGGGCTCCGCGCTCGCGGCCACCGTGATGTACGA
CGCCAACGTCCTCACCGTCCTGTCCGAGGTCCTGTCCGCCGTCTTCTGCGAGGTGATGAACGGCAA
GCCCGAGTACACCGACCACCTCACCCACAAGCTCAAGCACCACCCGGGGTCCATCGAGGCCGCCGC
CATCATGGAGCACATCCTGGACGGCAGCGCCTTCATGAAGCACGCCAAGAAGGTGAACGAGCTGG
ACCCGCTGCTCAAGCCCAAGCAGGACAGGTACGCGCTCCGCACGTCGCCGCAGTGGCTGGGCCCCC
AGATCGAGGTCATCCGCGCCGCCACCAAGTCCATCGAGCGCGAGGTCAACTCCGTCAACGACAACC
CGGTCATCGACGTCCACCGTGGCAAGGCGCTGCACGGCGGCAACTTCCAGGGCACGCCCATCGGC
GTGTCCATGGACAACGCTCGCCTCGCCATCGCCAACATCGGCAAGCTCATGTTCGCGCAGTTCTCGG
AGCTGGTCAACGAGTTCTACAACAACGGGCTCACCTCCAACCTGGCCGGCAGCCGCAACCCCAGCC
TGGACTACGGCTTCAAGGGCACGGAGATCGCCATGGCCTCCTACTGCTCTGAGCTGCAGTACCTGG
GCAACCCCATCACCAACCACGTNCAGAGCGCGGAGCAGCACAACCAGGACGTCAACTCCCTCGGCC
TCGTCTCCGCCAGGAAGACCGCCGAGGCCATCGACATCCTGAAGCTCATGTCGTCCACCTACATCGT
GGCGCTGTGCCAGGCCATCGACCTGCGCCACCTCGAGGAGAACATCAAGACGTCGGTGAAGAACA
CGGTGACCCAGGTGGCGAAGAAGGTGCTGACCATGAACCCGTCGGGCGACCTCTCCAGCGCGCGC
TTCAGCGAGAAGGAGCTCATCACCGCCATCGACCGCGAGGGCGTGTTCACCTACGCGGAGGACCC
GGCCAGCGGCAGCCTGCCGCTGATGCAGAAGCTGCGCTCCGTGCTGGTGGACCACGCCCTCAGCA
GCGGCGACGCGGGAACGGGAGCCCTCCGTGTTCTCCAAGATCACCAATTTCGAGGAGGAGCTCCG

CGCGGGGCTGGCCCGGGAGGTGGAAGGCGCCCCGCTTCGCCGTGGGCCGAGGGCACCGCCCCCG
GGCGAAACCGGAACTGGGACAGCCGGTCGTTCCCGCTGTACCGCTTCGTCCGCGAGGAGCTCGGC
TGCGTGTTCCTGACCGGCGAGAAGCTCAAGTCCCCCGGCGAGGAGTGCACCAAGGTGTTCAACGG
CATCAGCCAGGGCAAGCTCGTCGACCCCATGCTCGAGTGCCTCAAGGAGTGGGACGGCAAGCCGC
TGCCCATCAACGTCGTCAACTAA

Cinnamate 4-hydroxylase (C4H)

ATGGACCTCGTGCTCCTGGAGAAGGCCCTTCTGGGCCTGTTCGCGGCGGCGGTGGTGGCCATCGCC
GTCGCAAAGCTGACCGGCAAGCGGTACCGCCTCCCTCCCGGCCCGCCAGGCGCCCCCGTGGTGGG
CAACTGGCTACAGGTCGGCGACGACCTGAACCACCGCAACCTGATGGCCCTCGCGAAGCGGTTCG
GCGACATCTTCCTCCTGCGCATGGGCGTGCGCAACCTGGTGGTGGTCTCGACCCCCGAGCTCGCCA
AGGAGGTGCTCCACACGCAGGGCGTGGAGTTCGGGTCCCGCACCGCAACGTCGTCTTCGACATCTT
CACGGGGAAGGGCCAGGACATGGTGTTCACCGTGTACGGCGACCACTGGCGCAAGATGCGGCGG
ATCATGACGGTGCCCTTCTTCACCAACAAGGTGGTCGCGCAGAACCGCGCCGGGTGGGAGGAGGA
GGCCCGCCTCGTGGTGGAGGACGTGCGGCGGGACCCCAGGGCCGCCGCCGAGGGCGTCGTGATC
CGGAAGCG

Coumarate 3-hydroxylase (C3H)

AGCGGTTCGTGAACGCGGCCGGCGAGCTGGACGAGCAAGGGCGCGAGTTCAAGGGGATCGTTCA
CAACGGCATCAAGATCGGCGCGTCCCTCTCCATCGCGCAGCACATCCCGTGGCTGCGGTGGCTGGC
CCCCGTCGACGAGCAGGTCTTCAAAGCCCACGGCGAACGGCGCGACCGCCTCACCGTGAAGATCAT
GGAGGAGCACGCCAAGGCCCTCAAGCAGCGCGGCGCCCAGCAGCATTTCGTCGACGCGCTCTTCA
CTCTCAGGGACAAGTACGACCTCAGCGACGACACCGTCATAGGCCTCCTCTGGGACATGATCACCG
CCGGCACAGACACGACGGTGATCTCGGTGGAGTGGGCAATGGCGGAGCTGCTGAGGAACCCCAG
GGTGCAGGAGAAGCTGCAGGAGGAGCTGGACCACGTCGTCGGCCGCGAC

4-coumarate-CoA ligase (4CL)

ATGGGTTCCGTGGACACGGCGGTCGCGGTGCCGGTGCCGGTGGCGGAGCCGGCGGCGGAGGAGA
AGGCCGTGGTGTTCCGGTCCAAGCTTCCCGACATCGAGATCAACAACAGCCAGTCGCTGCACGCCT
ACTGCTTCGGGAAGATGAGCGAGGTGGCGGACCGCGCCTGCCTCATCGACGGGCAGACCGGCGCG
TCGTACACGTACGCGGAGGTGGAGTCCCTGTCCCGCCGCGCCGCGTCGGGCCTGCGCGCCATGGG
CGTGGGCAAGGGCGACGTGGTGATGAACCTGCTCCGCAACTGCCCCGAGTTCGCCTTCACCTTCCT
GGGCGCGGCCCGGCTGGGCGCCGCCACCACCACGGCCAACCCGTTCTACACCCCGCACGAGATCCA
CCGCCAGGCGGAGGCGGCCGGCGCCAAGCTCATCGTCACCGAGGCCTGCGCCGTGGAGAAGGTG
CGCGAGTTCGCGGCG

Caffeoyl CoA 3-*O*-methyltransferase (CCoAOMT)

ATGGCCACCACGGCGACCGAGGCGGCCAAGGCCGCGCCGGCGGAGCAGGCCAACGGCGAGCAGA
AGACGCGCCACTCCGAGGTCGGCCACAAGAGCCTGCTCAAGAGCGACGACCTCTACCAGTACATCC
TGGACACGAGCGTGTACCCGCGGGAGCCGGAGAGCATGAAGGAGCTCCGCGAGATCACCGCCAA
GCACCCATGGAACCTGATGACGACCTCCGCCGACGAGGGCCAGTTCCTCAACATGCTCATCAAGCT
CATCGGCGCCAAGAAGACCATGGAGATCGGCGTCTACACCGGCTACTCCCTCCTCGCCACCGCGCT
CGCTCTCCCGGAGGACGGCACGATCTTGGCCATGGACATCAACCGCGAGAACTACGAGCTGGGCCT
GCCCTGCATCGAGAAGGCCGGCGTCGCCCACAAGATCGACTTCCGCGAGGGCCCCGCGCTCCCCGT
CCTCGACGACCTCATCGCAGACGAGAAGAACCACGGGTCGTTCGACTTCGTCTTCGTGGACGCCGA
CAAGGACAACTACCTCAACTACCACGAGCGGCTGCTCAAGCTGGTGAAGCTGGGGGGCCTCATCG
GCTACGACAACACGCTGTGGAACGGCTCCGTCGTGCTCCCCGACGACGCCCCCATGCGCAAGTACA
TCCGCTTCTACCGCGACTTCGTGCTCGTCCTCAACAAGGCGCTCGCCGCCGACGAGCGCGTCGAGAT
CTGCCAGCTCCCCGTCGGCGACGGCGTCACCCTCTGCCGCCGCGTCAAGTGA

Ferulate 5-hydroxylase (F5H)

ATGGACCTGCAGGGGATCAACCGCCGCCTCCGCGCCGCCCGGTCCGCGCTGGACCGGTTCATCGAC
AAGATCATCGACGAGCACGTGAAGCGCGGGAAGAGCCCCGACGACGCCGACGCCGACATGGTCG
ACGACATGCTCGCCTTCTTCGCCGAGGCCAAGCCGCCCAAGAAGGGGCCCGCCGCCGCCGCGGAC
GGTGACGACCTGCACAACACCCTCCGGCTCACGCGCGACAATATCAAGGCTATCATCATGGACGTG
ATGTTTGGCGGGACGGAGACGGTGGCGTCGGCGATCGAGTGGGCGATGTCGGAGATGATGCACA
GCCCCGACGACCTCCGCCGCGTGCAGCAGGAGCTCGCCGACGTCGTGGGCCTGGACCGGAACGTG
AACGAGTCGGACCTGGACAAGCTCCCCTTCCTCAAGTGCGTCATCAAGGAGACGCTCCGGCTGCAC
CCGCCGATCCCGCTGCTCCTCCACGAGACCGCCGACGACTGCGTCGTGGGCGGCTACTCCGTGCCC
AAGGGCTCCCGCGTCATGATCAACGTGTGGGCCATCGGCCGCCACCGCGGGTCCTGGAAGGACGC
CGACGTGTTCCGGCCGTCCAGGTTCACGCCCGAGGGCGAGGCCGCGGGGCTCGACTTCAAGGGCG
GCTGCTTCGAGTTCCTGCCCTTCGGCTCCGGCCGCCGGTCCTGCCCCGGCACGGCGCTGGGCCTGT
ACGCGCTGGAGCTCGCCGTCGCCCAGCTCGCGCACGGCTTCAACTGGTCGCTGCCCGACGGGATGA
AGCCGTCGGAGCTGGACATGGGCGACATCTTCGGCCTCACCGCGCCACGCGCCACAAGGCTCTACG
CCGTGCCCACGCCCCGGCTCAACTGCCCCTTGTACTGA

Cinnamyl alcohol dehydrogenase (CAD)

ATGGGGAGCCTGGCGTCCGAGAGGAAGGTGGTCGGGTGGGCCGCCAGGGACGCCACCGGACACC
TCGCCCCCTACACCTACACCCTCAGGAGCACAGGCCCTGAAGATGTGGTGGTGAAGGTGCTCTACT
GTGGGATCTGCCACACAGACATCCACCAGGCCAAGAACCACCTCGGGGCTTCAAAGTACCCCATGG
TCCCTGGGCACGAAGTGGTCGGTGAGGTGGTGGAAGTCGGGCCCGAAGTGACCAAGTACGGCGTC

GGCGACGTGGTAGGCGTCGGGGTGATCGTCGGGTGCTGCCGCGAGTGCAAACCCTGCAAGGCCAA
CGTTGAGCAGTACTGCAACAAGAAGATCTGGTCCTACAACGATGTCTACACTGACGGCCGGCCCAC
GCAGGGCGGCTTCGCCTCCACCATGGTCGTCGACCAGAAGTTTGTGATGAAGATCCCGGCGGGTCT
GGCGCCGGAGCAAGCGGCGCCGCTGCTGTGCGCGGGCGTTACGGTGTACAGCCCGCTGAAAGCCT
TTGGGCTGACGACCCCGGGCCTCCGTGGCGCGATCCTGGGCCTCGGCGGCGTGGGCCACATGGGC
GTGAAGGTGGCCAAGGCCATGGGCCACCACGTGACGGTGATCAGCTCGTCGTCCAAGAAGCGCGC
GGAGGCGATGGACCACCTGGGCGCGGACGCGTACCTGGTGAGCTCGGACGCGGCGGCCATGGCG
GCGGCCGCCGACTCGCTGGACTACATCATCGACACGGTGCCCGTGCACCACCCGCTGGAGCCCTAC
CTGGCGCTGCTGAAGCTGGACGGCAAGCACGTTCTGCTGGGCGTCATCGGCGAGCCCCTCAGCTTC
GTGTCCCCAATGGTGATGCTGGGGCGGAAGGCCATCACGGGGAGCTTCATCGGCAGCATCGACGA
GACCGCCGAGGTGCTCCAGTTCTGCGTCGACAAGGGGCTCACCTCCCAGATCGAGGTGGTCAAGAT
GGGGTACGTGAACGAGGCGCTGGATCGGCTGGAGCGCAACGACGTCCGCTACCGCTTCGTCGTCG
ACGTCGCCGGCAGCAACGTCGAGGAGGTGGCGGCCGATGCGCCGAGCAACTGA

Cinnamoyl-CoA reductase (CCR)

ATGACCGTCGTCGACGCCGTGTCCACTGATGCCGCCGGCGCCCCTGCAGCCGCCGCGGCACCGGTG
CAGCAGCCCGGGAACGGGCAGACCGTGTGCGTCACCGGTGCGGCCGGGTACATCGCCTCGTGGCT
CGTCAAGCTGCTGCTCGAGAAGGGATACACTGTCAAGGGAACAGTCAGGAACCCAGATGACCCGA
AGAACGCGCACCTCAAGGCGCTGGACGGCGCCGCCGAGCGGCTGATCCTCTGCAAGGCCGACCTC
CTGGACTACGACGCCATCTGCCGCGCCGTGCAGGGCTGCCATGGCGTCTTCCACACCGCCTCCCCG
GTCACCGACGACCCGGAGCAAATGGTGGAGCCGGCGGTGCGCGGCACGGAGTACGTGATCAACG
CGGCGGCGGAGGCCGGCACGGTGCGGCGGGTGGTGTTCACGTCCTCCATCGGCGCGGTGACCATG
GACCCCAGCCGCGGGCCCGACGTCGTGGTCGACGAGTCGTGCTGGAGCGACCTCGAGTTCTGCAA
GAAAACCAGGAACTGGTACTGCTACGGCAAGGCGGTGGCGGAGCAGGCGGCGTGGGACGCGGCC
CGGCAGCGCGGCGTGGACCTGGTGGTGGTGAACCCGGTGCTGGTGGTGGGCCCGCTGCTGCAGCC
GACGGTGAACGCCAGCATCGCGCACGTGGTCAAGTACCTGGACGGCTCCGCGCGCACCTTCGCCAA
CGCCGTGCAGGCGTACGTGGACGTCCGCGACGTCGCCGACGCGCACCTCCGCGTCTTCGAGAGCCC
GCGCGCGTCCGGCCGATACCTCTGCGCCGAGCGCGTCCTCCACCGCGAGGACGTCGTCCGCATCCT
CGCCAAGCTCTTCCCCGAGTACCCCGTCCCCACCAGGTGCTCCGACGAGGTGAACCCGCGGAAGCA
GCCGTACAAGTTCTCGAACCAGAAGCTCCGGGACCTGGGCCTGGAGTTCCGGCCGGTGAGCCAGT
CGCTGTACGACACGGTGAAGAACCTTCAGGAGAAGGGCCACCTGCCGGTGCTCGGAGAGCAGACG
ACGGAGGCCGACGACAAGGAGGCGGCCCCCGCCGCCGCCGAGCTGCAGCAGGGAGGAATCGCCA
TCCGTGCGTAA

Caffeic acid 3-O-methyltransferase (COMT)

ATGGGCTCGACCGCCGAGGACGTGGCCGCGGTGGCGGACGAGGAGGCGTGCATGTACGCGATGC
AGCTGGCGTCGGCGTCCATCCTGCCCATGACGCTGAAGAACGCGCTGGAGCTGGGCCTGCTGGAG
GTGCTGCAGGCGGAGGCGCCTGCGGGGAAGGCGCTGGCGCCCGAGGAGGTGGTGGCGCGGCTG
CCCGTGGCGCCCACCAATCCCGACGCGGCGGACATGGTGGACCGCATGCTCCGCCTCCTCGCCTCC
TACGACGTCGTCAAGTGCCAGATGGAGGACAAGGACGGCAAGTACGAGCGGCGGTACTCCGCCGC
CCCCGTCGGCAAGTGGCTCACCCCCAACGAGGACGGCGTCTCCATGGCCGCGCTCACGCTCATGAA
CCAGGACAAGGTCCTCATGGAGAGCTGGTACTACCTCAAGGACGCGGTGCTTGACGGCGGCATCC
CGTTCAACAAGGCGTACGGGATGACGGCGTTCGAGTACCACGGCACGGACCCGCGCTTCAACCGC
GTGTTCAACGAGGGCATGAAGAACCACTCGGTCATCATCACCAAGAAGCTCCTCGAGTTCTACACG
GGCTTCGAGGGCGTCTCCACGCTCGTCGACGTGGGCGGCGGCATCGGCGCCACCCTGCACGCCATC
ACCTCGCACCACCCGCAGATCAAGGGCATCAACTTCGACCTCCCCCACGTGATCTCCGAGGCGCCGC
CGTTCCCCGGCGTGCAGCACGTCGGCGGGGACATGTTCAAGTCGGTGCCGGCGGGCGACGCCATC
CTCATGAAGTGGATCCTCCACGACTGGAGCGACGCGCACTGCGCCACGCTGCTCAAGAACTGCTAC
GACGCGCTGCCGGAGAACGGCAAGGTGATCGTCGTCGAGTGCGTGCTGCCGGTGAACACCGAGGC
CGTCCCGAAGGCGCAGGGCGTGTTCCACGTCGACATGATCATGCTCGCGCATAACCCCGGCGGCAG
GGAGCGGTACGAGCGGGAGTTCCACGACCTCGCCAAGGGCGCCGGGTTCTCCGGGTTCAAGGCCA
CCTACATCTACGCCAACGCCTGGGCCATCGAGTTCATCAAGTAA

APPENDIX C: qPCR ΔCt values of lignin biosynthesis genes in sugarcane stem sections and roots for expression profile

TABLE C.1: Normalised ΔCt values of lignin biosynthesis genes showing standard error of the mean. qPCR analysis was performed on lignin biosynthesis gene expression levels in five stem sections (A-E) and roots (n = 5 per tissue section per gene). Data was normalised against section A for each individual gene.

Gene	section A		section B		section C		section D		section E		Roots	
	ΔCt (norm.)	+/-	ΔCt (norm.)	+/-	ΔCt (norm.)	+/-	ΔCt (norm.)	+/-	ΔCt (norm.)	+/-	ΔCt (norm.)	+/-
PAL	1.00	0.16	0.22	0.06	0.08	0.02	0.06	0.02	0.04	0.01	1.00	0.08
C4H	1.00	0.12	1.21	0.09	1.03	0.05	1.14	0.09	1.31	0.11	8.97	0.57
C3H	1.00	0.10	2.33	0.50	1.66	0.20	1.33	0.09	1.53	0.22	1.69	0.10
4CL	1.00	0.27	0.36	0.04	0.29	0.04	0.37	0.07	0.33	0.03	1.68	0.18
CCoAOMT	1.00	0.18	0.65	0.11	0.81	0.13	0.73	0.13	0.84	0.12	0.72	0.03
F5H	1.00	0.28	1.83	0.23	1.18	0.28	1.32	0.45	1.30	0.30	1.22	0.16
CAD	1.00	0.23	0.58	0.10	0.25	0.06	0.27	0.08	0.39	0.07	0.64	0.10
CCR	1.00	0.23	0.28	0.07	0.16	0.04	0.15	0.05	0.09	0.03	0.78	0.06
COMT	1.00	0.19	0.37	0.07	0.25	0.04	0.24	0.03	0.25	0.04	0.62	0.04

TABLE C.2: ΔCt values of lignin biosynthesis genes showing standard error of the mean. qPCR analysis was performed on lignin biosynthesis gene expression levels in five stem sections (A-E) and roots (n = 5 per tissue section per gene). Results represent raw data before normalisation.

Gene	section A		section B		section C		section D		section E		Roots	
	ΔCt	+/-	ΔCt	+/-	ΔCt	+/-	ΔCt	+/-	ΔCt	+/-	ΔCt	+/-
PAL	97.57	15.63	21.11	5.81	8.00	1.73	6.05	1.56	3.46	1.01	97.56	7.50
C4H	0.21	0.03	0.26	0.02	0.22	0.01	0.24	0.02	0.28	0.02	1.90	0.12
C3H	0.08	0.01	0.19	0.04	0.14	0.02	0.11	0.01	0.13	0.02	0.14	0.01
4CL	0.31	0.09	0.11	0.01	0.09	0.01	0.12	0.02	0.11	0.01	0.53	0.06
CCoAOMT	33.79	6.08	21.89	3.83	27.45	4.32	24.81	4.38	28.50	4.07	24.30	1.17
F5H	0.17	0.05	0.32	0.04	0.20	0.05	0.23	0.08	0.22	0.05	0.21	0.03
CAD	1.66	0.38	0.97	0.16	0.42	0.10	0.45	0.13	0.64	0.12	1.06	0.16
CCR	14.68	3.37	4.11	1.02	2.41	0.55	2.19	0.73	1.30	0.47	11.50	0.92
COMT	24.79	4.78	9.08	1.85	6.16	0.95	5.96	2.09	6.59	0.82	15.31	0.96

APPENDIX D: Sequences of maize Ubiquitin promoter (Ubi) and 5' UTR intron (iUbi) (Christensen and Quail, 1996; Christensen *et al.*, 1992) and the nopaline synthase (nos) terminator (Bevan *et al.*, 1983a) of expression vectors used in this research.

Sequences of maize Ubiquitin promoter (Ubi) and 5' UTR intron (iUbi) underlined (5' - 3')

CTGCAGTGCAGCGTGACCCGGTCGTGCCCCTCTCTAGAGATAATGAGCATTGCATGTCTAAGTTATA
AAAAATTACCACATATTTTTTTTGTCACACTTGTTTGAAGTGCAGTTTATCTATCTTTATACATATATTT
AAACTTTACTCTACGAATAATATAATCTATAGTACTACAATAATATCAGTGTTTTAGAGAATCATATA
AATGAACAGTTAGACATGGTCTAAAGGACAATTGAGTATTTTGACAACAGGACTCTACAGTTTTATC
TTTTTAGTGTGCATGTGTTCTCCTTTTTTTTTGCAAATAGCTTCACCTATATAATACTTCATCCATTTTA
TTAGTACATCCATTTAGGGTTTAGGGTTAATGGTTTTTATAGACTAATTTTTTTAGTACATCTATTTTA
TTCTATTTTAGCCTCTAAATTAAGAAAACTAAAACTCTATTTTAGTTTTTTTATTTAATAATTTAGATAT
AAAATAGAATAAAATAAAGTGACTAAAAATTAAACAAATACCCTTTAAGAAATTAAAAAAACTAAG
GAAACATTTTTCTTGTTTCGAGTAGATAATGCCAGCCTGTTAAACGCCGTCGACGAGTCTAACGGAC
ACCAACCAGCGAACCAGCAGCGTCGCGTCGGGCCAAGCGAAGCAGACGGCACGGCATCTCTGTCG
CTGCCTCTGGACCCCTCTCGAGAGTTCCGCTCCACCGTTGGACTTGCTCCGCTGTCGGCATCCAGAA
ATTGCGTGGCGGAGCGGCAGACGTGAGCCGGCACGGCAGGCGGCCTCCTCCTCCTCTCACGGCAC
CGGCAGCTACGGGGGATTCCTTTCCCACCGCTCCTTCGCTTTCCCTTCCTCGCCCGCCGTAATAAATA
GACACCCCCTCCACACCCTCTTTCCCCAACCTCGTGTTGTTCGGAGCGCACACACACACAACCAGATC
TCCCCCAAATCCACCCGTCGGCACCTCCGCTTCAAGGTACGCCGCTCGTCCTCCCCCCCCCCCCCTCT
CTACCTTCTCTAGATCGGCGTTCCGGTCCATGGTTAGGGCCCGGTAGTTCTACTTCTGTTCATGTTTG
TGTTAGATCCGTGTTTGTGTTAGATCCGTGCTGCTAGCGTTCGTACACGGATGCGACCTGTACGTCA
GACACGTTCTGATTGCTAACTTGCCAGTGTTTCTCTTTGGGGAATCCTGGGATGGCTCTAGCCGTTC
CGCAGACGGGATCGATTTCATGATTTTTTTTGTTTCGTTGCATAGGGTTTGGTTTGCCCTTTTCCTTTA
TTTCAATATATGCCGTGCACTTGTTTGTCGGGTCATCTTTTCATGCTTTTTTTGTCTTGGTTGTGATG
ATGTGGTCTGGTTGGGCGGTCGTTCTAGATCGGAGTAGAATTCTGTTTCAAACTACCTGGTGGATTT
ATTAATTTTGGATCTGTATGTGTGTGCCATACATATTCATAGTTACGAATTGAAGATGATGGATGGA
AATATCGATCTAGGATAGGTATACATGTTGATGCGGGTTTTACTGATGCATATACAGAGATGCTTTT
TGTTCGCTTGGTTGTGATGATGTGGTGTGGTTGGGCGGTCGTTCATTCGTTCTAGATCGGAGTAGA
ATACTGTTTCAAACTACCTGGTGTATTTATTAATTTTGGAACTGTATGTGTGTGTCATACATCTTCATA
GTTACGAGTTTAAGATGGATGGAAATATCGATCTAGGATAGGTATACATGTTGATGTGGGTTTTACT
GATGCATATACATGATGGCATATGCAGCATCTATTCATATGCTCTAACCTTGAGTACCTATCTATTAT
AATAAACAAGTATGTTTTATAATTATTTTGATCTTGATATACTTGGATGATGGCATATGCAGCAGCTA
TATGTGGATTTTTTTAGCCCTGCCTTCATACGCTATTTATTTGCTTGGTACTGTTTCTTTTGTCGATGC
TCACCCTGTTGTTTGGTGTTACTTCTGCA

Sequence of the nopaline synthase (nos) terminator (5' - 3')

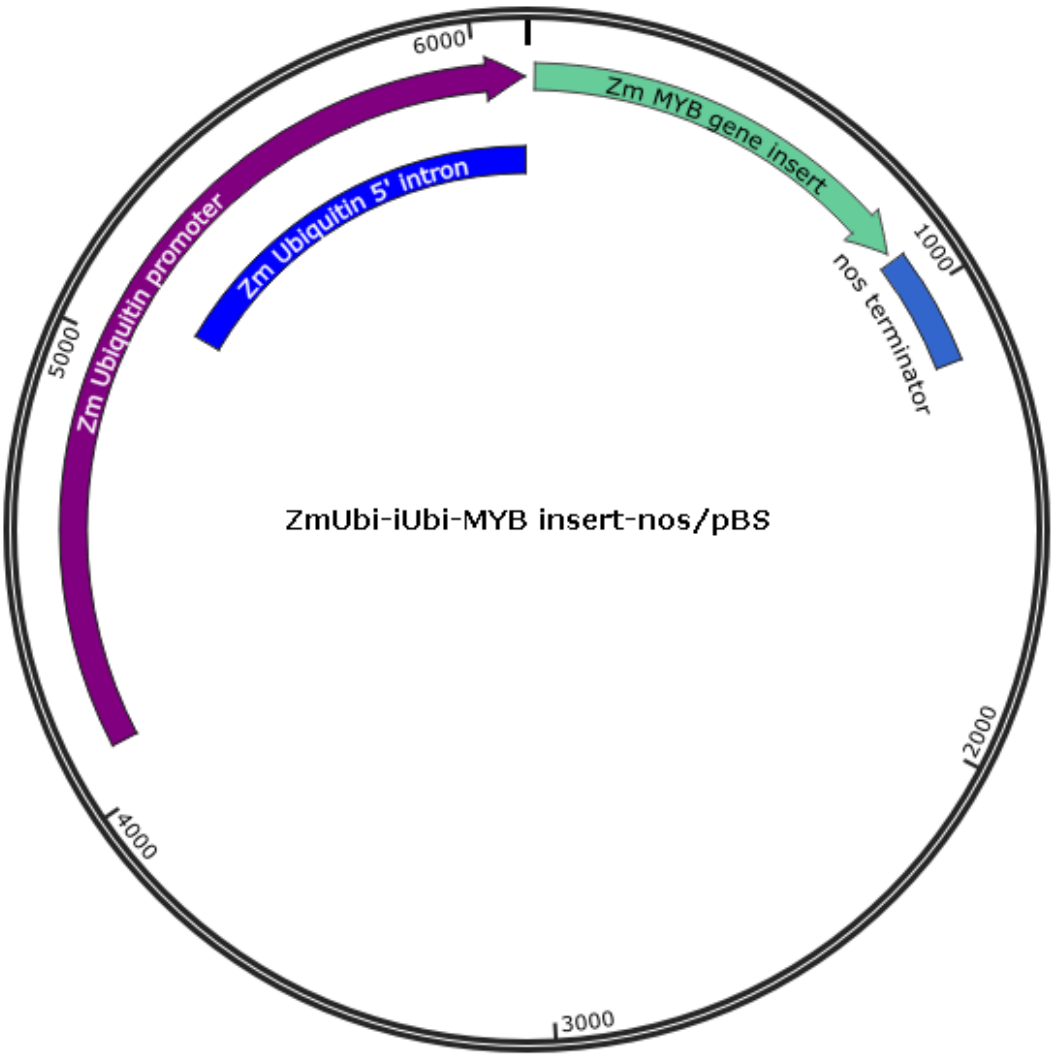
ATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATC

ATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAG

ATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGC

GCAAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATC

APPENDIX E: Generic MYB expression vector map (*Zm*Ubi-iUbi-MYB-nos/pBS) showing positions of maize Ubiquitin promoter (Ubi) and 5' UTR intron (iUbi) (Christensen and Quail, 1996; Christensen *et al.*, 1992) and the nopaline synthase (nos) terminator (Bevan *et al.*, 1983a). Individual MYB31 and MYB42 ORF and UTR sequences are seen in Appendix F and were cloned into the *Zm MYB gene insert* site of this expression vector.



APPENDIX F: Nucleotide and amino acid alignments of MYB31 and MYB42

Alignments of published (Fornalé *et al.*, 2006) and cloned nucleotide and amino acid sequences (ORF and UTR) for *Zm*MYB31 (NM_001112479) and *Zm*MYB42 (NM_001112539). Alignments were made using the Kyoto University Bioinformatics Center website (http://www.genome.jp/tools-bin/clustalw) and amino acid translations were made using the ORF region of each MYB nucleotide sequences and Vector NTI software. The start and stop codons are underlined in nucleotide sequence alignments with the 5' and 3' UTR regions being upstream and downstream of the start and stop codons respectively. The R2 and R3 motifs in each sequence are underlined with light grey and dark grey shading respectively.

* = identical match, : = conserved substitutions, . = semi-conserved substitution

MYB31 nucleotide sequence alignment

```
MYB31_ORF       ------------------------------------------------------------
MYB31_UTR       ------------ACAGCAGCAACAACAACAACAACCTCCACTGCCGCAACCCACCGAGAG
NM_001112479    GCATCGCACCTCACAGCAGCAGCAACAACAACAACCTCCACTGCCGCAACCCACCGAGAG
                         ********* ************************************

MYB31_ORF       ------------------------------------------------------------
MYB31_UTR       GCGAGACCGGCGGCGGCAAAAGGACGATACAAAAGCAGCCAGGGTTGCTGGCAACAGCGT
NM_001112479    GCGAGACCGGCGGCGGCAAAAGGACGATACAAAAGCAGCCAGGGTTGCTGGCAACAGCGT
                ************************************************************

MYB31_ORF       -------------------ATGGGGAGGTCGCCGTGCTGCGAGAAGGCGCACACCAACA
MYB31_UTR       CGGTCGCCCGCCCGCTCGCCATGGGGAGGTCGCCGTGCTGCGAGAAGGCGCACACCAACA
NM_001112479    CGGTCGCCCGCCCGTACGCCATGGGGAGGTCGCCGTGCTGCGAGAAGGCGCACACCAACA
                **************  ************************************

MYB31_ORF       AGGGCGCGTGGACCAAGGAGGAGGACGAGCGCCTGGTCGCGCACATCAGGGCGCACGGCG
MYB31_UTR       AGGGCGCGTGGACCAAGGAGGAGGACGAGCGCCTGGTCGCGCACATCAGGGCGCACGGCG
NM_001112479    AGGGCGCGTGGACCAAGGAGGAGGACGAGCGCCTGGTCGCGCACATCAGGGCGCACGGCG
                ************************************************************

MYB31_ORF       AGGGGTGCTGGCGCTCGCTGCCCAAGGCCGCCGGCCTCCTGCGCTGCGGCAAGAGCTGCC
MYB31_UTR       AGGGGTGCTGGCGCTCGCTGCCCAAGGCCGCCGGCCTCCTGCGCTGCGGCAAGAGCTGCC
NM_001112479    AGGGGTGCTGGCGCTCGCTGCCCAAGGCCGCCGGCCTCCTGCGCTGCGGCAAGAGCTGCC
                ************************************************************

MYB31_ORF       GCCTCCGCTGGATCAACTACCTCCGCCCCGACCTCAAGCGCGGCAACTTCACGGAGGAGG
MYB31_UTR       GCCTCCGCTGGATCAACTACCTCCGCCCCGACCTCAAGCGCGGCAACTTCACGGAGGAGG
NM_001112479    GCCTCCGCTGGATCAACTACCTCCGCCCCGACCTCAAGCGCGGCAACTTCACGGAGGAAG
                ********************************************************** *

MYB31_ORF       AGGACGAGCTCATCGTCAAGCTGCACAGCGTCCTCGGCAACAAGTGGTCCCTGATCGCCG
MYB31_UTR       AGGACGAGCTCATCGTCAAGCTGCACAGCGTCCTCGGCAACAAGTGGTCCCTGATCGCCG
NM_001112479    AGGACGAGCTCATCGTCAAGCTGCACAGCGTCCTCGGCAACAAGTGGTCCCTGATCGCCG
                ************************************************************

MYB31_ORF       GAAGGCTGCCCGGCAGGACGGACAACGAGATCAAGAACTACTGGAACACGCACATCCGGA
MYB31_UTR       GAAGGCTGCCCGGCAGGACGGACAACGAGATCAAGAACTACTGGAACACGCACATCCGGA
NM_001112479    GAAGGCTGCCCGGCAGGACGGACAACGAGATCAAGAACTACTGGAACACGCACATCCGGA
                ************************************************************
```

```
MYB31_ORF       GGAAGCTGCTGAGCAGGGGGATCGACCCGGTGACGCACCGCCCGGTCACGGAGCACCACG
MYB31_UTR       GGAAGCTGCTGAGCAGGGGGATCGACCCGGTGACGCACCGCCCGGTCACGGAGCACCACG
NM_001112479    GGAAGCTGCTGAGCAGGGGGATCGACCCGGTGACGCACCGCCCGGTCACGGAGCACCACG
                *****************************************************

MYB31_ORF       CGTCCAACATCACCATATCGTTCGAGACGGAGGTGGCCGCCGCCGCCCGTGATGATAAGA
MYB31_UTR       CGTCCAACATCACCATATCGTTCGAGACGGAGGTGGCCGCCGCCGCCCGTGATGATAAGA
NM_001112479    CGTCCAACATCACCATATCGTTCGAGACGGAAGTGGCCGCCGCTGCCCGTGATGATAAGA
                ****************************** *********** ****************

MYB31_ORF       AGGGCGCCGTCTTCCGGCTGGAGGACGAGGAG---------CGCAACAAGGCGACGATGG
MYB31_UTR       AGGGCGCCGTCTTCCGGCTGGAGGACGAGGAG---------CGCAACAAGGCGACGATGG
NM_001112479    AGGGCGCCGTCTTCCGGTTGGAGGACGAGGAGGAGGAGGAGCGCAACAAGGCGACGATGG
                ***************** **************         ******************

MYB31_ORF       TCGTCGGCCGCGACCGGCAGAGCCAGAGCCACAGCCACAGCCACCCCGCCGGCGAGTGGG
MYB31_UTR       TCGTCGGCCGCGACCGGCAGAGCCAGAGCCACAGCCACAGCCACCCCGCCGGCGAGTGGG
NM_001112479    TCGTCGGCCGCGACCGGCAGAGCCAGAGCCACAGCCACAGCCACCCCGCCGGCGAGTGGG
                ************************************************************

MYB31_ORF       GCCAGGGGAAGAGGCCGCTCAAGTGCCCCGACCTCAACCTGGACCTCTGCATCAGCCCGC
MYB31_UTR       GCCAGGGGAAGAGGCCGCTCAAGTGCCCCGACCTCAACCTGGACCTCTGCATCAGCCCGC
NM_001112479    GCCAGGGGAAGAGGCCGCTCAAGTGCCCCGACCTCAACCTGGACCTCTGCATCAGCCCGC
                ************************************************************

MYB31_ORF       CGTGCCAGGAGGAGGAGGAGATGGAGGAGGCTGCGATGAGAGTGAGACCGGCGGTGAAGC
MYB31_UTR       CGTGCCAGGAGGAGGAGGAGATGGAGGAGGCTGCGATGAGAGTGAGACCGGCGGTGAAGC
NM_001112479    CGTGCCAGGAGGAGGAGGAGATGGAGGAGGCTGCGATGAGAGTGAGACCGGCGGTGAAGC
                ************************************************************

MYB31_ORF       GGGAGGCCGGGCTCTGCTTCGGCTGCAGCCTGGGGCTCCCCAGGACCGCGGACTGCAAGT
MYB31_UTR       GGGAGGCCGGGCTCTGCTTCGGCTGCAGCCTGGGGCTCCCCAGGACCGCGGACTGCAAGT
NM_001112479    GGGAGGCCGGGCTCTGCTTCGGCTGCAGCCTGGGGCTCCCCAGGACCGCGGACTGCAAGT
                ************************************************************

MYB31_ORF       GCAGCAGCAGCAGCTTCCTCGGGCTCAGGACCGCCATGCTCGACTTCAGAAGCCTCGAGA
MYB31_UTR       GCAGCAGCAGCAGCTTCCTCGGGCTCAGGACCGCCATGCTCGACTTCAGAAGCCTCGAGA
NM_001112479    GCAGCAGCAGCAGCTTCCTCGGGCTCAGGACCGCCATGCTCGACTTCAGAAGCCTCGAGA
                ************************************************************

MYB31_ORF       TGAAATGA-----------------------------------------------------
MYB31_UTR       TGAAATGAGCGCGCTTCTACCCTCTCTGTGTAGCTTCTCCCCCCCGTCGTCCTCGTTTTT
NM_001112479    TGAAATGAGCGCGCTTCT-CCCTCTCTGTGTAGCTTCTCCCCCCCGTCGTCCTCGTTTTT
                ****************** *****************************************

MYB31_ORF       ------------------------------------------------------------
MYB31_UTR       GTTTTGCCACACCTCACATGGATGATGAATTGATGATACGTGGTTGGTTAGTTTTTTCGT
NM_001112479    GTTTTGCCACACCTCACATGGATGATGAATTGATGATACGTGGTTGGTTAGTTTTTTCGT
                ************************************************************

MYB31_ORF       ------------------------------------------------------------
MYB31_UTR       AGGTGAAAAATACGCGATGGTGAGCGAGTGAAAGAGAGATTTTGTGCCCTGGGTCCTCCT
NM_001112479    AGGTGAAAAATACGCGATGGTGAGCGCGTGAAAGAGAGATTTTG-GCCCTGGGTCCTCCT
                ************************** ***************** ***************

MYB31_ORF       ------------------------------------------------------------
MYB31_UTR       CCCTGCTCTCTCGTGGTGCCCCATTGCGCCTCCTCTGTCCCCCCTCTCTCTCTCTGTATC
NM_001112479    CCGTGCTCTCTCGTGGTGCCCCATTGCGCCTCCTCTGTCCCCCCTCTCTCTCTCTGTATC
                ** ********************************************************

MYB31_ORF       ------------------------------------------------------------
MYB31_UTR       TCTGTAATCACCATCGCCA--------------------------------------------
NM_001112479    TCTGTAATCACCATCGCCAAAATGATCATGGGGGGCAATATAATATAATACATGCTGCTAA
                ********************

MYB31_ORF       ------
MYB31_UTR       ------
NM_001112479    TGCTAT
```

MYB31 amino acid sequence alignment

```
MYB31_ORF       MGRSPCCEKAHTNKGAWTKEEDERLVAHIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY
MYB31_UTR       MGRSPCCEKAHTNKGAWTKEEDERLVAHIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY
NM_001112479    MGRSPCCEKAHTNKGAWTKEEDERLVAHIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY
                ************************************************************

MYB31_ORF       LRPDLKRGNFTEEEDELIVKLHSVLGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLSRG
MYB31_UTR       LRPDLKRGNFTEEEDELIVKLHSVLGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLSRG
NM_001112479    LRPDLKRGNFTEEEDELIVKLHSVLGNKWSLIAGRLPGRTDNEIKNYWNTHIRRKLLSRG
                ************************************************************

MYB31_ORF       IDPVTHRPVTEHHASNITISFETEVAAAARDDKKGAVFRLEDEE---RNKATMVVGRDRQ
MYB31_UTR       IDPVTHRPVTEHHASNITISFETEVAAAARDDKKGAVFRLEDEE---RNKATMVVGRDRQ
NM_001112479    IDPVTHRPVTEHHASNITISFETEVAAAARDDKKGAVFRLEDEEEEERNKATMVVGRDRQ
                ****************************************   *************

MYB31_ORF       SQSHSHSHPAGEWGQGKRPLKCPDLNLDLCISPPCQEEEEMEEAAMRVRPAVKREAGLCF
MYB31_UTR       SQSHSHSHPAGEWGQGKRPLKCPDLNLDLCISPPCQEEEEMEEAAMRVRPAVKREAGLCF
NM_001112479    SQSHSHSHPAGEWGQGKRPLKCPDLNLDLCISPPCQEEEEMEEAAMRVRPAVKREAGLCF
                ************************************************************

MYB31_ORF       GCSLGLPRTADCKCSSSSFLGLRTAMLDFRSLEMK
MYB31_UTR       GCSLGLPRTADCKCSSSSFLGLRTAMLDFRSLEMK
NM_001112479    GCSLGLPRTADCKCSSSSFLGLRTAMLDFRSLEMK
                ***********************************
```

MYB42 nucleotide sequence alignment

```
MYB42_ORF       ------------------------------------------------------------
MYB42_UTR       ------------------------------------ACTCGCTGCCTTCTCAAATCCA
NM_001112539    CGGGCGCAGGAATTCGGCACGAGGGGAAACCCGCGCCCACTCGCTGCCTTCTCAAATCCA
                                                     **********************

MYB42_ORF       ----------------------------------------ATGGGGCGGTCGCCGTGC
MYB42_UTR       AACGCGAAGTAGCAACAAGCAAAAGCCCAGATCGATAATACGATGGGGCGGTCGCCGTGC
NM_001112539    AACGCGAAGTAGCAACAAGCAAAAGCCCAGATCGATAATACGATGGGGCGGTCGCCGTGC
                ************************************************************

MYB42_ORF       TGCGAGAAGGCGCACACCAACAGGGGCGCGTGGACCAAGGAGGAGGACGAGCGGCTGGTG
MYB42_UTR       TGCGAGAAGGCGCACACCAACAGGGGCGCGTGGACCAAGGAGGAGGACGAGCGGCTGGTG
NM_001112539    TGCGAGAAGGCGCACACCAACAGGGGCGCGTGGACCAAGGAGGAGGACGAGCGGCTGGTG
                ****************************************************************

MYB42_ORF       GCCTACGTCCGCGCGCACGGCGAAGGGTGCTGGCGCTCGCTGCCCAGGGCGGCGGGCCTG
MYB42_UTR       GCCTACGTCCGCGCGCACGGCGAAGGGTGCTGGCGCTCGCTGCCCAGGGCGGCGGGCCTG
NM_001112539    GCCTACGTCCGCGCGCACGGCGAAGGGTGCTGGCGCTCGCTGCCCAGGGCGGCGGGCCTG
                ************************************************************

MYB42_ORF       CTGCGCTGCGGCAAGAGCTGCCGCCTGCGCTGGATCAACTACCTCCGCCCGGACCTCAAG
MYB42_UTR       CTGCGCTGCGGCAAGAGCTGCCGCCTGCGCTGGATCAACTACCTCCGCCCGGACCTCAAG
NM_001112539    CTGCGCTGCGGCAAGAGCTGCCGCCTGCGCTGGATCAACTACCTCCGCCCGGACCTCAAG
                ************************************************************

MYB42_ORF       CGAGGCAACTTCACCGCCGACGAGGACGACCTCATCGTCAAGCTGCACAGCCTGCTCGGG
MYB42_UTR       CGCGGCAACTTCACCGCCGACGAGGACGACCTCATCGTCAAGCTGCACAGCCTGCTCGGG
NM_001112539    CGAGGCAACTTCACCGCCGACGAGGACGACCTCATCGTCAAGCTGCACAGCCTCCTCGGG
                ** ********************************************** ******
```

```
MYB42_ORF       AACAAGTGGTCGCTCATCGCCGCGCGGCTCCCGGGGCGGACGGACAACGAGATCAAGAAC
MYB42_UTR       AACAAGTGGTCGCTCATCGCCGCGCGGCTCCCGGGGCGGACGGACAACGAGATCAAGAAC
NM_001112539    AACAAGTGGTCGCTCATCGCCGCGCGGCTCCCGGGGCGGACGGACAACGAGATCAAGAAC
                ************************************************************

MYB42_ORF       TACTGGAACACGCACATCCGGCGCAAGCTGCTGTGCAGCGGCATCGACCCCGTCACGCAC
MYB42_UTR       TACTGGAACACGCACATCCGGCGCAAGCTGCTGTGCAGCGGCATCGACCCCGTCACGCAC
NM_001112539    TACTGGAACACGCACATCCGGCGCAAGCTGCTGGGCAGCGGCATCGACCCCGTCACGCAC
                ********************************** **************************

MYB42_ORF       CGCCGTGTCGCGGGTGGCGCCGCGACCACCATCTCGTTCCAGCCCAGCCCCAACTCCGCC
MYB42_UTR       CGCCGTGTCGCGGGTGGCGCCGCGACCACCATCTCGTTCCAGCCCAGCCCCAACTCCGCC
NM_001112539    CGCCGCGTCGCGGGGGGCGCCGCGACCACCATCTCGTTCCAGCCCAGCCCCAACTCCGCC
                ***** ******** ******************************************

MYB42_ORF       GCCGCCGCCGCC---GCAGAAGCAGCAGCGCAGGCGCCGATCAAGGCCGAGGAGACGGCG
MYB42_UTR       GCCGCCGCCGCC---GCAGAAGCAGCAGCGCAGGCGCCGATCAAGGCCGAGGAGACGGCG
NM_001112539    GCCGCCGCCGCCGCCGCAGAAACAGCAGCGCAGGCGCCGATCAAGGCCGAGGAGACGGCG
                ************   ****** ************************************

MYB42_ORF       GGCGTCAAGGCGCCCAGGTGCCCTGACCTCAACCTGGACCTCTGCATCAGCCCGCCGTGC
MYB42_UTR       GGCGTCAAGGCGCCCAGGTGCCCTGACCTCAACCTGGACCTCTGCATCAGCCCGCCGTGC
NM_001112539    GCCGTCAAGGCGCCCAGGTGCCCCGACCTCAACCTGGACCTCTGCATCAGCCCGCCGTGC
                * ********************* ************************************

MYB42_ORF       CAGCATGAGGACGACGGCGAGGAGGAGGACGAGGAGCTGGACCTCAAGCCCGCCTTCGTC
MYB42_UTR       CAGCATGAGGACGACGGCGAGGAGGAGGACGAGGAGCTGGACCTCAAGCCCGCCTTCGTC
NM_001112539    CAGCATGAGGACGACGGCGAGGAGGAGGACGAGGAGCTGGACCTCAAGCCCGCCTTCGTC
                ************************************************************

MYB42_ORF       AAGCGGGAGGCGCTGCAGGCCGGCCACGGCCACGGCCACGGCCTCTGCCTCGGCTGCGGC
MYB42_UTR       AAGCGGGAGGCGCTGCAGGCCGGCCACGGCCACGGCCACGGCCTCTGCCTCGGCTGCGGC
NM_001112539    AAGCGGGAGGCGCTGCAGGCCGGCCACGGCCACGGCCACGGCCTCTGCCTCGGCTGCGGC
                ************************************************************

MYB42_ORF       CTGGGCGGACAGAAGGGAGCGGCCGGGTGCAGCTGCAGCAACGGCCACCACTTCCTGGGG
MYB42_UTR       CTGGGCGGACAGAAGGGAGCGGCCGGGTGCAGCTGCAGCAACGGCCACCACTTCCTGGGG
NM_001112539    CTGGGCGGACAGAAGGGAGCGGCCGGGTGCAGCTGCAGCAACGGCCACCACTTCCTGGGG
                ************************************************************

MYB42_ORF       CTCAGGACCAGCGTGCTCGACTTCAGAGGCCTGGAGATGAAGTGA--------------
MYB42_UTR       CTCAGGACCAGCGTGCTCGACTTCAGAGGCCTGGAGATGAAGTGAACGAAACGAAGCCCA
NM_001112539    CTCAGGACCAGCGTGCTCGACTTCAGAGGCCTGGAGATGAAGTGAACGAAACGAAGCCCA
                ************************************************************

MYB42_ORF       ------------------------------------------------------------
MYB42_UTR       CACGTCCTTTCTTCTCC--------------------------------------------
NM_001112539    CACGTCCTTTCTTCTCCTTTTGTTGTCGGTTGTAGTCTTGGCTTGTTGGATTTGGATAGA
                *****************

MYB42_ORF       ------------------------------------------------------------
MYB42_UTR       ------------------------------------------------------------
NM_001112539    GCTAGTTGGTTACTAGTTGTTAGTTAGAAGATAGTGCAGGATGATCACTAGCTACTGGCT


MYB42_ORF       ------------------------------------------------------------
MYB42_UTR       ------------------------------------------------------------
NM_001112539    ACCTCAACAGTACAGTAGCTGCTCCCTTCTCTTCCATTCTATGTAAAAAAGAAACAAAAA


MYB42_ORF       ------------------------------------------------
MYB42_UTR       ------------------------------------------------
NM_001112539    TACTTATAAGGTGTTTGGGTTGAGAAATGAACTAGTCTATTATCTTTTC
```

MYB42 amino acid sequence alignment

```
MYB42_ORF       MGRSPCCEKAHTNRGAWTKEEDERLVAYVRAHGEGCWRSLPRAAGLLRCGKSCRLRWINY
MYB42_UTR       MGRSPCCEKAHTNRGAWTKEEDERLVAYVRAHGEGCWRSLPRAAGLLRCGKSCRLRWINY
NM_001112539    MGRSPCCEKAHTNRGAWTKEEDERLVAYVRAHGEGCWRSLPRAAGLLRCGKSCRLRWINY
                ************************************************************

MYB42_ORF       LRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKNYWNTHIRRKLLCSG
MYB42_UTR       LRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKNYWNTHIRRKLLCSG
NM_001112539    LRPDLKRGNFTADEDDLIVKLHSLLGNKWSLIAARLPGRTDNEIKNYWNTHIRRKLLGSG
                ************************************************** ** **

MYB42_ORF       IDPVTHRRVAGGAATTISFQPSPNSAAAAA-AEAAAQAPIKAEETAGVKAPRCPDLNLDL
MYB42_UTR       IDPVTHRRVAGGAATTISFQPSPNSAAAAA-AEAAAQAPIKAEETAGVKAPRCPDLNLDL
NM_001112539    IDPVTHRRVAGGAATTISFQPSPNSAAAAAAAETAAQAPIKAEETAAVKAPRCPDLNLDL
                ****************************** **:************.*************

MYB42_ORF       CISPPCQHEDDGEEEDEELDLKPAFVKREALQAGHGHGHGLCLGCGLGGQKGAAGCSCSN
MYB42_UTR       CISPPCQHEDDGEEEDEELDLKPAFVKREALQAGHGHGHGLCLGCGLGGQKGAAGCSCSN
NM_001112539    CISPPCQHEDDGEEEDEELDLKPAFVKREALQAGHGHGHGLCLGCGLGGQKGAAGCSCSN
                ************************************************************

MYB42_ORF       GHHFLGLRTSVLDFRGLEMK
MYB42_UTR       GHHFLGLRTSVLDFRGLEMK
NM_001112539    GHHFLGLRTSVLDFRGLEMK
                ********************
```

APPENDIX G: Expression levels of lignin biosynthesis genes in sugarcane after MYB31 and MYB42 expression

TABLE G.1: Normalised lignin biosynthesis genes ΔCt values for MYB expressing plants that underwent cell wall compositional analysis. Values normalised against average GFP ΔCt value for each gene. Values represent initial expression screening of leaf tissue and post-harvest expression results from young internode tissue and maturing internode tissue (Table 5.3). NE: Normalised expression with standard error of the mean shown.

		PAL						C4H						C3H					
		leaf		young		maturing		leaf		young		maturing		leaf		young		maturing	
		NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-
WT	n=9	2.27	1.02	1.78	0.25	1.05	0.20	0.94	0.33	0.24	0.09	0.61	0.10	0.47	0.11	0.16	0.04	0.18	0.01
GFP	n=3	1.00	0.34	1.00	0.16	1.00	0.53	1.00	0.06	1.00	0.46	1.00	0.24	1.00	0.30	1.00	0.38	1.00	0.57
MYB31 ORF	13	0.34	0.04	0.34	0.00	0.40	0.00	0.48	0.01	0.04	0.00	0.17	0.01	0.19	0.02	0.46	0.03	0.33	0.02
	11	0.23	0.00	0.96	0.07	0.55	0.01	0.51	0.01	0.11	0.00	0.18	0.00	0.08	0.01	0.29	0.01	0.13	0.00
	2	0.08	0.03	0.08	0.00	0.01	0.00	0.37	0.02	0.37	0.00	0.34	0.01	0.22	0.02	0.40	0.02	0.15	0.01
	7	0.03	0.01	0.93	0.01	0.63	0.00	0.12	0.01	0.06	0.00	0.16	0.00	0.23	0.02	0.20	0.01	0.06	0.00
	1	0.27	0.01	2.68	0.03	0.66	0.01	0.39	0.02	0.21	0.00	0.23	0.01	0.49	0.05	0.04	0.00	0.07	0.00
	8	0.33	0.05	0.20	0.00	0.54	0.00	0.28	0.03	0.04	0.00	0.21	0.01	0.37	0.04	0.04	0.00	0.22	0.01
	9	0.85	0.02	2.34	0.02	0.60	0.05	0.53	0.23	0.07	0.00	0.30	0.00	0.77	0.06	0.16	0.01	0.27	0.03
MYB31 UTR	27	0.63	0.01	0.30	0.00	0.21	0.00	0.41	0.01	0.03	0.00	0.14	0.01	0.04	0.00	0.05	0.00	0.19	0.01
	2	0.33	0.07	0.63	0.00	0.60	0.02	0.35	0.02	0.04	0.00	0.12	0.00	0.21	0.05	0.04	0.00	0.32	0.01
	18	0.53	0.02	1.08	0.01	0.23	0.00	0.36	0.02	0.06	0.00	0.12	0.00	0.06	0.00	0.23	0.01	0.08	0.00
	11	0.73	0.01	1.08	0.00	1.94	0.01	0.28	0.07	0.09	0.00	0.35	0.01	0.11	0.00	0.15	0.00	0.11	0.01
	12	0.80	0.02	1.02	0.00	0.64	0.01	0.50	0.06	0.06	0.00	0.20	0.01	0.04	0.00	0.15	0.02	0.08	0.00
	7	0.55	0.03	8.02	0.61	0.69	0.02	0.40	0.03	0.25	0.01	0.46	0.01	0.22	0.01	0.29	0.01	0.42	0.02
	20	0.45	0.02	5.94	0.06	2.02	0.06	0.44	0.01	0.34	0.02	0.64	0.01	0.13	0.01	0.40	0.02	0.39	0.01

TABLE G.1: Normalised lignin biosynthesis genes ΔCt values for MYB expressing plants that underwent cell wall compositional analysis (continued)

		4CL						CCoAOMT						F5H					
		leaf		young		maturing		leaf		young		maturing		leaf		young		maturing	
		NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-
WT	n=9	0.76	0.18	1.22	0.45	1.22	0.25	0.53	0.09	1.22	0.26	1.34	0.12	0.98	0.19	0.72	0.20	0.83	0.15
GFP	n=3	1.00	0.18	1.00	0.08	1.00	0.33	1.00	0.14	1.00	0.13	1.00	0.26	1.00	0.16	1.00	0.25	1.00	0.36
MYB31 ORF	13	0.51	0.02	0.16	0.01	0.41	0.01	0.29	0.01	0.37	0.01	1.22	0.01	0.57	0.05	0.07	0.00	0.32	0.01
	11	0.40	0.02	0.32	0.02	0.39	0.01	0.26	0.00	0.48	0.01	0.76	0.00	0.38	0.02	0.18	0.00	0.47	0.00
	2	0.47	0.01	0.15	0.01	0.03	0.00	0.40	0.02	1.13	0.03	0.22	0.01	0.83	0.01	0.15	0.00	0.07	0.00
	7	0.14	0.01	0.44	0.01	0.88	0.02	0.20	0.02	0.75	0.02	0.69	0.00	0.43	0.05	0.29	0.01	0.48	0.00
	1	0.39	0.02	1.57	0.03	0.48	0.02	0.37	0.00	1.71	0.06	0.74	0.02	0.82	0.02	1.01	0.01	0.51	0.00
	8	0.25	0.02	0.31	0.01	0.41	0.01	0.26	0.01	0.24	0.01	0.87	0.01	0.91	0.06	0.11	0.00	0.56	0.01
	9	0.67	0.04	0.61	0.03	0.65	0.05	1.00	0.03	1.08	0.08	0.89	0.00	1.02	0.10	0.38	0.00	0.40	0.00
MYB31 UTR	27	1.52	0.05	0.27	0.00	0.13	0.01	0.65	0.01	0.28	0.01	0.59	0.01	0.40	0.02	0.13	0.00	0.20	0.01
	2	1.05	0.05	0.74	0.04	0.53	0.02	0.26	0.01	0.42	0.01	0.57	0.01	1.30	0.11	0.11	0.00	0.25	0.00
	18	0.83	0.02	0.36	0.01	0.31	0.00	0.35	0.02	0.59	0.00	0.28	0.00	0.44	0.01	0.22	0.00	0.34	0.00
	11	0.52	0.03	0.56	0.01	1.54	0.04	0.30	0.02	0.61	0.02	0.84	0.01	0.79	0.02	0.36	0.01	1.17	0.04
	12	0.62	0.02	0.43	0.02	1.55	0.02	0.24	0.02	0.63	0.01	0.67	0.03	0.42	0.01	0.38	0.01	0.40	0.00
	7	0.78	0.04	1.54	0.03	1.26	0.02	0.36	0.01	2.16	0.04	1.28	0.03	1.01	0.07	2.25	0.10	1.68	0.02
	20	0.88	0.07	3.27	0.26	5.42	0.08	0.48	0.00	4.47	0.11	2.67	0.03	0.89	0.03	4.18	0.08	4.35	0.04

TABLE G.1: Normalised lignin biosynthesis genes ΔCt values for MYB expressing plants that underwent cell wall compositional analysis (continued)

		CAD						CCR						COMT					
		leaf		young		maturing		leaf		young		maturing		leaf		young		maturing	
		NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-
WT	n=9	0.48	0.11	0.58	0.23	0.73	0.30	1.05	0.21	1.04	0.21	1.30	0.20	0.60	0.07	1.68	0.28	1.83	0.36
GFP	n=3	1.00	0.15	1.00	0.12	1.00	0.32	1.00	0.18	1.00	0.13	1.00	0.38	1.00	0.23	1.00	0.08	1.00	0.35
MYB31 ORF	13	0.52	0.04	0.52	0.03	1.13	0.08	0.19	0.01	0.32	0.00	0.82	0.01	0.25	0.03	0.50	0.00	0.82	0.01
	11	0.47	0.02	0.64	0.02	0.60	0.03	0.15	0.00	0.50	0.03	0.64	0.01	0.19	0.04	0.54	0.00	0.89	0.07
	2	0.37	0.03	0.98	0.02	0.16	0.01	0.31	0.01	0.03	0.00	0.00	0.00	0.28	0.08	0.16	0.00	0.02	0.00
	7	0.19	0.01	1.20	0.09	1.22	0.11	0.14	0.01	0.71	0.01	1.17	0.07	0.09	0.03	0.47	0.01	0.81	0.01
	1	0.47	0.04	2.53	0.19	1.36	0.10	0.24	0.02	1.01	0.01	0.82	0.01	0.41	0.11	0.95	0.10	0.46	0.00
	8	0.28	0.03	0.52	0.02	1.93	0.02	0.26	0.02	0.29	0.00	1.15	0.00	0.21	0.07	0.33	0.03	0.79	0.08
	9	0.12	0.03	2.55	0.09	1.13	0.07	0.85	0.08	1.06	0.01	0.99	0.01	0.58	0.10	1.00	0.01	0.88	0.00
MYB31 UTR	27	0.97	0.08	0.48	0.04	0.50	0.03	0.61	0.01	0.26	0.00	0.33	0.02	0.21	0.05	0.37	0.01	0.26	0.00
	2	0.14	0.01	0.61	0.02	1.51	0.07	0.33	0.05	0.46	0.04	1.03	0.06	0.50	0.12	0.49	0.01	0.78	0.00
	18	0.77	0.01	1.32	0.07	0.78	0.02	0.41	0.02	0.77	0.01	0.35	0.01	0.17	0.05	0.88	0.01	0.56	0.00
	11	0.77	0.02	1.94	0.05	2.44	0.06	0.42	0.01	0.47	0.04	1.25	0.08	0.58	0.07	1.00	0.01	1.57	0.01
	12	0.65	0.02	1.15	0.01	1.14	0.07	0.31	0.02	0.79	0.06	1.03	0.02	0.38	0.07	0.79	0.01	1.12	0.01
	7	0.43	0.03	5.40	0.26	3.36	0.09	0.94	0.07	1.58	0.01	1.25	0.01	0.24	0.08	3.20	0.48	1.76	0.01
	20	0.97	0.04	6.39	0.17	6.67	0.25	0.69	0.01	3.04	0.02	5.47	0.38	0.46	0.07	4.39	0.04	3.35	0.31

TABLE G.1: Normalised lignin biosynthesis genes ΔCt values for MYB expressing plants that underwent cell wall compositional analysis (continued)

		PAL						C4H						C3H					
		leaf		young		maturing		leaf		young		maturing		leaf		young		maturing	
		NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-
WT	n=9	2.27	1.02	1.78	0.25	1.05	0.20	0.94	0.33	0.24	0.09	0.61	0.10	0.47	0.11	0.16	0.04	0.18	0.01
GFP	n=3	1.00	0.34	1.00	0.16	1.00	0.53	1.00	0.06	1.00	0.46	1.00	0.24	1.00	0.30	1.00	0.38	1.00	0.57
MYB42 ORF	14	0.76	0.10	1.31	0.01	0.09	0.00	0.24	0.08	0.06	0.00	0.07	0.01	0.23	0.01	0.20	0.02	0.45	0.02
	16	0.61	0.01	0.85	0.01	1.28	0.02	0.20	0.00	0.04	0.01	0.44	0.03	0.13	0.02	0.56	0.05	0.27	0.01
	23	0.34	0.01	1.41	0.01	0.73	0.02	0.18	0.00	0.18	0.01	0.42	0.01	0.08	0.01	0.36	0.03	0.30	0.02
	11	0.64	0.01	3.31	0.03	0.17	0.02	0.69	0.02	0.42	0.01	0.24	0.01	0.13	0.01	0.14	0.00	0.11	0.00
	18	0.41	0.04	1.91	0.01	1.14	0.01	0.64	0.01	0.33	0.03	0.29	0.02	0.37	0.01	0.46	0.01	0.52	0.01
	21	0.65	0.01	2.46	0.01	1.00	0.01	0.87	0.11	0.21	0.00	0.48	0.04	0.26	0.04	0.26	0.00	0.35	0.02
	26	1.06	0.10	1.94	0.17	1.07	0.01	0.65	0.09	0.15	0.00	0.29	0.01	0.21	0.03	0.12	0.00	0.12	0.00
MYB42 UTR	28	0.22	0.01	2.65	0.02	0.94	0.01	0.47	0.03	0.14	0.01	0.27	0.01	0.12	0.01	0.39	0.01	0.29	0.02
	6	0.09	0.03	0.40	0.00	0.53	0.00	0.11	0.00	0.09	0.00	0.30	0.01	0.14	0.01	0.75	0.04	0.25	0.02
	32	0.82	0.09	3.88	0.06	0.55	0.03	0.39	0.07	0.30	0.02	0.18	0.00	0.34	0.02	0.13	0.01	0.08	0.00
	30	0.32	0.04	2.34	0.04	0.40	0.01	0.29	0.02	0.18	0.01	0.30	0.00	0.16	0.00	0.39	0.02	0.18	0.00
	15	0.72	0.01	6.26	0.61	0.79	0.05	0.30	0.01	0.35	0.03	0.30	0.02	0.17	0.01	0.17	0.00	0.15	0.01
	26	0.70	0.10	1.55	0.02	2.28	0.03	0.47	0.02	0.05	0.00	0.16	0.00	0.15	0.02	0.06	0.00	0.14	0.00
	16	0.46	0.01	0.87	0.01	1.73	0.02	0.56	0.01	0.05	0.00	0.63	0.02	0.23	0.03	0.15	0.01	0.15	0.01

TABLE G.1: Normalised lignin biosynthesis genes ΔCt values for MYB expressing plants that underwent cell wall compositional analysis (continued)

		4CL						CCoAOMT						F5H					
		leaf		young		maturing		leaf		young		maturing		leaf		young		maturing	
		NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-
WT	n=9	0.76	0.18	1.22	0.45	1.22	0.25	0.53	0.09	1.22	0.26	1.34	0.12	0.98	0.19	0.72	0.20	0.83	0.15
GFP	n=3	1.00	0.18	1.00	0.08	1.00	0.33	1.00	0.14	1.00	0.13	1.00	0.26	1.00	0.16	1.00	0.25	1.00	0.36
MYB42 ORF	14	2.61	0.06	0.25	0.01	0.18	0.01	1.27	0.02	1.05	0.01	0.43	0.01	3.57	0.34	0.27	0.01	0.31	0.01
	16	1.50	0.04	0.41	0.05	0.17	0.01	0.72	0.05	0.52	0.01	1.02	0.02	0.71	0.05	0.34	0.02	0.24	0.00
	23	2.06	0.07	0.39	0.03	0.41	0.01	0.73	0.00	1.42	0.03	1.41	0.03	0.37	0.04	0.32	0.01	0.42	0.01
	11	5.95	2.55	0.28	0.03	0.29	0.01	2.87	0.01	0.75	0.03	0.47	0.01	0.39	0.04	1.35	0.03	1.01	0.03
	18	3.45	0.08	0.41	0.01	0.90	0.03	0.61	0.03	0.96	0.01	1.90	0.01	0.77	0.10	0.47	0.02	0.91	0.04
	21	2.36	0.05	1.00	0.02	1.21	0.02	1.01	0.01	2.06	0.04	1.21	0.01	0.84	0.12	0.32	0.01	1.46	0.01
	26	1.93	0.13	1.46	0.07	0.69	0.01	0.90	0.01	0.74	0.01	0.61	0.01	0.42	0.07	0.42	0.01	0.57	0.01
MYB42 UTR	28	3.06	0.12	0.82	0.08	0.34	0.03	0.47	0.03	1.72	0.18	1.06	0.02	3.84	0.59	0.56	0.02	0.45	0.00
	6	0.80	0.02	0.71	0.03	0.53	0.02	0.54	0.00	0.39	0.03	0.93	0.02	0.86	0.04	0.59	0.01	0.34	0.00
	32	6.69	0.03	0.57	0.04	0.20	0.01	1.50	0.05	3.30	0.09	1.09	0.02	2.11	0.23	1.35	0.07	1.06	0.01
	30	2.46	0.21	1.45	0.10	0.11	0.01	0.69	0.00	1.61	0.02	0.49	0.01	0.48	0.07	0.63	0.01	0.21	0.01
	15	1.38	0.02	1.72	0.14	1.14	0.03	0.90	0.03	2.32	0.11	0.87	0.01	0.43	0.01	2.11	0.02	1.29	0.01
	26	3.15	0.27	0.39	0.02	1.15	0.07	0.61	0.06	0.67	0.03	1.21	0.03	5.39	0.30	0.60	0.03	1.99	0.02
	16	1.54	0.15	0.36	0.01	2.14	0.08	1.11	0.02	0.61	0.02	1.65	0.08	0.50	0.02	0.25	0.00	1.61	0.03

TABLE G.1: Normalised lignin biosynthesis genes ΔCt values for MYB expressing plants that underwent cell wall compositional analysis (continued)

		CAD						CCR						COMT					
		leaf		young		maturing		leaf		young		maturing		leaf		young		maturing	
		NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-	NE	+/-
WT	n=9	0.48	0.11	0.58	0.23	0.73	0.30	1.05	0.21	1.04	0.21	1.30	0.20	0.60	0.07	1.68	0.28	1.83	0.36
GFP	n=3	1.00	0.15	1.00	0.12	1.00	0.32	1.00	0.18	1.00	0.13	1.00	0.38	1.00	0.23	1.00	0.08	1.00	0.35
MYB42 ORF	14	0.87	0.16	1.06	0.03	0.33	0.01	1.24	0.01	0.73	0.00	0.30	0.01	0.13	0.02	0.79	0.01	0.20	0.00
	16	0.48	0.03	0.86	0.03	0.75	0.05	0.56	0.02	0.43	0.00	0.87	0.01	0.18	0.04	0.68	0.00	0.96	0.01
	23	0.39	0.02	1.28	0.02	1.39	0.02	0.89	0.03	0.90	0.02	1.24	0.02	0.28	0.06	0.96	0.00	0.97	0.01
	11	0.93	0.08	3.16	0.15	1.17	0.08	1.70	0.06	0.60	0.03	0.30	0.00	0.52	0.11	1.61	0.18	0.41	0.06
	18	0.28	0.03	2.30	0.12	1.93	0.11	0.76	0.02	1.59	0.09	1.71	0.03	0.41	0.08	1.51	0.03	0.99	0.01
	21	0.40	0.04	2.39	0.11	1.79	0.17	0.93	0.11	1.00	0.02	0.74	0.01	0.53	0.14	1.82	0.01	1.11	0.10
	26	0.24	0.01	2.70	0.14	1.61	0.07	1.54	0.09	1.21	0.08	1.55	0.01	0.65	0.13	1.33	0.02	0.97	0.00
MYB42 UTR	28	0.07	0.02	1.19	0.07	1.74	0.16	0.89	0.06	0.62	0.01	0.99	0.01	0.21	0.05	1.02	0.01	1.99	0.01
	6	0.17	0.01	0.80	0.07	0.87	0.10	0.28	0.01	0.35	0.00	1.22	0.07	0.18	0.05	0.54	0.00	0.72	0.00
	32	0.12	0.01	2.82	0.03	0.92	0.06	1.70	0.19	2.04	0.22	0.35	0.01	0.23	0.02	1.33	0.01	0.44	0.01
	30	0.13	0.01	1.14	0.03	0.61	0.01	0.50	0.02	1.27	0.02	0.39	0.03	0.73	0.31	1.03	0.01	0.45	0.00
	15	0.57	0.05	4.31	0.18	1.99	0.06	0.96	0.01	2.63	0.01	1.17	0.02	0.33	0.11	2.48	0.01	1.10	0.00
	26	0.13	0.05	2.12	0.14	2.70	0.16	0.93	0.04	0.75	0.01	1.55	0.01	0.30	0.05	0.83	0.00	1.24	0.11
	16	0.56	0.03	2.06	0.14	3.19	0.15	1.10	0.10	0.37	0.03	2.48	0.06	0.40	0.07	0.55	0.05	1.39	0.02

APPENDIX H: Cellulose crystallinity index of MYB bagasse

Plant		Crystallinity Index (%)
WT	90	50.41
WT	113	51.81
WT	130	46.96
WT	134	49.40
GFP	8	49.03
GFP	10	51.49
MYB31 ORF	27	45.56
	11	47.50
MYB 31 UTR	18	45.60
MYB42 ORF	23	47.41
MYB42 UTR	32	45.97

APPENDIX I: Phenotypic measurements of MYB31 and MYB42 expressing sugarcane

TABLE I.1: Phenotypic measurements of MYB31 ORF and UTR expressing sugarcane. Overall averages for wild type (WT) (n = 9 individual plants) and GFP (n = 3 individual plants) control measurements are presented with standard deviation. *Z* scores represent the number of standard deviations each MYB plant measurement is from the GFP control average, with *z* scores greater than 2 or -2 highlighted in bold font. Plants are listed in ascending order of total lignin content.

Plant		**Height (cm)**		**Total number of internodes**		**Average internode diameter (mm)**		**Average internode length (cm)**	
	WT	145.89 ± 15.08		14.22 ± 1.62		14.36 ± 1.16		10.29 ± 0.79	
	GFP	129 ± 24.26		15 ± 2.16		14.11 ± 0.85		8.56 ± 0.60	
			***z* score**		***z* score**		***z* score**		***z* score**
MYB31 ORF	13	114	-0.62	14	-0.46	12.29	**-2.15**	8.14	-0.70
	11	132	0.12	13	-0.93	13.64	-0.56	10.15	**2.63**
	2	100	-1.20	10	**-2.31**	12.13	**-2.34**	10.00	**2.38**
	7	205	**3.13**	15	0.00	14.05	-0.08	13.67	**8.45**
	1	140	0.45	14	-0.46	10.18	**-4.64**	10.00	**2.38**
	8	158	1.20	14	-0.46	12.59	-1.80	11.29	**4.51**
	9	159	1.24	15	0.00	11.48	**-3.10**	10.60	**3.37**
			***z* score**		***z* score**		***z* score**		***z* score**
MYB31 UTR	27	140	0.45	16	0.46	10.63	**-4.11**	8.75	0.31
	2	128	-0.04	13	-0.93	12.15	**-2.32**	9.85	**2.12**
	18	131	0.08	13	-0.93	13.86	-0.30	10.08	**2.51**
	11	160	1.28	15	0.00	12.90	-1.43	10.67	**3.48**
	12	147	0.74	14	-0.46	11.98	**-2.52**	10.50	**3.21**
	7	136	0.29	15	0.00	14.91	0.94	9.07	0.83
	20	108	-0.87	15	0.00	13.20	-1.08	7.20	**-2.26**

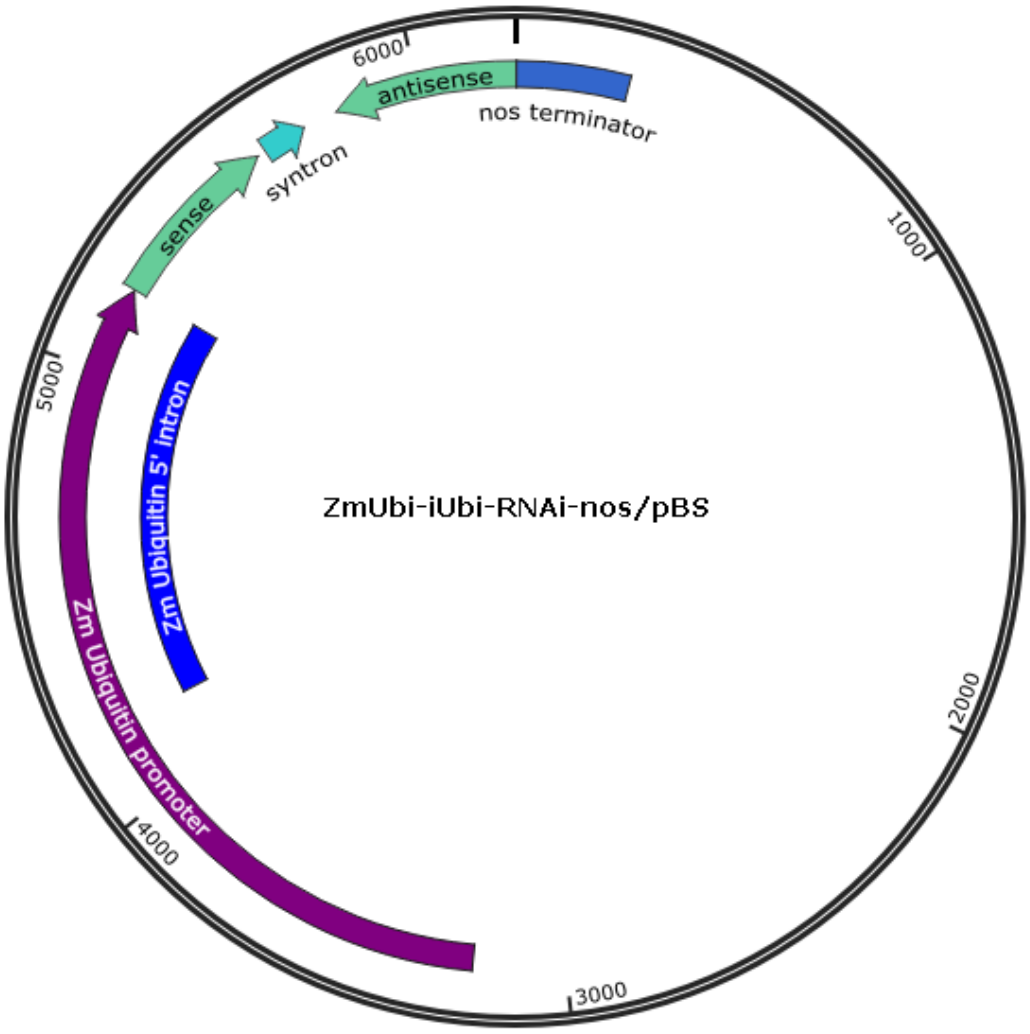
TABLE I.2: Phenotypic measurements of MYB42 ORF and UTR expressing sugarcane. Overall averages for wild type (WT) (n = 9 individual plants) and GFP (n = 3 individual plants) control measurements are presented with standard deviation. *Z* scores represent the number of standard deviations each MYB plant measurement is from the GFP control average, with *z* scores greater than 2 or -2 highlighted in bold font. Plants are listed in ascending order of total lignin content.

Plant		**Height (cm)**		**Total number of internodes**		**Average internode diameter (mm)**		**Average internode length (cm)**	
	WT	145.89 ± 15.08		14.22 ± 1.62		14.36 ± 1.16		10.29 ± 0.79	
	GFP	129 ± 24.26		15 ± 2.16		14.11 ± 0.85		8.56 ± 0.60	
			***z* score**		***z* score**		***z* score**		***z* score**
MYB42 ORF	14	162	1.36	17	0.93	12.77	-1.58	9.53	1.60
	16	130	0.04	17	0.93	16.15	**2.40**	7.65	-1.52
	23	134	0.21	15	0.00	14.00	-0.13	8.93	0.61
	11	138	0.37	15	0.00	10.97	**-3.70**	9.20	1.05
	18	124	-0.21	14	-0.46	12.27	**-2.17**	8.86	0.49
	21	163	1.40	16	0.46	12.67	-1.70	10.19	**2.69**
	26	145	0.66	13	-0.93	14.56	0.52	11.15	**4.29**
			***z* score**		***z* score**		***z* score**		***z* score**
MYB42 UTR	28	92	-1.52	12	-1.39	14.07	-0.05	7.67	-1.49
	6	109	-0.82	15	0.00	11.73	**-2.81**	7.27	**-2.15**
	32	137	0.33	14	-0.46	15.04	1.09	9.79	**2.02**
	30	71	**-2.39**	11	-1.85	12.25	**-2.20**	6.45	**-3.50**
	15	147	0.74	15	0.00	16.10	**2.34**	9.80	**2.05**
	26	174	1.85	18	1.39	14.52	0.48	9.67	1.83
	16	162	1.36	17	0.93	12.72	-1.64	9.53	1.60

APPENDIX J: Generic RNAi expression vector map (*Zm*Ubi-iUbi-sense/syntron/antisense-nos/pBS) showing positions of maize Ubiquitin promoter (Ubi) and 5' UTR intron (iUbi) (Christensen and Quail, 1996; Christensen *et al.*, 1992), syntron and the nopaline synthase (nos) terminator (Bevan *et al.*, 1983a). Individual sense and antisense sequences for CCoAOMT, F5H and COMT are underlined in Appendix B and were cloned into the *sense* and *antisense* sites of the expression vector respectively.

Sequence of synthetic intron (syntron) (5' - 3')

CTGCAAGAAAACAAAATAAAAAATAAAAAGTTAGATTTAAATGGCGCGCCGATCGAGAATTCAACA
GATCGAATTAATTAAATAAAAAATAAAAATCTTAC



APPENDIX K: Normalised qPCR ΔCt quantified gene expression levels of RNAi targeted lignin biosynthetic genes. Values represent initial screening of leaf tissue and post-harvest expression results from young internode tissue and maturing internode tissue. All data normalised against UKN transgenic controls with standard error of the mean shown. Plants are listed in ascending order of expression in maturing internode tissue then ascending order of expression in leaf tissue of plants not selected for further analysis. NE: Normalised expression

TABLE K.1: Normalised qPCR expression levels for CCoAOMT in CCoAOMT-RNAi targeted plants

		Leaf		**Young**		**Maturing**	
Plant		NE	+/-	NE	+/-	NE	+/-
WT		1.41	0.18	1.14	0.22	1.15	0.33
UKN		1.00	0.10	1.00	0.14	1.00	0.37
CCoAOMT	9	1.03	0.01	0.98	0.02	0.03	0.00
	10	0.31	0.00	0.98	0.01	0.09	0.00
	5	0.20	0.00	0.54	0.01	0.14	0.00
	11	0.95	0.02	1.00	0.01	0.17	0.00
	1	1.44	0.17	1.05	0.02	0.30	0.01
	13	1.69	0.06	1.07	0.03	0.39	0.00
	7	0.69	0.01	1.02	0.03	1.53	0.10
	2	0.63	0.01	0.55	0.02	1.76	0.03
	8	0.56	0.02	1.73	0.04	1.89	0.03
	6	0.48	0.01				
	12	0.57	0.01				
	4	3.13	0.01				

TABLE K.2: Normalised qPCR expression levels for F5H in F5H-RNAi targeted plants

		Leaf		Young		Maturing	
Plant		NE	+/-	NE	+/-	NE	+/-
WT		1.70	0.61	1.33	0.37	0.78	0.22
UKN		1.00	0.10	1.00	0.14	1.00	0.35
F5H	7	0.76	0.05	5.40	0.81	0.16	0.00
	2	1.23	0.20	0.83	0.02	0.17	0.00
	4	0.74	0.03	0.85	0.01	0.29	0.04
	1	1.06	0.08	0.84	0.01	0.96	0.01
	13	0.79	0.12	0.74	0.04	1.12	0.06
	3	0.72	0.03	0.77	0.04	1.23	0.02
	14	1.29	0.05	1.33	0.02	1.98	0.04
	6	0.93	0.14	1.35	0.02	2.18	0.02
	8	0.57	0.09	1.77	0.07	4.75	0.05
	10	0.37	0.02				
	11	0.52	0.05				
	9	0.86	0.04				
	15	0.90	0.11				
	12	0.95	0.12				
	5	1.41	0.05				

TABLE K.3: Normalised qPCR expression levels for COMT in COMT-RNAi targeted plants

		Leaf		Young		Maturing	
Plant		NE	+/-	NE	+/-	NE	+/-
WT		1.49	0.09	1.35	0.54	0.98	0.32
UKN		1.00	0.10	1.00	0.41	1.00	0.46
COMT	4	1.37	0.02	3.82	0.12	0.68	0.01
	10	2.33	0.13	1.80	0.03	0.79	0.03
	3	1.19	0.05	0.88	0.01	1.26	0.02
	2	0.41	0.06	0.98	0.02	4.01	0.11
	13	3.00	0.07	14.06	0.48	4.23	0.12
	8	2.09	0.12	2.25	0.03	5.27	0.17
	14	2.92	0.06	7.72	0.17	6.59	0.11
	7	1.61	0.06	4.40	0.08	10.90	0.07
	6	0.49	0.06				
	11	3.04	0.13				
	5	3.45	0.14				
	12	3.89	0.10				
	1	5.04	0.14				

APPENDIX L: Normalised ΔCt expression levels of genes related to RNAi targeted lignin biosynthetic genes to assess specificity of RNAi vectors**.** Values represent post-harvest expression results from young internode tissue and maturing internode tissue. All data normalised against UKN transgenic controls with standard error of the mean shown. Plants are listed in ascending order of expression of the targeted RNAi gene in maturing internode (Appendix K). NE: Normalised expression

TABLE L.1: Normalised qPCR expression levels for COMT in CCoAOMT-RNAi targeted plants

		Young tissue		Maturing tissue	
Plant		NE	+/-	NE	+/-
WT		1.35	0.54	0.98	0.32
UKN		1.00	0.41	1.00	0.46
CCoAOMT	9	0.58	0.02	0.74	0.02
	10	1.60	0.04	3.49	0.06
	5	5.82	0.13	7.18	0.21
	11	4.99	0.09	0.83	0.02
	1	3.31	0.06	1.90	0.02
	13	3.09	0.02	3.42	0.12
	7	0.52	0.01	7.84	0.13
	2	0.95	0.01	7.87	0.06
	8	6.52	0.09	4.09	0.16

TABLE L.2: Normalised qPCR expression levels for C3H in F5H-RNAi targeted plants

		Young tissue		Maturing tissue	
Plant		NE	+/-	NE	+/-
WT		1.33	0.28	0.91	0.08
UKN		1.00	0.22	1.00	0.15
F5H	7	0.70	0.03	1.73	0.02
	2	1.15	0.05	1.64	0.07
	4	0.75	0.02	1.20	0.01
	1	1.41	0.03	2.42	0.24
	13	1.64	0.02	1.95	0.09
	3	0.73	0.06	1.56	0.06
	14	0.42	0.02	1.65	0.13
	6	1.04	0.01	1.94	0.11
	8	0.96	0.01	4.18	0.16

TABLE L.3: Normalised qPCR expression levels for CCoAOMT in COMT-RNAi targeted plants

		Young tissue		Maturing tissue	
Plant		NE	+/-	NE	+/-
WT		1.14	0.22	1.15	0.33
UKN		1.00	0.14	1.00	0.37
COMT	4	1.92	0.14	1.39	0.04
	10	4.58	0.03	3.56	0.22
	3	2.85	0.02	7.49	0.51
	2	5.17	0.16	11.62	0.93
	13	6.97	0.25	1.30	0.02
	8	1.09	0.01	1.67	0.03
	14	5.01	0.39	3.69	0.28
	7	4.24	0.31	1.45	0.00

APPENDIX M: Cellulose crystallinity index of RNAi bagasse

Plant		Crystallinity Index (%)
WT	23	48.21
UKN	4	53.51
CCoAOMT	9	53.89
	10	57.86
F5H	2	55.62
COMT	10	54.71

APPENDIX N

TABLE N.1: Phenotypic measurements of CCoAOMT, F5H and COMT-RNAi sugarcane. Overall averages for wild type (WT) (n = 7 individual plants) and UKN (n = 6 individual plants) control measurements are presented with standard deviation. *Z* scores represent the number of standard deviations each RNAi plant measurement is from the UKN control average, with *z* scores greater than 2 or -2 highlighted in bold font. Plants are listed in ascending order of total lignin content.

Plant		Height (cm)		Total number of internodes		Average internode diameter (mm)		Average internode length (cm)	
	WT	170.29 ± 10.47		20.43 ± 1.05		12.56 ± 0.49		8.35 ± 0.48	
	UKN	155.17 ± 28.23		19.67 ± 1.89		12.48 ± 0.94		7.89 ± 1.20	
			***z* score**		***z* score**		***z* score**		***z* score**
CCoAOMT	11	163	0.28	21	0.71	12.36	-0.13	7.76	-0.10
	5	80	**-2.66**	20	0.18	10.68	-1.92	4.00	**-3.23**
	10	150	-0.18	18	-0.88	12.95	0.50	8.33	0.37
	9	105	-1.78	17	-1.41	12.61	0.14	6.18	-1.42
F5H	4	131	-0.86	24	**2.30**	11.36	-1.19	5.46	**-2.02**
	2	130	-0.89	18	-0.88	13.07	0.62	7.22	-0.55
	7	104	-1.81	16	-1.94	11.65	-0.89	6.50	-1.15
	1	151	-0.15	22	1.24	11.80	-0.72	6.86	-0.85
COMT	2	107	-1.71	22	1.24	11.75	-0.78	4.86	**-2.51**
	10	113	-1.49	16	-1.94	13.49	1.07	7.06	-0.68
	3	172	0.60	20	0.18	14.36	**2.00**	8.60	0.59
	4	139	-0.57	19	-0.35	14.41	**2.04**	7.32	-0.47

Literature Cited

Adney, B. and Baker, J. (2008). Measurement of cellulase activities: Laboratory analytical procedure (LAP) *National Renewable Energy Laboratory,* Technical report: NREL/TP-510-42628.

Alexandrov, N. N., Brover, V. V., Freidin, S., Troukhan, M. E., Tatarinova, T. V., Zhang, H., Swaller, T. J., Lu, Y., Bouck, J., Flavell, R. B. and Feldmann, K. A. (2009). Insights into corn genes derived from large-scale cDNA sequencing. *Plant Molecular Biology,* 69: 179-194.

Ambawat, S., Sharma, P., Yadav, N. R. and Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: An overview. *Physiology and Molecular Biology of Plants,* 19: 307-321.

Andersen, J. R., Zein, I., Wenzel, G., Darnhofer, B., Eder, J., Ouzunova, M. and Lübberstedt, T. (2008). Characterization of phenylpropanoid pathway genes within European maize (*Zea mays* L.) inbreds. *BMC Plant Biology,* 8: 2.

Anterola, A. M., Jeon, J., Davin, L. B. and Lewis, N. G. (2002). Transcriptional control of monolignol biosynthesis in *Pinus taeda*: Factors affecting monolignol ratios and carbon allocation in phenylpropanoid metabolism. *The Journal of Biological Chemistry,* 277: 18272-18280.

Barrière, Y., Ralph, J., Méchin, V., Guillaumie, S., Grabber, J. H., Argillier, O., Chabbert, B. and Lapierre, C. (2004). Genetic and molecular basis of grass cell wall biosynthesis and degradability. II. Lessons from brown-midrib mutants. *Comptes Rendus Biologies,* 327: 847-860.

Barrière, Y., Riboulet, C., Méchin, V., Maltese, S., Pichon, M., Cardinal, A., Lapierre, C., Lübberstedt, T. and Martinant, J. (2007). Genetics and genomics of lignification in grass cell walls based on maize as model species. *Genes, Genomes and Genomics,* 1: 133-156.

Baucher, M., Halpin, C., Petit-Conil, M. and Boerjan, W. (2003). Lignin: Genetic engineering and impact on pulping. *Critical Reviews in Biochemistry and Molecular Biology,* 38: 305-350.

Baumberger, N. and Baulcombe, D. C. (2005). *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proceedings of the National Academy of Sciences of the United States of America,* 102: 11928-11933.

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

Baxter, H. L. and Stewart, C. N. (2013). Effects of altered lignin biosynthesis on phenylpropanoid metabolism and plant stress. *Biofuels,* 4: 635-650.

Bedon, F., Grima-Pettenati, J. and Mackay, J. (2007). Conifer R2R3-MYB transcription factors: Sequence analyses and gene expression in wood-forming tissues of white spruce (*Picea glauca*). *BMC Plant Biology,* 7: 17.

Benjamin, Y., Cheng, H. and Görgens, J. F. (2013). Evaluation of bagasse from different varieties of sugarcane by dilute acid pretreatment and enzymatic hydrolysis. *Industrial Crops and Products,* 51: 7-18.

Benjamin, Y., Görgens, J. F. and Joshi, S. V. (2014). Comparison of chemical composition and calculated ethanol yields of sugarcane varieties harvested for two growing seasons. *Industrial Crops and Products,* 58: 133-141.

Bernstein, E., Caudy, A. A., Hammond, S. M. and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature,* 409: 363-366.

Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *Journal of Bacteriology,* 62: 293-300.

Besseau, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B. and Legrand, M. (2007). Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. *The Plant Cell,* 19: 148-162.

Bevan, M., Barnes, W. M. and Chilton, M. (1983a). Structure and transcription of the nopaline synthase gene region of T-DNA. *Nucleic Acids Research,* 11: 369-385.

Bevan, M. W., Flavell, R. B. and Chilton, M. (1983b). A chimaeric antibiotic resistance gene as a selectable marker for plant cell transformation. *Nature,* 304: 184-187.

Bhinu, V., Li, R., Huang, J., Kaminskyj, S., Sharpe, A. and Hannoufa, A. (2009). Perturbation of lignin biosynthesis pathway in *Brassica napus* (canola) plants using RNAi. *Canadian Journal of Plant Science,* 89: 441-453.

Bi, C., Chen, F., Jackson, L., Gill, B. S. and Li, W. (2011). Expression of lignin biosynthetic genes in wheat during development and upon infection by fungal pathogens. *Plant Molecular Biology Reporter,* 29: 149-161.

Bimboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research,* 7: 1513-1523.

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

Bohula, E. A., Salisbury, A. J., Sohail, M., Playford, M. P., Riedemann, J., Southern, E. M. and Macaulay, V. M. (2003). The efficacy of small interfering RNAs targeted to the type 1 insulin-like growth factor receptor (IGF1R) is influenced by secondary structure in the IGF1R transcript. *The Journal of Biological Chemistry*, 278: 15991–15997.

Bomal, C., Bedon, F., Caron, S., Mansfield, S. D., Levasseur, C., Cooke, J. E. K., Blais, S., Tremblay, L., Morency, M., Pavy, N., Grima-Pettenati, J., Séguin, A. and Mackay, J. (2008). Involvement of *Pinus taeda MYB1* and *MYB8* in phenylpropanoid metabolism and secondary cell wall biogenesis: A comparative *in planta* analysis. *Journal of Experimental Botany,* 59: 3925-3939.

Bonawitz, N. D. and Chapple, C. (2010). The genetics of lignin biosynthesis: Connecting genotype to phenotype. *Annual Review of Genetics,* 44: 337-363.

Bonawitz, N. D. and Chapple, C. (2013). Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Current Opinion in Biotechnology,* 24: 336-343.

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A. and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *The Plant Cell,* 12: 2383-2393.

Borrion, A. L., McManus, M. C. and Hammond, G. P. (2012). Environmental life cycle assessment of lignocellulosic conversion to ethanol: A review. *Renewable and Sustainable Energy Reviews,* 16: 4638-4650.

Botha, F. C. and Moore, P. H. (2014). *Biomass and bioenergy*, in Moore, P. H. and Botha, F. C. (Eds.). *Sugarcane: Physiology, biochemistry and functional biology*. Somerset, NJ: Wiley-Blackwell, p. 521-540.

Bourquin, V., Nishikubo, N., Abe, H., Brumer, H., Denman, S., Eklund, M., Christiernin, M., Teeri, T. T., Sundberg, B. and Mellerowicz, E. J. (2002). Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues. *The Plant Cell,* 14: 3073-3088.

Boutet, S., Vazquez, F., Liu, J., Béclin, C., Fagard, M., Gratias, A., Morel, J., Crété, P., Chen, X. and Vaucheret, H. (2003). *Arabidopsis HEN1*: A genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. *Current Biology,* 13: 843-848.

Bower, R. and Birch, R. G. (1992). Transgenic sugarcane plants via microprojectile bombardment. *The Plant Journal,* 2: 409–416.

Bower, R., Elliott, A. R., Potier, B. A. M. and Birch, R. G. (1996). High-efficiency, microprojectile-mediated cotransformation of sugarcane, using visible or selectable markers. *Molecular Breeding,* 2: 239-249.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* 72: 248-254.

Buanafina, M. M. de O. (2009). Feruloylation in grasses : Current and future perspectives. *Molecular Plant,* 2: 861-872.

Canegrowers (2010). Canegrowers is the peak representative body for Australian sugarcane growers. Available online: www.canegrowers.com.au Last accessed 14/10/14.

Canilha, L., Chandel, A. K., Milessi, T. S. dos S., Antunes, F. A. F., Freitas, W. L. da C., Felipe, M. das G. A. and da Silva, S. S. (2012). Bioconversion of sugarcane biomass into ethanol: An overview about composition, pretreatment methods, detoxification of hydrolysates, enzymatic saccharification, and ethanol fermentation. *Journal of Biomedicine and Biotechnology,* 2012: 1-15.

Capellades, M., Torres, M. A., Bastisch, I., Stiefel, V., Vignols, F., Bruce, W. B., Peterson, D., Puigdomènech, P. and Rigau, J. (1996). The maize caffeic acid *O*-methyltransferase gene promoter is active in transgenic tobacco and maize plant tissues. *Plant Molecular Biology,* 31: 307-322.

Casu, R. E., Dimmock, C. M., Chapman, S. C., Grof, C. P. L., McIntyre, C. L., Bonnett, G. D. and Manners, J. M. (2004). Identification of differentially expressed transcripts from maturing stem of sugarcane by *in silico* analysis of stem expressed sequence tags and gene expression profiling. *Plant Molecular Biology,* 54: 503-517.

Casu, R. E., Jarmey, J. M., Bonnett, G. D. and Manners, J. M. (2007). Identification of transcripts associated with cell wall metabolism and development in the stem of sugarcane by Affymetrix GeneChip Sugarcane Genome Array expression profiling. *Functional and Integrative Genomics,* 7: 153-167.

Century, K., Reuber, T. L. and Ratcliffe, O. J. (2008). Regulating the regulators: The future prospects for transcription-factor-based agricultural biotechnology products. *Plant Physiology,* 147: 20-29.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science,* 263: 802-805.

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

Chen, F., Reddy, M. S. S., Temple, S., Jackson, L., Shadle, G. and Dixon, R. A. (2006). Multi-site genetic modulation of monolignol biosynthesis suggests new routes for formation of syringyl lignin and wall-bound ferulic acid in alfalfa (*Medicago sativa* L.). *The Plant Journal,* 48: 113-124.

Chen, L., Auh, C., Dowling, P., Bell, J., Lehmann, D. and Wang, Z. (2004). Transgenic down-regulation of caffeic acid *O*-methyltransferase (COMT) led to improved digestibility in tall fescue (*Festuca arundinacea*). *Functional Plant Biology,* 31: 235-245.

Chen, Y., Zein, I., Brenner, E. A., Andersen, J. R., Landbeck, M., Ouzunova, M. and Lübberstedt, T. (2010). Polymorphisms in monolignol biosynthetic genes are associated with biomass yield and agronomic traits in European maize (*Zea mays* L.). *BMC Plant Biology,* 10: 12.

Cheng, H., Li, L., Xu, F., Cheng, S., Cao, F., Wang, Y., Yuan, H., Jiang, D. and Wu, C. (2013). Expression patterns of a cinnamyl alcohol dehydrogenase gene involved in lignin biosynthesis and environmental stress in *Ginkgo biloba*. *Molecular Biology Reports,* 40: 707-721.

Chiu, W., Niwa, Y., Zeng, W., Hirano, T., Kobayashi, H. and Sheen, J. (1996). Engineered GFP as a vital reporter in plants. *Current Biology,* 6: 325-330.

Chong, B. F. and O'Shea, M. G. (2012). Developing sugarcane lignocellulosic biorefineries: Opportunities and challenges. *Biofuels,* 3: 307-319.

Chopra, S., Cocciolone, S. M., Bushman, S., Sangar, V., McMullen, M. D. and Peterson, T. (2003). The maize *Unstable factor for orange1* is a dominant epigenetic modifier of a tissue specifically silent allele of *pericarp color1*. *Genetics,* 163: 1135-1146.

Christensen, A. H. and Quail, P. H. (1996). Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research,* 5: 213-218.

Christensen, A. H., Sharrock, R. A. and Quail, P. H. (1992). Maize polyubiquitin genes: Structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Molecular Biology,* 18: 675-689.

Civardi, L., Rigau, J. and Puigdomenech, P. (1999). Nucleotide sequence of two cDNAs coding for caffeoyl-Coenzyme A *O*-methyltransferase and their expression in maize (Accession nos. AJ242980 and AJ242981) (PGR 99–113). *Plant Physiology,* 120: 1205-1207.

Colbére-Garapin, F., Horodniceanu, F., Kourilsky, P. and Garapin, A. (1981). A new dominant hybrid selective marker for higher eukaryotic cells. *Journal of Molecular Biology,* 150: 1-14.

Coleman, H. D., Park, J., Nair, R., Chapple, C. and Mansfield, S. D. (2008). RNAi-mediated suppression of *p*-coumaroyl-CoA 3'-hydroxylase in hybrid poplar impacts lignin deposition and soluble secondary metabolism. *Proceedings of the National Academy of Sciences of the United States of America,* 105: 4501-4506.

Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology,* 6: 850-861.

Craven-Bartle, B., Pascual, M. B., Cánovas, F. M. and Avila, C. (2013). A Myb transcription factor regulates genes of the phenylalanine pathway in maritime pine. *The Plant Journal,* 74: 755-766.

Dallas, P. B., Gottardo, N. G., Firth, M. J., Beesley, A. H., Hoffmann, K., Terry, P. A., Freitas, J. R., Boag, J. M., Cummings, A. J. and Kees, U. R. (2005). Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR – how well do they correlate? *BMC Genomics*, 6: 59.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S. and Baulcombe, D. C. (2000). An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell,* 101: 543-553.

Day, A., Neutelings, G., Nolin, F., Grec, S., Habrant, A., Crônier, D., Maher, B., Rolando, C., David, H., Chabbert, B. and Hawkins, S. (2009). Caffeoyl coenzyme A *O*-methyltransferase down-regulation is associated with modifications in lignin and cell-wall architecture in flax secondary xylem. *Plant Physiology and Biochemistry,* 47: 9-19.

de Souza, A. P., Grandis, A., Leite, D. C. C. and Buckeridge, M. S. (2014). Sugarcane as a bioenergy source: History, performance, and perspectives for second-generation bioethanol. *BioEnergy Research,* 7: 24-35.

de Souza, A. P., Leite, D. C. C., Pattathil, S., Hahn, M. G. and Buckeridge, M. S. (2013). Composition and structure of sugarcane cell wall polysaccharides: Implications for second-generation bioethanol production. *BioEnergy Research,* 6: 564-579.

Delmer, D. P. and Haigler, C. H. (2002). The regulation of metabolic flux to cellulose, a major sink for carbon in plants. *Metabolic Engineering,* 4: 22-28.

Deluc, L., Barrieu, F., Marchive, C., Lauvergeat, V., Decendit, A., Richard, T., Carde, J., Mérillon, J. and Hamdi, S. (2006). Characterization of a grapevine R2R3-MYB transcription factor that regulates the phenylpropanoid pathway. *Plant Physiology,* 140: 499-511.

Du, H., Feng, B., Yang, S., Huang, Y. and Tang, Y. (2012a). The R2R3-MYB transcription factor gene family in maize. *PloS ONE,* 7: e37463.

Du, H., Yang, S., Liang, Z., Feng, B., Liu, L., Huang, Y. and Tang, Y. (2012b). Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC Plant Biology,* 12: 106.

Eudes, A., Liang, Y., Mitra, P. and Loqué, D. (2014). Lignin bioengineering. *Current Opinion in Biotechnology,* 26: 189-198.

Fagard, M., Boutet, S., Morel, J., Bellini, C. and Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. *Proceedings of the National Academy of Sciences of the United States of America,* 97: 11650–11654.

Feng, C., Andreasson, E., Maslak, A., Mock, H. P., Mattsson, O. and Mundy, J. (2004). *Arabidopsis MYB68* in development and responses to environmental cues. *Plant Science,* 167: 1099-1107.

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

Filipowicz, W. (2005). RNAi : The nuts and bolts of the RISC machine. *Cell,* 122: 17-20.

Firon, N., LaBonte, D., Villordon, A., Kfir, Y., Solis, J., Lapis, E., Perlman, T. S., Doron-Faigenboim, A., Hetzroni, A., Althan, L. and Nadir, L. A. (2013). Transcriptional profiling of sweetpotato (*Ipomoea batatas*) roots indicates down-regulation of lignin biosynthesis and up-regulation of starch biosynthesis at an early stage of storage root formation. *BMC Genomics,* 14: 460.

Fornalé, S., Capellades, M., Encina, A., Wang, K., Irar, S., Lapierre, C., Ruel, K., Joseleau, J., Berenguer, J., Puigdomènech, P., Rigau, J. and Caparrós-Ruiz, D. (2012). Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. *Molecular Plant,* 5: 817-830.

Fornalé, S., Shi, X., Chai, C., Encina, A., Irar, S., Capellades, M., Fuguet, E., Torres, J., Rovira, P., Puigdomènech, P., Rigau, J., Grotewold, E., Gray, J. and Caparrós-Ruiz, D. (2010). *Zm*MYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *The Plant Journal,* 64: 633-644.

Fornalé, S., Sonbol, F., Maes, T., Capellades, M., Puigdomènech, P., Rigau, J. and Caparrós-Ruiz, D. (2006). Down-regulation of the maize and *Arabidopsis thaliana* caffeic acid *O*-methyl-transferase genes by two new maize R2R3-MYB transcription factors. *Plant Molecular Biology,* 62: 809-823.

Frizzi, A. and Huang, S. (2010). Tapping RNA silencing pathways for plant biotechnology. *Plant Biotechnology Journal,* 8: 1-23.

Fu, C., Mielenz, J. R., Xiao, X., Ge, Y., Hamilton, C. Y., Rodriguez, M., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R. A. and Wang, Z. (2011a). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proceedings of the National Academy of Sciences of the United States of America,* 108: 3803-3808.

Fu, C., Xiao, X., Xi, Y., Ge, Y., Chen, F., Bouton, J., Dixon, R. A. and Wang, Z. (2011b). Downregulation of cinnamyl alcohol dehydrogenase (CAD) leads to improved saccharification efficiency in switchgrass. *BioEnergy Research,* 4: 153-164.

Gabrielsen, O. S., Sentenac, A. and Fromageot, P. (1991). Specific DNA binding by c-Myb: Evidence for a double helix-turn-helix-related motif. *Science,* 253: 1140-1143.

Gardiner, J., Schroeder, S., Polacco, M. L., Sanchez-Villeda, H., Fang, Z., Morgante, M., Landewe, T., Fengler, K., Useche, F., Hanafey, M., Tingey, S., Chou, H., Wing, R., Soderlund, C. and Coe, E. H. (2004). Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization. *Plant Physiology,* 134: 1317-1326.

Git, A., Dvinge, H., Salmon-Divon, M., Osborne, M., Kutter, C., Hadfield, J., Bertone, P. and Caldas, C. (2010). Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA*, 16: 991-1006.

Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., Lapierre, C., Pollet, B., Verhaegen, D., Chaubet-Gigot, N. and Grima-Pettenati, J. (2005). *Eg*MYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *The Plant Journal,* 43: 553-567.

Goldemberg, J. (2007). Ethanol for a sustainable energy future. *Science,* 315: 808-810.

Grabber, J. H. (2005). How do lignin composition, structure, and cross-linking affect degradability? A review of cell wall model studies. *Crop Science,* 45: 820-831.

Gray, J., Caparrós-Ruiz, D. and Grotewold, E. (2012). Grass phenylpropanoids: Regulate before using! *Plant Science,* 184: 112-120.

Grivet, L. and Arruda, P. (2001). Sugarcane genomics: Depicting the complex genome of an important tropical crop. *Current Opinion in Plant Biology*, 5: 122-127.

GUDSF (2005). Ethanol/EDTA/Sodium Acetate Precipitation. Available online: www.griffith.edu.au/science-aviation/dna-sequencing-facility Last accessed 19/10/2014.

Guillaumie, S., San-Clemente, H., Deswarte, C., Martinez, Y., Lapierre, C., Murigneux, A., Barrière, Y., Pichon, M. and Goffner, D. (2007). MAIZEWALL. Database and developmental gene expression profiling of cell wall biosynthesis and assembly in maize. *Plant Physiology,* 143: 339-363.

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

Guo, D., Chen, F., Wheeler, J., Winder, J., Selman, S., Peterson, M. and Dixon, R. A. (2001b). Improvement of in-rumen digestibility of alfalfa forage by genetic manipulation of lignin *O*-methyltransferases. *Transgenic Research,* 10: 457-464.

Halpin, C. (2004). Investigating and manipulating lignin biosynthesis in the postgenomic era. *Advances in Botanical Research,* 41: 63-106.

Hames, B., Ruiz, R., Scarlata, C., Sluiter, A., Sluiter, J. and Templeton, D. (2008). Preparation of samples for compositional analysis: Laboratory analytical procedure (LAP). *National Renewable Energy Laboratory,* Technical report: NREL/TP-510-42620.

Hamilton, A., Voinnet, O., Chappell, L. and Baulcombe, D. (2002). Two classes of short interfering RNA in RNA silencing. *The European Molecular Biology Organization Journal,* 21: 4671-4679.

Hamilton, A. J. and Baulcombe, D. C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science,* 286: 950-952.

Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature,* 404: 293-296.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology,* 166: 557-580.

Harrison, M. D., Zhang, Z., Shand, K., O'Hara, I. M., Doherty, W. O. S. and Dale, J. L. (2013). Effect of pretreatment on saccharification of sugar cane bagasse by complex and simple enzyme mixtures. *Bioresource Technology,* 148: 105-113.

He, X., Hall, M. B., Gallo-Meagher, M. and Smith, R. L. (2003). Improvement of forage quality by downregulation of maize *O*-methyltransferase. *Crop Science,* 43: 2240-2251.

Hill, J., Nelson, E., Tilman, D., Polasky, S. and Tiffany, D. (2006). Environmental , economic , and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America,* 103: 11206-11210.

Himber, C., Dunoyer, P., Moissiard, G., Ritzenthaler, C. and Voinnet, O. (2003). Transitivity-dependent and -independent cell-to-cell movement of RNA silencing. *The European Molecular Biology Organization Journal,* 22: 4523-4533.

Hirano, K., Kondo, M., Aya, K., Miyao, A., Sato, Y., Antonio, B. A., Namiki, N., Nagamura, Y. and Matsuoka, M. (2013). Identification of transcription factors involved in rice secondary cell wall formation. *Plant and Cell Physiology,* 54: 1791-1802.

Hisano, H., Nandakumar, R. and Wang, Z. (2009). Genetic modification of lignin biosynthesis for improved biofuel production. *In Vitro Cellular & Developmental Biology - Plant,* 45: 306-313.

Ho, J. H. and Hong, C. (2011). Salvianolic acids: Small compounds with multiple mechanisms for cardiovascular protection. *Journal of Biomedical Science,* 18: 30.

Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., Pollet, B. and Legrand, M. (2004). Silencing of hydroxycinnamoyl-Coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *The Plant Cell,* 16: 1446-1465.

Huang, H., Xu, L., Tong, Z., Lin, E., Liu, Q., Cheng, L. and Zhu, M. (2012). De novo characterization of the Chinese fir (*Cunninghamia lanceolata*) transcriptome and analysis of candidate genes involved in cellulose and lignin biosynthesis. *BMC Genomics,* 13: 648.

Hughes, T. A. (2006). Regulation of gene expression by alternative untranslated regions. *Trends in Genetics,* 22: 119-122.

Huntley, S. K., Ellis, D., Gilbert, M., Chapple, C. and Mansfield, S. D. (2003). Significant increases in pulping efficiency in C4H-F5H-transformed poplars: Improved chemical savings and reduced environmental toxins. *Journal of Agricultural and Food Chemistry,* 51: 6178-6183.

Inman-Bamber, N. G., Bonnett, G. D., Spillman, M. F., Hewitt, M. L. and Jackson, J. (2008). Increasing sucrose accumulation in sugarcane by manipulating leaf extension and photosynthesis with irrigation. *Australian Journal of Agricultural Research,* 59: 13-26.

Inoue, H., Nojima, H. and Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene,* 96: 23-28.

Jackson, D., Culianez-Macia, F., Prescott, A. G., Roberts, K. and Martin, C. (1991). Expression patterns of myb genes from Antirrhinum flowers. *The Plant Cell,* 3: 115-125.

Jackson, L. A., Shadle, G. L., Zhou, R., Nakashima, J., Chen, F. and Dixon, R. A. (2008). Improving saccharification efficiency of alfalfa stems through modification of the terminal stages of monolignol biosynthesis. *BioEnergy Research,* 1: 180-192.

Jackson, P. (2013). *Energy Cane*, in Saha, M. C., Bhandari, H. S. and Bouton, J. H. (Eds.). *Bioenergy feedstocks: Breeding and genetics* (1st ed.). Somerset, NJ: John Wiley & Sons, Inc., p. 117–149.

Jefferson, R. A. (1987). Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Molecular Biology Reporter,* 5: 387-405.

Jefferson, R. A., Kavanagh, T. A. and Bevan, M. W. (1987). GUS fusions: *β*-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The European Molecular Biology Organization Journal,* 6: 3901-3907.

Jiang, C., Gu, X. and Peterson, T. (2004). Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of *Arabidopsis* and *Oryza sativa* L. ssp. *indica*. *Genome Biology,* 5: R46.

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B. and Martin, C. (2000). Transcriptional repression by *At*MYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *The European Molecular Biology Organization Journal,* 19: 6150-6161.

Jouanin, L., Goujon, T., de Nadaï, V., Martin, M., Mila, I., Vallet, C., Pollet, B., Yoshinaga, A., Chabbert, B., Petit-Conil, M. and Lapierre, C. (2000). Lignification in transgenic poplars with extremely reduced caffeic acid *O*-methyltransferase activity. *Plant Physiology,* 123: 1363-1374.

Jung, H. G. and Casler, M. D. (2006). Maize stem tissues: Cell wall concentration and composition during development. *Crop Science,* 46: 1793-1800.

Jung, J. H., Fouad, W. M., Vermerris, W., Gallo, M. and Altpeter, F. (2012). RNAi suppression of lignin biosynthesis in sugarcane reduces recalcitrance for biofuel production from lignocellulosic biomass. *Plant Biotechnology Journal,* 10: 1067-1076.

Jung, J. H., Vermerris, W., Gallo, M., Fedenko, J. R., Erickson, J. E. and Altpeter, F. (2013). RNA interference suppression of lignin biosynthesis increases fermentable sugar yields for biofuel production from field-grown sugarcane. *Plant Biotechnology Journal,* 11: 709-716.

Kanehisa, M. and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research,* 28: 27-30.

Karpinska, B., Karlsson, M., Srivastava, M., Stenberg, A., Schrader, J., Sterky, F., Bhalerao, R. and Wingsle, G. (2004). MYB transcription factors are differentially expressed and regulated during secondary vascular tissue development in hybrid aspen. *Plant Molecular Biology,* 56: 255-270.

Kim, H. and Ralph, J. (2010). Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6/pyridine-d_5. *Organic & Biomolecular Chemistry,* 8: 576-591.

Kim, H., Ralph, J. and Akiyama, T. (2008). Solution-state 2D NMR of ball-milled plant cell wall gels in DMSO-d_6. *BioEnergy Research,* 1: 56-66.

Kim, W., Ko, J. and Han, K. (2012). Identification of a *cis*-acting regulatory motif recognized by MYB46, a master transcriptional regulator of secondary wall biosynthesis. *Plant Molecular Biology,* 78: 489-501.

Kim, W., Ko, J., Kim, J., Kim, J., Bae, H. and Han, K. (2013). MYB46 directly regulates the gene expression of secondary wall-associated cellulose synthases in *Arabidopsis*. *The Plant Journal,* 73: 26-36.

Kiyota, E., Mazzafera, P. and Sawaya, A. C. H. F. (2012). Analysis of soluble lignin in sugarcane by ultrahigh performance liquid chromatography-tandem mass spectrometry with a do-it-yourself oligomer database. *Analytical Chemistry,* 84: 7015-7020.

Klempnauer, K., Gonda, T. J. and Bishop, J. M. (1982). Nucleotide sequence of the retroviral leukemia gene *v-myb* and its cellular progenitor *c-myb*: The architecture of a transduced oncogene. *Cell,* 31: 453-463.

Ko, J., Kim, W. and Han, K. (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in *Arabidopsis*. *The Plant Journal,* 60: 649-665.

Kolahi, M., Jonoubi, P., Majd, A., Tabandeh, M. R. and Hashemitabar, M. (2013). Differential expression of phenylalanine ammonia-lyase in different tissues of sugarcane (*Saccharum officinarum* L.) during development. *BioResources,* 8: 4912-4922.

Koutaniemi, S., Warinowski, T., Kärkönen, A., Alatalo, E., Fossdal, C. G., Saranpää, P., Laakso, T., Fagersted, K. V., Simola, L. K., Paulin, L., Rudd, S. and Teeri, T. H. (2007). Expression profiling of the lignin biosynthetic pathway in Norway spruce using EST sequencing and real-time RT-PCR. *Plant Molecular Biology,* 65: 311-328.

Kranz, H. D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meissner, R. C., Petroni, K., Urzainqui, A., Bevan, M., Martin, C., Smeekens, S., Tonelli, C., Paz-Ares, J. and Weisshaar, B. (1998). Towards functional characterisation of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *The Plant Journal,* 16: 263-276.

Lacombe, E., Hawkins, S., Van Doorsselaere, J., Piquemal, J., Goffner, D., Poeydomenge, O., Boudet, A. and Grima-Pettenati, J. (1997). Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: Cloning, expression and phylogenetic relationships. *The Plant Journal,* 11: 429-441.

Lakshmanan, P., Geijskes, R. J., Aitken, K. S., Grof, C. L. P., Bonnett, G. D. and Smith, G. R. (2005). Sugarcane biotechnology: The challenges and opportunities. *In Vitro Cellular & Developmental Biology - Plant,* 41: 345-363.

Lauvergeat, V., Rech, P., Jauneau, A., Guez, C., Coutos-Thevenot, P. and Grima-Pettenati, J. (2002). The vascular expression pattern directed by the *Eucalyptus gunnii* cinnamyl alcohol dehydrogenase *EgCAD2* promoter is conserved among woody and herbaceous plant species. *Plant Molecular Biology,* 50: 497-509.

Legay, S., Lacombe, E., Goicoechea, M., Brière, C., Séguin, A., Mackay, J. and Grima-Pettenati, J. (2007). Molecular characterization of *EgMYB1*, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Science,* 173: 542-549.

Legay, S., Sivadon, P., Blervacq, A., Pavy, N., Baghdady, A., Tremblay, L., Levasseur, C., Ladouce, N., Lapierre, C., Séguin, A., Hawkins, S., Mackay, J. and Grima-Pettenati, J. (2010). *Eg*MYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in *Arabidopsis* and poplar. *New Phytologist,* 188: 774-786.

Levy, M., Edelbaum, O. and Sela, I. (2004). Tobacco mosaic virus regulates the expression of its own resistance gene *N*. *Plant Physiology,* 135: 2392-2397.

Li, C. and Lu, S. (2014). Genome-wide characterization and comparative analysis of R2R3-MYB transcription factors shows the complexity of MYB-associated regulatory networks in *Salvia miltiorrhiza*. *BMC Genomics,* 15: 277-288.

Li, J., Yang, Z., Yu, B., Liu, J. and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. *Current Biology,* 15: 1501-1507.

Li, X., Chen, W., Zhao, Y., Xiang, Y., Jiang, H., Zhu, S. and Cheng, B. (2013). Downregulation of caffeoyl-CoA O-methyltransferase (*CCoAOMT*) by RNA interference leads to reduced lignin production in maize straw. *Genetics and Molecular Biology,* 36: 540-546.

Li, X., Weng, J. and Chapple, C. (2008). Improvement of biomass through lignin modification. *The Plant Journal,* 54: 569-581.

Limayem, A. and Ricke, S. C. (2012). Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects. *Progress in Energy and Combustion Science,* 38: 449-467.

Lingle, S. E. and Thomson, J. L. (2012). Sugarcane internode composition during crop development. *BioEnergy Research,* 5: 168-178.

Liu, C. (2012). Deciphering the enigma of lignification: Precursor transport, oxidation, and the topochemistry of lignin assembly. *Molecular Plant,* 5: 304-317.

Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods,* 25: 402-408.

Louie, G. V., Bowman, M. E., Tu, Y., Mouradov, A., Spangenberg, G. and Noel, J. P. (2010). Structure-function analyses of a caffeic acid *O*-methyltransferase from perennial ryegrass reveal the molecular basis for substrate preference. *The Plant Cell,* 22: 4114–4127.

Luo, K. Q and Chang, D. C. (2004). The gene-silencing efficiency of siRNA is strongly dependent on the local structure of mRNA at the targeted region. *Biochemical and Biophysical Research Communications,* 318: 303–310.

Ma, Q. (2007). Characterization of a cinnamoyl-CoA reductase that is associated with stem development in wheat. *Journal of Experimental Botany,* 58: 2011-2021.

Ma, Q., Wang, C. and Zhu, H. (2011). *Ta*MYB4 cloned from wheat regulates lignin biosynthesis through negatively controlling the transcripts of both cinnamyl alcohol dehydrogenase and cinnamoyl-CoA reductase genes. *Biochimie,* 93: 1179-1186.

Margis, R., Fusaro, A. F., Smith, N. A., Curtin, S. J., Watson, J. M., Finnegan, E. J. and Waterhouse, P. M. (2006). The evolution and diversification of Dicers in plants. *FEBS Letters,* 580: 2442-2450.

Marita, J. M., Ralph, J., Hatfield, R. D., Guo, D., Chen, F. and Dixon, R. A. (2003). Structural and compositional modifications in lignin of transgenic alfalfa down-regulated in caffeic acid 3-*O*-methyltransferase and caffeoyl coenzyme A 3-*O*-methyltransferase. *Phytochemistry,* 62: 53-65.

Masarin, F., Gurpilhares, D. B., Baffa, D. C. F., Barbosa, M. H. P., Carvalho, W., Ferraz, A. and Milagres, A. M. F. (2011). Chemical composition and enzymatic digestibility of sugarcane clones selected for varied lignin content. *Biotechnology for Biofuels,* 4: 55.

Mazourek, M., Pujar, A., Borovsky, Y., Paran, I., Mueller, L. and Jahn, M. M. (2009). A dynamic interface for capsaicinoid systems biology. *Plant Physiology,* 150: 1806-1821.

McCarthy, R. L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., Spicer, C. and Ye, Z. (2010). The poplar MYB transcription factors, *Ptr*MYB3 and *Ptr*MYB20 , are involved in the regulation of secondary wall biosynthesis. *Plant and Cell Physiology,* 51: 1084-1090.

McCarthy, R. L., Zhong, R. and Ye, Z. (2009). MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant and Cell Physiology,* 50: 1950-1964.

Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A. and Matzke, A. J. M. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *The European Molecular Biology Organization Journal,* 19: 5194-5201.

Meyermans, H., Morreel, K., Lapierre, C., Pollet, B., De Bruyn, A., Busson, R., Herdewijn, P., Devreese, B., Van Beeumen, J., Marita, J. M., Ralph, J., Chen, C., Burggraeve, B., Van Montagu, M., Messens, E. and Boerjan, W. (2000). Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-Coenzyme A *O*-methyltransferase, an enzyme involved in lignin biosynthesis. *The Journal of Biological Chemistry,* 275: 36899-36909.

Mignone, F., Gissi, C., Liuni, S. and Pesole, G. (2002). Untranslated regions of mRNAs. *Genome Biology,* 3: 0004.0001-0004.0010.

Miller, J. H. (1972). *Experiments in molecular genetics* (3rd edition). New York: Cold Spring Harbor Laboratory.

Mlotshwa, S., Voinnet, O., Mette, M. F., Matzke, M., Vaucheret, H., Ding, S. W., Pruss, G. and Vance, V. B. (2002). RNA silencing and the mobile silencing signal. *The Plant Cell,* 14 (suppl 1): S289-S301.

Moore, P. H. (1987). *Anatomy and morphology,* in Heinz, D. J. (Ed.). *Developments in crop science: Sugarcane improvement through breeding*. Amsterdam: Elsevier, p. 85-142.

Morel, J., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., Proux, F. and Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (*ago1*) mutants impaired in post-transcriptional gene silencing and virus resistance. *The Plant Cell,* 14: 629-639.

Morey, J. S., Ryan, J. C. and Van Dolah, F. M. (2006). Microarray validation: Factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biological Procedures Online*, 8: 175-193.

Mudge, S. R., Osabe, K., Casu, R. E., Bonnett, G. D., Manners, J. M. and Birch, R. G. (2009). Efficient silencing of reporter transgenes coupled to known functional promoters in sugarcane, a highly polyploid crop species. *Planta,* 229: 549-558.

Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum,* 15: 473-497.

Murray, J. and King, D. (2012). Oil's tipping point has passed. *Nature,* 481: 433-435.

Nakashima, J., Chen, F., Jackson, L., Shadle, G. and Dixon, R. A. (2008). Multi-site genetic modification of monolignol biosynthesis in alfalfa (*Medicago sativa*): Effects on lignin composition in specific cell types. *New Phytologist,* 179: 738-750.

Nelson, D. R. (2006). Plant cytochrome P450s from moss to poplar. *Phytochemistry Reviews,* 5: 193-204.

Newman, L. J., Perazza, D. E., Juda, L. and Campbell, M. M. (2004). Involvement of the R2R3-MYB, *At*MYB61, in the ectopic lignification and dark-photomorphogenic components of the *det3* mutant phenotype. *The Plant Journal,* 37: 239-250.

Ogata, K., Chie, K., Sasaki, M., Hatanaka, H., Nagadoi, A., Enari, M., Nakamura, H., Nishimura, Y., Ishii, S. and Sarai, A. (1996). The cavity in the hydrophobic core of Myb DNA-binding domain is reserved for DNA recognition and *trans*-activation. *Nature Structural Biology,* 3: 178-187.

Ogata, K., Morikawa, S., Nakamura, H., Hojo, H., Yoshimura, S., Zhang, R., Aimoto, S., Ametani, Y., Hirata, Z., Sarai, A., Ishii, S. and Nishimura, Y. (1995). Comparison of the free and DNA-complexed forms of the DNA-binding domain from c-Myb. *Nature Structural Biology,* 2: 309-320.

Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S. and Nishimura, Y. (1994). Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell,* 79: 639-648.

Öhman, D., Demedts, B., Kumar, M., Gerber, L., Gorzsás, A., Goeminne, G., Hedenström, M., Ellis, B., Boerjan, W. and Sundberg, B. (2013). MYB103 is required for ferulate-5-hydroxylase expression and syringyl lignin biosynthesis in *Arabidopsis* stems. *The Plant Journal,* 73: 63-76.

Omer, S., Kumar, S. and Khan, B. M. (2013). Over-expression of a subgroup 4 R2R3 type MYB transcription factor gene from *Leucaena leucocephala* reduces lignin content in transgenic tobacco. *Plant Cell Reports,* 32: 161-171.

Ong, S. S. and Wickneswari, R. (2012). Characterization of microRNAs expressed during secondary wall biosynthesis in *Acacia mangium*. *PloS ONE,* 7: e49662.

Oraby, H. F. and Ramadan, M. F. (2014). Impact of suppressing the caffeic acid *O*-methyltransferase (*COMT*) gene on lignin, fiber, and seed oil composition in *Brassica napus* transgenic plants. *European Food Research and Technology*, DOI 10.1007/s00217-014-2397-3. Published online: 04 December 2014.

Osabe, K., Mudge, S. R., Graham, M. W. and Birch, R. G. (2009). RNAi mediated down-regulation of *PDS* gene expression in sugarcane (*Saccharum*), a highly polyploid crop. *Tropical Plant Biology,* 2: 143-148.

Pan, Y., Michael, T. P., Hudson, M. E., Kay, S. A., Chory, J. and Schuler, M. A. (2009). Cytochrome P450 monooxygenases as reporters for circadian-regulated pathways. *Plant Physiology,* 150: 858-878.

Pandey, A., Soccol, C. R., Nigam, P. and Soccol, V. T. (2000). Biotechnological potential of agro-industrial residues. I: Sugarcane bagasse. *Bioresource Technology,* 74: 69-80.

Papini-Terzi, F. S., Rocha, F. R., Vêncio, R. Z. N., Felix, J. M., Branco, D. S., Waclawovsky, A. J., Del Bem, L. E. V., Lembke, C. G., Costa, M. D. L., Nishiyama, M. Y., Vicentini, R., Vincentz, M. G. A., Ulian, E. C., Menossi, M. and Souza, G. M. (2009). Sugarcane genes associated with sucrose content. *BMC Genomics,* 10: 120.

Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L. J., Smith, C., Bevan, M. W., Mansfield, S., Whetten, R. W., Sederoff, R. R. and Campbell, M. M. (2003a). Characterisation of a pine MYB that regulates lignification. *The Plant Journal,* 36: 743-754.

Patzlaff, A., Newman, L. J., Dubos, C., Whetten, R. W., Smith, C., McInnis, S., Bevan, M. W., Sederoff, R. R. and Campbell, M. M. (2003b). Characterisation of *Pt*MYB1 , an R2R3-MYB from pine xylem. *Plant Molecular Biology,* 53: 597-608.

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A. and Saedler, H. (1987). The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. *The European Molecular Biology Organization Journal,* 6: 3553-3558.

Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F. and Liuni, S. (2001). Structural and functional features of eukaryotic mRNA untranslated regions. *Gene,* 276: 73-81.

Pinçon, G., Maury, S., Hoffmann, L., Geoffroy, P., Lapierre, C., Pollet, B. and Legrand, M. (2001). Repression of *O*-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth. *Phytochemistry,* 57: 1167-1176.

Piquemal, J., Chamayou, S., Nadaud, I., Beckert, M., Barrière, Y., Mila, I., Lapierre, C., Rigau, J., Puigdomenech, P., Jauneau, A., Digonnet, C., Boudet, A., Goffner, D. and Pichon, M. (2002). Down-regulation of caffeic acid *O*-methyltransferase in maize revisited using a transgenic approach. *Plant Physiology,* 130: 1675-1685.

Plomion, C., Leprovost, G. and Stokes, A. (2001). Wood formation in trees. *Plant Physiology,* 127: 1513-1523.

Porter, T. D. and Coon, M. J. (1991). Cytochrome P-450: Multiplicity of isoforms, substrates, and catalytic and regulatory mechanisms. *The Journal of Biological Chemistry,* 266: 13469-13472.

Prashant, S., Sunita, M. S., Pramod, S., Gupta, R. K., Anil Kumar, S., Rao Karumanchi, S., Rawal, S. K. and Kishor, P. B. K. (2011). Down-regulation of *Leucaena leucocephala* cinnamoyl CoA reductase (*LlCCR*) gene induces significant changes in phenotype, soluble phenolic pools and lignin in transgenic tobacco. *Plant Cell Reports,* 30: 2215-2231.

Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G. and Cormier, M. J. (1992). Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene,* 111: 229-233.

Prouse, M. B. and Campbell, M. M. (2012). The interaction between MYB proteins and their target DNA binding sites. *Biochimica et Biophysica Acta,* 1819: 67-77.

Prouse, M. B. and Campbell, M. M. (2013). Interactions between the R2R3-MYB transcription factor, *At*MYB61, and target DNA binding sites. *PloS ONE,* 8: e65132.

Puigdomenech, P. C., Perez, P., Murigneux, A., Martinant, J. P., Tixier, M. H., Rigau, J., Civardi, L. and Maes, T. (2001). Identifying genes associated with a qtl corn digestibility locus. Patent WO 0155395-A 2. Biogemma. France.

Raes, J., Rohde, A., Christensen, J. H., Van de Peer, Y. and Boerjan, W. (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiology,* 133: 1051-1071.

Reddy, M. S. S., Chen, F., Shadle, G., Jackson, L., Aljoe, H. and Dixon, R. A. (2005). Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.). *Proceedings of the National Academy of Sciences of the United States of America,* 102: 16573-16578.

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O. J., Samaha, R. R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J. Z., Ghandehari, D., Sherman, B. K. and Yu, G. (2000). *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes. *Science,* 290: 2105-2110.

Robbins, M. L., Roy, A., Wang, P., Gaffoor, I., Sekhon, R. S., Buanafina, M. M. de O., Rohila, J. S. and Chopra, S. (2013). Comparative proteomics analysis by DIGE and iTRAQ provides insight into the regulation of phenylpropanoids in maize. *Journal of Proteomics,* 93: 254-275.

Rodrigues, F. A., de Laia, M. L. and Zingaretti, S. M. (2009). Analysis of gene expression profiles under water stress in tolerant and sensitive sugarcane plants. *Plant Science,* 176: 286-302.

Rogers, L. A. and Campbell, M. M. (2004). The genetic control of lignin deposition during plant growth and development. *New Phytologist,* 164: 17-30.

Rogers, L. A., Dubos, C., Cullis, I. F., Surman, C., Poole, M., Willment, J., Mansfield, S. D. and Campbell, M. M. (2005). Light, the circadian clock, and sugar perception in the control of lignin biosynthesis. *Journal of Experimental Botany,* 56: 1651-1663.

Rozen, S. and Skaletsky, H. J. (2000). *Primer3 on the WWW for general users and for biologist programmers,* in Krawetz, S. and Misener, S. (Eds.). *Bioinformatics methods and protocols: Methods in molecular biology*. Totowa, NJ: Humana Press, p. 365-386.

Saathoff, A. J., Sarath, G., Chow, E. K., Dien, B. S. and Tobias, C. M. (2011). Downregulation of cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. *PloS ONE,* 6: e16416.

Sablowski, R. W. M., Moyano, E., Culianez-Macia, F. A., Martin, C. and Bevan, M. (1994). A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. *The European Molecular Biology Organization Journal,* 13: 128-137.

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H. and Siomi, M. C. (2007). Pimet, the *Drosophila* homolog of HEN1, mediates 2'-*O*-methylation of Piwi-interacting RNAs at their 3' ends. *Genes and Development,* 21: 1603-1608.

Salmona, J., Dussert, S., Descroix, F., de Kochko, A., Bertrand, B. and Joët, T. (2008). Deciphering transcriptional networks that govern *Coffea arabica* seed development using combined cDNA array and real-time RT-PCR approaches. *Plant Molecular Biology,* 66: 105-124.

Sambrook, J. and Russell, D. W. (2001). *Molecular cloning: A laboratory manual* (3rd edition). New York: Cold Spring Harbor Laboratory Press.

Samuel, R., Pu, Y., Jiang, N., Fu, C., Wang, Z. and Ragauskas, A. (2014). Structural characterization of lignin in wild-type versus COMT down-regulated switchgrass. *Frontiers in Energy Research,* 1: 1-9.

Schell, D. J., Farmer, J., Newman, M. and McMillan, J. D. (2003). Dilute-sulfuric acid pretreatment of corn stover in pilot-scale reactor: Investigation of yields, kinetics, and enzymatic digestibilities of solids. *Applied Biochemistry and Biotechnology*, 105-108: 69-85.

Schwartz, T. J. and Lawoko, M. (2010). Removal of acid-soluble lignin from biomass extracts using Amberlite XAD-4 resin. *BioResources*, 5: 2337-2347.

Segal, L., Creely, J. J., Martin, A. E. and Conrad, C. M. (1959). An empirical method for estimating the degree of crystallinity of native cellulose using the x-ray diffractometer. *Textile Research Journal,* 29: 786-794.

Selman-Housein, G., López, M. A., Miranda, F., Hernández, D., Civardi, L., Rigau, J. and Puigdome, P. (1999). Molecular cloning of cDNAs coding for three sugarcane enzymes involved in lignification. *Plant Science,* 143: 163-171.

Shadle, G., Chen, F., Reddy, M. S. S., Jackson, L., Nakashima, J. and Dixon, R. A. (2007). Down-regulation of hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry,* 68: 1521-1529.

Shang, Q., Li, L. and Dong, C. (2012). Multiple tandem duplication of the phenylalanine ammonia-lyase genes in *Cucumis sativus* L. *Planta,* 236: 1093-1105.

Sharma, A. K. and Sharma, M. K. (2009). Plants as bioreactors: Recent developments and emerging opportunities. *Biotechnology Advances,* 27: 811-832.

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

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

Shen, H., Poovaiah, C. R., Ziebell, A., Tschaplinski, T. J., Pattathil, S., Gjersing, E., Engle, N. L., Katahira, R., Pu, Y., Sykes, R., Chen, F., Ragauskas, A. J., Mielenz, J. R., Hahn, M. G., Davis, M., Stewart, C. N. and Dixon, R. A. (2013). Enhanced characteristics of genetically modified switchgrass (*Panicum virgatum* L.) for high biofuel production. *Biotechnology for Biofuels,* 6: 71.

Shi, R., Sun, Y., Li, Q., Heber, S., Sederoff, R. and Chiang, V. L. (2010). Towards a systems approach for lignin biosynthesis in *Populus trichocarpa*: Transcript abundance and specificity of the monolignol biosynthetic genes. *Plant and Cell Physiology,* 51: 144-163.

Silverstein, R. A., Chen, Y., Sharma-Shivappa, R. R., Boyette, M. D. and Osborne, J. (2007). A comparison of chemical pretreatment methods for improving saccharification of cotton stalks. *Bioresource Technology*, 98: 3000–3011.

Simmons, B. A., Loqué, D. and Ralph, J. (2010). Advances in modifying lignin for enhanced biofuel production. *Current Opinion in Plant Biology,* 13: 312-319.

Sluiter, A., Hames, B., Hyman, D., Payne, C., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D. and Wolfe, J. (2008a). Determination of total solids in biomass and total dissolved solids in liquid process samples: Laboratory analytical procedure (LAP). *National Renewable Energy Laboratory,* Technical report: NREL/TP-510-42621.

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., Templeton, D. and Crocker, D. (2008b). Determination of structural carbohydrates and lignin in biomass: Laboratory analytical procedure (LAP). *National Renewable Energy Laboratory,* Technical report: NREL/TP-510-42618.

Sluiter, A., Ruiz, R., Scarlata, C., Sluiter, J. and Templeton, D. (2008c). Determination of extractives in biomass: Laboratory analytical procedure (LAP). *National Renewable Energy Laboratory,* Technical report: NREL/TP-510-42619.

Soderlund, C., Descour, A., Kudrna, D., Bomhoff, M., Boyd, L., Currie, J., Angelova, A., Collura, K., Wissotski, M., Ashley, E., Morrow, D., Fernandes, J., Walbot, V. and Yu, Y. (2009). Sequencing, mapping, and analysis of 27,455 maize full-length cDNAs. *PLoS Genetics,* 5: e1000740.

Somerville, C., Youngs, H., Taylor, C., Davis, S. C. and Long, S. P. (2010). Feedstocks for lignocellulosic biofuels. *Science*, 329: 790-792.

Sonbol, F., Fornalé, S., Capellades, M., Encina, A., Touriño, S., Torres, J., Rovira, P., Ruel, K., Puigdomènech, P., Rigau, J. and Caparrós-Ruiz, D. (2009). The maize *Zm*MYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in *Arabidopsis thaliana*. *Plant Molecular Biology,* 70: 283-296.

Stewart, J. J., Akiyama, T., Chapple, C., Ralph, J. and Mansfield, S. D. (2009). The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiology,* 150: 621-635.

Stracke, R., Werber, M. and Weisshaar, B. (2001). The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology,* 4: 447-456.

Studer, M. H., DeMartini, J. D., Davis, M. F., Sykes, R. W., Davison, B., Keller, M., Tuskan, G. A. and Wyman, C. E. (2011). Lignin content in natural *Populus* variants affects sugar release. *Proceedings of the National Academy of Sciences of the United States of America,* 108: 6300-6305.

Suprasanna, P., Patade, V. Y., Desai, N. S., Devarumath, R. M., Kawar, P. G., Pagariya, M. C., Ganapathi, A., Manickavasagam, M. and Babu, K. H. (2011). Biotechnological developments in sugarcane improvement: An overview. *Sugar Tech,* 13: 322-335.

Tamagnone, L., Merida, A., Parr, A., Mackay, S., Culianez-Macia, F. A., Roberts, K. and Martin, C. (1998). The *Am*MYB308 and *Am*MYB330 transcription factors from Antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *The Plant Cell,* 10: 135-154.

Taylor, S., Wakem, M., Dijkman, G., Alsarraj, M. and Nguyen, M. (2011). A practical approach to RT-qPCR: Publishing data that conform to the MIQE guidelines. *Bio-Rad Laboratories, Inc.,* Bulletin 5859: tech note 5859.

Thévenin, J., Pollet, B., Letarnec, B., Saulnier, L., Gissot, L., Maia-Grondard, A., Lapierre, C. and Jouanin, L. (2011). The simultaneous repression of CCR and CAD, two enzymes of the lignin biosynthetic pathway, results in sterility and dwarfism in *Arabidopsis thaliana*. *Molecular Plant,* 4: 70-82.

Thomas, N. R., Kosh, S., Simsek, M. and Abraham, A. K. (1988). A precaution when preparing very large plasmids by alkaline lysis procedure. *Biotechnology and Applied Biochemistry,* 10: 402-407.

Thomson, D. and Henry, R. (1995). Single-step protocol for preparation of plant tissue for analysis by PCR. *BioTechniques,* 19: 394-400.

Tian, Q., Wang, X., Li, C., Lu, W., Yang, L., Jiang, Y. and Luo, K. (2013). Functional characterization of the poplar R2R3-MYB transcription factor *Pto*MYB216 involved in the regulation of lignin biosynthesis during wood formation. *PloS ONE,* 8: e76369.

Tolia, N. H. and Joshua-Tor, L. (2007). Slicer and the Argonautes. *Nature Chemical Biology,* 3: 36-43.

Tsai, C., Popko, J. L., Mielke, M. R., Hu, W., Podila, G. K. and Chiang, V. L. (1998). Suppression of *O*-methyltransferase gene by homologous sense transgene in quaking aspen causes red-brown wood phenotypes. *Plant Physiology,* 117: 101-112.

Van Acker, R., Vanholme, R., Storme, V., Mortimer, J. C., Dupree, P. and Boerjan, W. (2013). Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnology for Biofuels,* 6: 46.

van Dillewijn, C. (1952) *Botany of Sugarcane*. Waltham, MA: Chronica Botanica Co.

Vanholme, R., Demedts, B., Morreel, K., Ralph, J. and Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiology,* 153: 895-905.

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

Vaucheret, H. (2006). Post-transcriptional small RNA pathways in plants: Mechanisms and regulations. *Genes and Development,* 20: 759-771.

Vettore, A. L., da Silva, F. R., Kemper, E. L. and Arruda, P. (2001). The libraries that made SUCEST. *Genetics and Molecular Biology,* 24: 1-7.

Vieira, J. and Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene,* 19: 259-268.

Voelker, S. L., Lachenbruch, B., Meinzer, F. C., Jourdes, M., Ki, C., Patten, A. M., Davin, L. B., Lewis, N. G., Tuskan, G. A., Gunter, L., Decker, S. R., Selig, M. J., Sykes, R., Himmel, M. E., Kitin, P., Shevchenko, O. and Strauss, S. H. (2010). Antisense down-regulation of *4CL* expression alters lignification, tree growth, and saccharification potential of field-grown poplar. *Plant Physiology,* 154: 874-886.

Vogel, J. (2008). Unique aspects of the grass cell wall. *Current Opinion in Plant Biology,* 11: 301-307.

Vogt, T. (2010). Phenylpropanoid biosynthesis. *Molecular Plant,* 3: 2-20.

Wagner, A., Ralph, J., Akiyama, T., Flint, H., Phillips, L., Torr, K., Nanayakkara, B. and Kiri, L. T. (2007). Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in *Pinus radiata*. *Proceedings of the National Academy of Sciences of the United States of America,* 104: 11856-11861.

Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Torr, K., Donaldson, L., Pears, L. and Ralph, J. (2011). *CCoAOMT* suppression modifies lignin composition in *Pinus radiata*. *The Plant Journal,* 67: 119-129.

Wang, H. and Dixon, R. A. (2012). On-off switches for secondary cell wall biosynthesis. *Molecular Plant,* 5: 297-303.

Wang, Z., Cui, L., Chen, C., Liu, X., Yan, Y. and Wang, Z. (2012). Downregulation of cinnamoyl CoA reductase affects lignin and phenolic acids biosynthesis in *Salvia miltiorrhiza* Bunge. *Plant Molecular Biology Reporter,* 30: 1229-1236.

Weng, J. and Chapple, C. (2010). The origin and evolution of lignin biosynthesis. *New Phytologist,* 187: 273-285.

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

Wesley, S. V., Helliwell, C., Wang, M. and Waterhouse, P. (2004). *Posttranscriptional gene silencing in plants,* in Gott, J. M. (Ed.). *Methods in molecular biology: RNA interference, editing, and modification: Methods and protocols*. Totowa, NJ: Humana Press Inc. 265, p117-129.

Wesley, S. V., Liu, Q., Wielopolska, A., Ellacott, G., Smith, N., Singh, S. and Helliwell, C. (2003). *Custom knock-outs with hairpin RNA-mediated gene silencing,* in Grotewold, E. (Eds.). *Methods in molecular biology: Plant functional genomics: Methods and protocols*. Totowa, NJ: Humana Press, Inc. 236, p. 273-286.

Wilkie, G. S., Dickson, K. S. and Gray, N. K. (2003). Regulation of mRNA translation by 5'-and 3'-UTR-binding factors. *Trends in Biochemical Sciences,* 28: 182-188.

Wilkins, O., Nahal, H., Foong, J., Provart, N. J. and Campbell, M. M. (2009). Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors. *Plant Physiology,* 149: 981-993.

Xiong, Y., Liu, T., Tian, C., Sun, S., Li, J. and Chen, M. (2005). Transcription factors in rice: A genome-wide comparative analysis between monocots and eudicots. *Plant Molecular Biology,* 59: 191-203.

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

Yang, F., Mitra, P., Zhang, L., Prak, L., Verhertbruggen, Y., Kim, J., Sun, L., Zheng, K., Tang, K., Auer, M., Scheller, H. V. and Loqué, D. (2013). Engineering secondary cell wall deposition in plants. *Plant Biotechnology Journal,* 11: 325-335.

Yang, Z., Ebright, Y. W., Yu, B. and Chen, X. (2006). HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Research,* 34: 667-675.

Yasuda, S., Fukushima, K. and Kakehi, A. (2001). Formation and chemical structures of acid-soluble lignin I: Sulfuric acid treatment time and acid-soluble lignin content of hardwood. *Journal of Wood Science*, 47: 69-72.

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

Zhang, K., Bhuiya, M., Pazo, J. R., Miao, Y., Kim, H., Ralph, J. and Liu, C. (2012). An engineered monolignol 4-*O*-methyltransferase depresses lignin biosynthesis and confers novel metabolic capability in *Arabidopsis*. *The Plant Cell,* 24: 3135-3152.

Zhang, Y., Yan, Y. and Wang, Z. (2010). The *Arabidopsis PAP1* transcription factor plays an important role in the enrichment of phenolic acids in *Salvia miltiorrhiza*. *Journal of Agricultural and Food Chemistry,* 58: 12168-12175.

Zhao, H., Wei, J., Zhang, J., Liu, H., Wang, T. and Song, Y. (2002). Lignin biosynthesis by suppression of two O-methyltransferases. *Chinese Science Bulletin,* 47: 1092-1095.

Zhao, Q. and Dixon, R. A. (2011). Transcriptional networks for lignin biosynthesis: More complex than we thought? *Trends in Plant Science,* 16: 227-233.

Zhao, X., Zhang, L. and Liu, D. (2012). Biomass recalcitrance. Part I: The chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuels, Bioproducts and Biorefining,* 6: 465-482.

Zhong, R., Lee, C., McCarthy, R. L., Reeves, C. K., Jones, E. G. and Ye, Z. (2011). Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant and Cell Physiology,* 52: 1856-1871.

Zhong, R., Lee, C., Zhou, J., McCarthy, R. L. and Ye, Z. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *The Plant Cell,* 20: 2763-2782.

Zhong, R., McCarthy, R. L., Haghighat, M. and Ye, Z. (2013). The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PloS ONE,* 8: e69219.

Zhong, R., Morrison, W. H., Himmelsbach, D. S., Poole, F. L. and Ye, Z. (2000). Essential role of caffeoyl Coenzyme A *O*-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiology,* 124: 563-577.

Zhong, R., Richardson, E. A. and Ye, Z. (2007). The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis*. *The Plant Cell,* 19: 2776-2792.

Zhou, J., Lee, C., Zhong, R. and Ye, Z. (2009). MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *The Plant Cell,* 21: 248-266.

Zhu, L., Shan, H., Chen, S., Jiang, J., Gu, C., Zhou, G., Chen, Y., Song, A. and Chen, F. (2013). The heterologous expression of the Chrysanthemum R2R3-MYB transcription factor *CmMYB1* alters lignin composition and represses flavonoid synthesis in *Arabidopsis thaliana*. *PloS ONE,* 8: e65680.

Ziebell, A., Gracom, K., Katahira, R., Chen, F., Pu, Y., Ragauskas, A., Dixon, R. A. and Davis, M. (2010). Increase in 4-coumaryl alcohol units during lignification in alfalfa (*Medicago sativa*) alters the extractability and molecular weight of lignin. *The Journal of Biological Chemistry,* 285: 38961-38968.