



**UNICAMP**

**UNIVERSIDADE ESTADUAL DE CAMPINAS**

**Instituto de Biología**

**JUAN PABLO PORTILLA LLERENA**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE FATORES DE  
TRANSCRIÇÃO QUE REGULAM A BIOSÍNTESE DE LIGNINA  
EM *SACCHARUM* SPP.**

**IDENTIFICATION AND CHARACTERIZATION OF  
TRANSCRIPTION FACTORS THAT REGULATE LIGNIN  
BIOSYNTHESIS IN *SACCHARUM* SPP.**

**Campinas**

**2021**

**JUAN PABLO PORTILLA LLERENA**

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE FATORES DE  
TRANSCRIÇÃO QUE REGULAM A BIOSÍNTESE DE LIGNINA  
EM *SACCHARUM* SPP.**

**IDENTIFICATION AND CHARACTERIZATION OF  
TRANSCRIPTION FACTORS THAT REGULATE  
LIGNIN BIOSYNTHESIS IN *SACCHARUM* SPP.**

*Tese apresentada ao  
Instituto de Biologia da  
Universidade Estadual de  
Campinas como parte dos  
requisitos exigidos para a  
obtenção do título de Doutor  
em Biologia Vegetal*

**Orientador: Prof. Dr. PAULO MAZZAFERA**

ESTE TRABALHO CORRESPONDE À  
VERSÃO FINAL DA TESE DEFENDIDA PELO  
ALUNO JUAN PABLO PORTILLA LLERENA, E  
ORIENTADA PELO PROF. DR. PAULO  
MAZZAFERA

**Campinas**

**2021**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Biologia  
Mara Janaina de Oliveira - CRB 8/6972

P836i Portilla Llerena, Juan Pablo, 1985-  
Identificação e caracterização de fatores de transcrição que regulam a biossíntese de lignina em *Saccharum* spp. / Juan Pablo Portilla Llerena. – Campinas, SP : [s.n.], 2021.

Orientador: Paulo Mazzafera.

Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. *Saccharum*. 2. Lignina. 3. Sacarificação. 4. Fatores de transcrição. 5. Parede celular. I. Mazzafera, Paulo, 1961-. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Identification and characterization of transcription factors that regulate lignin biosynthesis in *Saccharum* spp.

**Palavras-chave em inglês:**

*Saccharum*

Lignin

Saccharification

Transcription factors

Cell wall

**Área de concentração:** Biologia Vegetal

**Titulação:** Doutor em Biologia Vegetal

**Banca examinadora:**

Paulo Mazzafera [Orientador]

Renato Vicentini dos Santos

Igor Cesarino

Maria Magdalena Rossi

Elisson Antonio da Costa Romanel

**Data de defesa:** 01-10-2021

**Programa de Pós-Graduação:** Biologia Vegetal

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0001-8140-1292>

- Currículo Lattes do autor: <http://lattes.cnpq.br/6660167766284087>

Campinas, 1 de outubro de 2021

## **COMISSÃO EXAMINADORA**

Prof. Dr. Paulo Mazzafera

Prof. Dr. Renato Vicentini dos Santos

Prof.Dr. Igor Cesarino

Profa. Dra. Maria Magdalena Rossi

Prof. Dr. Elisson Antonio da Costa Romanel

*Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.*

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa (inserir o nome do programa de pós) da Unidade (inserir o nome da faculdade/instituto).

## DEDICATORIA

Dedico esta tese de Doutorado...

A minha família...

*A minha querida mãe **Fabiola**, por estar comigo sempre nos momentos bons e maus. Por me ensinar a enfrentar com diligência e sabedoria os desafios da vida e pelo apoio incondicional na minha formação acadêmica. Estou orgulhoso de você, por ser uma mulher com muita coragem e perseverança.*

*As minhas queridas irmãs: Karina, Carolina e Sofia, quanta saudade, carinho e afeto. Os momentos da vida são mais maravilhosos ao lado de vocês.*

**«Y a mí parece súbitamente un  
sueño estar frente a ustedes, y a  
mí me parece súbitamente una  
añoranza cumplida estar frente a  
ustedes, y a mí parece que quizás  
he muerto y estoy frente a  
ustedes. Yo no sé y ni quiero  
despertar, Ni odios ni rencores, ni  
gritos Ni vejámenes, perdono a  
todos los que me gritan, perdono  
a todos los que me injuriaron,  
perdono a todos los que me  
dejaron, los perdono en nombre  
del Perú, no sé a dónde me  
conduzca la vida, no sé si me  
lleve a la muerte, pero aquí estoy  
entregando todo lo que soy otra  
vez al servicio de la Patria.**

**Arriba los**

**Corazones**

**Arriba las**

**esperanzas**

**Arriba la**

**juventud**

**Viva el PERÚ**

**Viva Haya de**

**la Torre**

**Viva el**

**APRA»**

## AGRADECIMENTOS

Agradeço a Deus, Jesus Cristo e ao Espírito Santo, por ter me permitido caminhar até aqui, pela vida, por ter me dado força e paz, nos momentos difíceis.

Ao Prof. Dr. Paulo Mazzafera, pela recepção no seu laboratório e oportunidade de me orientar e me iniciar no longo caminho da ciência. Obrigado pela paciência e confiança desde o primeiro momento, pelas argumentações críticas e sugestões e por me ensinar que a ciência precisa sempre do rigor científico.

À Universidade Estadual de Campinas, particularmente ao Instituto de Biologia e ao Programa de Pós-Graduação em Biologia Vegetal, pela estrutura física para a realização desta pesquisa. A CAPES e a FAPESP com o processo 2018/25058-5 pela bolsa de doutorado concedida.

Aos membros da pre-banca e Banca, por terem aceitado o convite para participar da minha banca examinadora, e valorizado o trabalho realizado.

Aos Doutores: Eduardo Kiyota, Pedro Araújo, Rafael Gallinari, Dra. Raquel Figueiredo, Prof. Dr. Michael dos Santos Brito, Dra. Paula Nobile e Dr. Raul Lima, pela ajuda na determinação da Razão S/G, compostos fenólicos, identificação de fatores de transcrição e isolamento de promotores, desenho de construções genéticas e coleta de material transgênico...Muito obrigado amigos, pela sua ajuda decisiva !!!

Ao Prof. Dr. Jose Carlos del Rio, Dr. Jorge Rencoret e Dra. Ana Gutierrez, pela colaboração com os análises da parede celular por espectroscopia 2D HSQC NMR

Ao Prof. Dr. Igor Cesarino e ao Dr. Savio de Siqueira Ferreira, pela colaboração com os experimentos de transativação de Protoplastos, muito obrigado...pelo imenso aprendizado.

Ao Prof. Dr. Markus Pauly pela colaboração na validação da determinação do conteúdo e composição de Lignina.

À Profa. Dra. Juliana Lischka Sampaio Mayer e à Estudante de Doutorado: Laís Soêmis Sisti pela colaboração nas análises histoquímicas.

Aos todos os colegas do Departamento de Fisiologia Vegetal, em especial aos membros do LaFiMP, que de alguma forma participaram na realização deste tese de Doutorado.

## RESUMO

O bagaço de cana-de-açúcar é uma matéria prima promissória para a produção de etanol de segunda geração, porém o seu uso, apresenta desafios desde que a biomassa desta cultura é lignificada (~23%), limitando o processo de sacarificação. Um dos principais enfoques para superar a recalcitrância é a modificação biotecnológica dos reguladores (Fatores de Transcrição; TFs) da deposição de lignina, dirigido a alterar o conteúdo ou a composição da lignina. A regulação transcricional da deposição de lignina, tem sido amplamente estudada em *Arabidopsis* do que em monocotiledôneas e muito pouco ou nada se sabe em cana-de-açúcar. O presente trabalho teve como objetivo identificar e caracterizar fatores de transcrição que regulem enzimas da biossíntese da lignina em *Saccharum* spp. visando a alteração da deposição de lignina. O isolamento dos promotores de genes representativos da rota biosintética da lignina (*CAD8*, *F5H* e *COMT*) nas espécies de *Saccharum*, permitiu identificar TFs repressores e ativadores. Informações originadas a partir de uma análise de componente principal (PCA), correlação de Pearson e de redes baseianas confrontando dados de expressão de genes de biossíntese de lignina, razão S/G, conteúdo de lignina, frente ao perfil transcricional dos TFs de ambos parentais, confirmadas com ensaios de transativação de protoplastos, evidenciou em destaque que os TFs *ShMYB85* e *ShMYB58/63* poderiam possuir a potencial função como ativadores específicos da deposição de lignina nas espécies de *Saccharum*. Para corroborar estas preliminares evidências foram desenvolvidos transgênicos com silenciamento para estes TFs (*Shmyb85* e *Shmyb58/63*) em *Saccharum* spp. cultivar híbrido SP-3280. A caracterização, reforçada por análises 2D-RMN, revelou que *Shmyb85* e *Shmyb58/63* tiveram similaridades e diferenças em alterações relacionadas ao conteúdo e composição de lignina, testes histoquímicos, perfil de expressão de genes da biossíntese de lignina, e teor de compostos fenólicos. A sacarificação foi mais eficiente em plantas *Shmyb85*. Foi observado que em *Shmyb85* e *Shmyb58/63* houve alterações em mais de um polissacarídeo da parede celular (SCW), sendo *ShMYB85* um ativador mais amplo que *ShMYB58/63*, desde que o primeiro controla especificadamente o teor de pectinas. Nossos dados enfatizam as oportunidades e desafios de usar estes TFs MYBs para redesenhar a parede celular com a finalidade de se produzir bioenergia. O conhecimento gerado aqui é importante para impulsionar as práticas tradicionais de melhoramento em cana-de-açúcar e abordagens biotecnológicas para o aproveitamento de biomassa lignocelulósica no desenvolvimento do etanol de segunda geração.

**Palavras-chave:** *Saccharum* spp, lignina, sacarificação, fator de transcrição, pectina, parede celular.

## ABSTRACT

Sugarcane bagasse is a promising raw material for the production of second-generation ethanol. Still, its use presents challenges since the biomass of this crop is lignified (~23%), limiting the saccharification process. One of the main approaches to overcome recalcitrance is the biotechnological modification of the regulators (Transcription Factors; TFs) of lignin deposition, aimed at altering the content or composition of lignin. The transcriptional regulation of lignin deposition has been more studied in *Arabidopsis* than in monocots, and very little or nothing is known in sugarcane. The present work aimed to identify and characterize transcription factors that regulate lignin biosynthesis enzymes in *Saccharum* spp. The isolation of promoters of genes of the lignin biosynthetic pathway (*CAD8*, *F5H*, and *COMT*) in *Saccharum* species allowed the identification of repressor and activator TFs. Information originated from a principal component analysis (PCA), Pearson correlation and based networks comparing expression of genes of the lignin biosynthesis, S/G ratio, lignin content, against the transcriptional profile of the TFs of both parents, indicated *ShMYB85* and *ShMYB58/63* as potential regulators of lignin biosynthesis in sugarcane. Such indication was supported with protoplast transactivation assays. To have functional evidence of the role of these TFs, transgenics plants silenced for these TFs were produced using *Saccharum* spp. cultivar SP-3280 hybrid. The characterization, reinforced by 2D-NMR analysis, revealed that *Shmyb85* and *Shmyb58/63* had similarities and differences in changes related to lignin content and composition, histochemical tests, gene expression, and phenolic compounds content. Saccharification was more efficient in *Shmyb85*. It was observed that in *Shmyb85* and *Shmyb58/63* there were alterations in more than one cell wall polysaccharide (SCW), with *ShMYB85* being a broader activator than *ShMYB58/63*, since the first precisely controls the pectin content. Our data emphasize the opportunities and challenges of using these TFs MYBs to redesign the cell wall for bioenergy purposes. The knowledge generated here is important to boost traditional sugarcane breeding practices and approaches technologies for using lignocellulosic biomass in the development of second-generation ethanol.

**Keywords:** *Saccharum* spp, lignin, saccharification, transcription factor, pectin, cell wall.

## LISTA DE FIGURAS

- Figura 1 – A provável história evolutiva das espécies de *Saccharum* (Grivet et al., 2006).....29
- Figura 2 – Principais espécies do gênero *Saccharum* que apresentam características comuns, principalmente quanto ao acúmulo de sacarose e produção de fibras, tais como *S. barberi* - *S. officinarum* e *S. spontaneum* - *S. robustum*, respectivamente, para o seu uso com fins de bioenergia..... 30
- Figura 3 – Representação esquemática da biossíntese da lignina, adaptada de Cesarino et al (2006), apresentando os mais recentes descobrimentos em gramíneas representadas por as setas coloridas. A seta que representa a conversão de cafeoil-chiquimato em ácido cafeico está tracejada porque a função da *CSE* caffeoyl shikimate esterase, não foi confirmada em gramíneas por meio da genética reversa. *PAL* phenylalanine ammonia-lyase, *PTAL* bifunctional phenylalanine tyrosine ammonia-lyase, *C4H* cinnamate 4-hydroxylase, *4CL* 4-coumarate:CoA ligase, *C3H* p-coumarate 3-hydroxylase, *HCT* p-hydroxycinnamoyl-CoA:quinic acid/ shikimate hydroxycinnamoyltransferase, *CSE* caffeoyl shikimate esterase, *CCoAOMT* caffeoyl-CoA O-methyltransferase, *CCR* cinnamoyl-CoA reductase, *CAD* cinnamyl alcohol dehydrogenase, *COMT* caffeic acid O-methyltransferase, *F5H* ferulate 5-hydroxylase, *PMT* p-coumaroyl-CoA:monolignol transferase, *CHS* chalcone synthase, *CHI* chalcone isomerase, *F3'H* flavonoid 3'-hydroxylase, *OMT* O-methyltransferase, *FPGS* folylpolyglutamate synthase, *MTHFR* methylenetetrahydrofolate reductase, *PRX* peroxidase, *LAC* laccase..... 36
- Figura 4 – Redes transcricionais hierárquica que controlam a deposição da parede celular secundária (SCW) (e lignina) em *Arabidopsis* e gramíneas. Os níveis hierárquicos são destacados com cores diferentes. Os fatores de Transcrição (TFs) ortólogos entre *Arabidopsis* e gramíneas são denotados pelo mesmo cor. Em *Arabidopsis*, os master switches, NAC (NAM, ATAF1/2 e CUC2) ativam todo o programa de formação da parede celular secundária em uma maneira específica ao tipo de célula, enquanto nas gramíneas a atividade específica dos NACs para um tipo de célula é ainda pouco conhecida. TFs upstream e downstream aos NACs em gramíneas permanecem ainda pouco estudados em comparação a *Arabidopsis* (Cesarino et al.,2016)..... 38
- Figura 5 – Conteúdo de (A) lignina solúvel e (B) insolúvel nos internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significantes entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie.

Letras minúsculas diferentes indicam diferenças significantes entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes..... 45

Figura 6 – Conteúdo de açúcares solúveis e razão S/G nos internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significantes entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie. Letras minúsculas diferentes indicam diferenças significantes entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes..... 45

Figura 7 – Análises filogenéticas de sequências de aminoácidos das enzimas da via biossintética dos monolignóis, obtidas da base de dados do NCBI, GeneBank, Phytozme e homólogos identificados neste estudo para as espécies de *Saccharum*. As anotações em azul mostram os genes (e respectivas SASs) identificados por Bottcher et al. (2013)..... 47

Figura 8 – Otimização do isolamento de Promotores e Fatores de transcrição em espécies de *Saccharum*. 1: *pShCAD8*: (1A-B) 1ro round PCR das bibliotecas DL1, DL2, DL3, DL4 e Ctrl+(Pvull) . (1C-D) 2do round PCR das bibliotecas DL1(1880 bp) em *S. officinarum* (OFF) e DL1(1150 bp) e DL3(2212 bp) em *S. spontaneum* (SPON). (1E) Amplificação de DL1 (OFF), DL1 e DL3 (SPON) a partir do 2do round de PCR isolado de gel de agarose ao 1 %. 2: *pShCOMT*: 2(A-B) 2do Round de PCR usando 4 combinações de primers (1-4) a partir de DNA gnômico de OFF e SPON, 2 (1428 bp) e 4 (1277 bp) apresentaram as bandas mais intensas. 2(C-D) Avaliação de minipreps por digestão com EcoRI para verificar a presença do produto de PCR da combinação 2 em OFF e SPON. 3: *pShF5H*: (3A) Bandas de distintos tamanhos obtidas a partir de distintas combinações de primers usando DNA gnômico. 1-3 e 2-4 representam as duas bandas de maior tamanho em OFF e SPON, respectivamente. (3B-D) Serie de otimizações para a obtenção de bandas específicas de maior tamanho (1774 bp) mudando parâmetros de PCR como descritos no texto em OFF e SPON. 4: Amplificação de *ShMYB85* e *ShMYB58/63* usando como template o cDNA do 5to internódio de SPON. As bandas visualizadas para *pShCAD8*, *pShCOMT*, *pShF5H*, *ShMYB85* e *ShMYB58/63* foram separadas em gel de agarose ao 1% contendo brometo de etideo. M= Marcador Molecular; 1kb Ladder, Sinapse Inc - M1181..... 74

Figura 9 – Análise filogenética dos fatores de transcrição da família R2R3-MYB,NAC, WRKY, KNOX e AP2 de cana-de-açúcar e seus homólogos em outros organismos. As

análises foram realizadas utilizando o programa ClustalX pelo método de NeighborJoining com bootstrap =1000. A árvore foi visualizada pelo programa MEGA 5.05. As sequências homólogas de *Sorghum bicolor* (Sb), *Zea mays* (Zm), *Brachypodium distachyon* (Bd), *Panicum virgatum* (Pv), *Oryza sativa* (Os) *Populus trichocarpa* (Pt), *Arabidopsis thaliana* (At) foram obtidas do Phytozome v12.1. As marcações em cores mostram os cladogramas filogenéticos formados. As sequências marcadas com círculos negros correspondem aos homólogos identificados para cada fator de transcrição em *Saccharum* spp.(cana-de-açúcar)..... 79

Figura 10 – Perfil de expressão dos genes da via biossintética dos monolignóis analisados por qRT-PCR em espécies de *Saccharum* em relação ao gene de referência GADPH. YL= Folha Jovem, ML= Folha Madura, R3= córtex do internódio 3, R5= córtex do internódio 5, P3= medula do internódio 3, P5= medula do internódio 5 e R= Raiz. Letras diferentes indicam diferenças significativas na expressão relativa do gene entre os tecidos de um mesmo genótipo. As médias foram comparadas pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das médias de triplicatas biológicas independentes..... 89

Figura 11 – Perfil de expressão de TFs relacionados com a regulação da deposição de SCW, analisado por qRT-PCR em espécies de *Saccharum* em relação ao gene de referência GADPH. YL= Folha Jovem, ML= Folha Madura, R3= córtex do internódio 3, R5= córtex do internódio 5, P3= medula do internódio 3, P5= medula do internódio 5 e R= Raiz. Letras diferentes indicam diferenças significativas na expressão relativa do gene entre os tecidos de um mesmo genótipo. As médias foram comparadas pelo teste pos-hoc de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das médias de triplicatas biológicas independentes..... 92

Figura 12 – Análise de correlação de Pearson entre dados do perfil de expressão dos TFs identificados, a deposição de lignina (conteúdo e composição) e genes da biossíntese de lignina nas espécies de *Saccharum* sendo todos avaliados nas mesmas amostras correspondentes a 7 tecidos. Os dados variam desde 1 (100% de positiva correlação- azul escuro) até -1(100% de negativa correlação- vermelho escuro).O valor de 0 indica nenhuma relação..... 96

Figura 13 – PCA do metabolismo em espécies de *Saccharum*. Gráficos de variáveis de correlação de 11 genes da rota biossintética de lignina, 12 TFs reguladores da biossíntese da parede celular (R2R3-MYB, NAC, WRKY, KNOX e AP2), conteúdo de lignina e razão S/G de 7 tecidos de (A) *S. officinarum* e (B) *S. spontaneum*. Números em parêntese correspondem ao percentual de variação explicado pelo primeiro e segundo componente principal. As cores do texto indicam o algoritmo de agrupamento de médias usado para classificar as variáveis em 3 grupos..... 99

Figura 14 – Rede Bayesiana correspondente à expressão de Fatores de Transcrição (TFs) relacionados com a regulação da parede celular secundária, genes da biossíntese de lignina, e o conteúdo de lignina acetyl bromide de amostras de 7 tecidos em ambas as espécies de *Saccharum*. YL e ML (folha nova e velha). R3; R5 e P3; P5, corresponde a córtex e medula do terceiro e quinto internódio, respectivamente. R(raiz). O círculo em vermelho indica o conteúdo de lignina. Sof e Spo, indicam: *S. officinarum* e *S. spontaneum*. Os genes relacionados à biossíntese da lignina são: *cinamato 4-hidroxilase (C4H)*, *4-hidroxicinamoil CoA ligase (4CL)*, *hidroxicinamoil CoA: shiquimato/quinato hidroxicinamoiltransferase (HCT)*, *p-coumarato 3-hidroxilase (C3H)*, *cafeoil-CoA O-metiltransferase (CCoAOMT A e B)*, *cinamoil-CoA redutase (CCR)*, *ferulato 5-hidroxilase (F5H)*, *cafeato O-metiltransferase (COMT)* e *cinamil álcool desidrogenase (CAD A e B)*. Nas redes obtidas, foi assumido que a expressão de genes pode depender de condições (características genéticas, morfológicas e bioquímicas), mas não o contrário. Além disso, a orientação das interações regulatórias foi inferidas. Foi considerado os dados quantitativos de PCR como variáveis contínuas e outros dados (genótipo, tipo de tecido, razão S/G, conteúdo de lignina) como variáveis discretas. Para os TFs, os nomes dos genes identificados para os parentais foram baseados em sequencias caracterizadas de *Arabidopsis thaliana* e *Zea mays*..... 102

Figura 15 – Avaliação de plasmídeos pGWL7 (vetores de destino) depois da recombinação via LR clonase para verificar a presença do Promotor dos genes *SoF5H*, *SoCOMT* e *SoCAD8* em *S. officinarum* por PCR usando primers *attb1* e *attb2* em gel de agarose ao 1% com brometo de etídeo.....103

Figura 16 – Avaliação de plasmídeos p2GW7 e pGWL7 (vetores de destino) depois da recombinação via LR clonase para verificar a presença dos TFs: *ShMYB85* e *ShMYB58/63* e os Promotores dos genes *ShCAD8*, *ShCOMT* e *ShF5H* em *S. spontaneum* por PCR usando primers *attb1* e *attb2* em gel de agarose ao 1% com brometo de etídeo..... 104

Figura 17 – Ensaio de Transativação em protoplastos BY2 de *N. tabacum*. A figura mostra o fold change da ativação dos promotores *pShCOMT*, *pShF5H* e *pShCAD8* induzidos por *ShMYB58/63* e *ShMYB85* em *S. officinarum* (So) e *S. spontaneum* (Ss), em comparação da a sua atividade basal (controle negativo, gene GUS que não é Fator de transcrição (TF)). Os dados são o produto da meia de 6 replicatas ± SE. Diferentes letras minúsculas indicam diferenças na ativação de um promotor baixo interação de um TF em particular. Diferenças significativas para a ativação de *pShCOMT*, *pShCAD8* e *pSHF5H* foram determinadas pelo t-test (P<0.05)..... 106

Figura 18 – Sequência codificante completa de <i>ShMYB85</i> e <i>ShMYB58/63</i> identificadas em <i>Saccharum</i> spp. cana-de-açúcar SP-3280. A região da CDS grifada em amarelo e as setas em vermelho representar as sequencias empregadas e a extensão, respectivamente para a construção do RNAi nos TFs para o seu silenciamento.....	110
Figura 19 – Esquema da construção pUbi_hp_GAI usada como ponto de partida para o silenciamento de <i>ShMYB85</i> e <i>ShMYB58/63</i> .....	111
Figura 20 – Esquema do construção <i>Shmyb85</i> (s+as+intron) e <i>Shmyb58/63</i> (s+as+intron).....	113
Figura 21 – Amplificação por PCR convencional de <i>Shmyb85</i> e <i>Shmyb58/63</i> tanto para o lado sense como antisense visualizados em gel de agarose ao 1.5% com brometo de etideo. L=padrão de bandas 100pb, os valores da esquerda do padrão de bandas representam os tamanhos em pares de bases.....	113
Figura 22 – Verificação da presença no plasmídeo pUBI das construções (s+as+intron) tanto para <i>Shmyb85</i> como para <i>Shmyb58/63</i> depois da reação de ligação com T4 DNA ligase. As verificações foram feitas pela digestão do pUBI com enzimas de restrição BamHI + KpnI as quais digeram dois sítios específicos que flanqueiam a construção s+as+intron e por sequenciamento Sanger. L=padrão de bandas 100pb, os valores da esquerda do padrão de bandas representam os tamanhos em pares de bases.....	114
Figura 23 – A) (Lado esquerdo) Digestão com KpnI e BamHI no pUBI:Shmyb para liberar a construção <i>Shmyb85</i> (S+AS+intron) e <i>Shmyb58/63</i> (s+as+intron); (Lado direito) Digestão de pCR8/GW/TOPO com KpnI e BamHI para ligar <i>Shmyb85</i> (s+as+intron) e <i>Shmyb58/63</i> (s+as+intron) com T4 DNA ligase. B) Verificação das construções <i>Shmyb85</i> (s+as+intron) e <i>Shmyb58/63</i> (s+as+intron) no pCR8/GW/TOPO por Digestão com KpnI e BamHI. Gel de agarose 1.5 % com brometo de etideo. L=padrão de bandas 1kb, os valores da esquerda do padrão de bandas representam os tamanhos em pares de bases.....	115
Figura 24 – Esquema do Vector pCR8_hp_ <i>Shmyb85</i> (s+as+intron) e pCR8_hp_ <i>Shmyb58/63</i> (s+as+intron).....	116
Figura 25 – Digestão com KpnI e BamHI no pGVG para confirmar a presença das construções <i>Shmyb85</i> (s+as+intron) e <i>Shmyb58/63</i> (s+as+intron). Gel de agarose 1.5 % com brometo de etideo. L=padrão de bandas 1kb, os valores da esquerda do padrão de bandas representam os tamanhos em pares de bases.....	117

Figura 26 – Esquema do Vector pGVG_hp_ <i>Shmyb85</i> (s+as+intron) e pGVG_hp_ <i>Shmyb58/63</i> (s+as+intron).....	117
Figura 27 – Pipeline apresentando os passos depois da cultura de tecido e transplante de plântulas de <i>Saccharum</i> spp. cultivar híbrido SP-3280 para vasos com substrato.....	119
Figura 28 – Esquema geral do sistema de injúria estabelecido em folhas de linhagens transgênicas e controles de <i>Saccharum</i> spp. cultivar híbrido SP-3280: <i>Shmyb85</i> , <i>Shmyb58/63</i> . WT e EV são os controles. O sistema de numeração das folhas proposto, foi de acordo com Moore (1987). As Folhas (+3) injúriadas depois de 5 dias foram seccionadas na sua região média do limbo. As Folhas (+4) serviram como controles. As folhas injúriadas e não injúriadas foram processadas para a avaliação de expressão genica e deposição de lignina AcetylBromide.....	121
Figura 29 – A e B) Obtenção de Calos embriogênicos e linhagens transgênicas de <i>Saccharum</i> spp. 3280-SP silenciadas para <i>ShMYB85</i> e <i>ShMYB58/63</i> , em meio MS. C e D) Linhagens transgênicas e controles transplantados em tubetes e transferidas a estufas para aclimação. E e F) Linhagens transgênicas e controles dos tubetes transplantados a vasos de 10L baixo condições de estufa. G e H) Linhagens transgênicas e controles crescendo em vasos de 10L com 6 e 12 meses de idade respectivamente.....	123
Figura 30 – A e B) Conteúdo, C e D) grau de redução de lignina Acetyl bromide e E e F) razão S/G no 5° internódio em linhagens transgênicas de <i>Shmyb85</i> , <i>Shmyb58/63</i> e controles (EV e WT) <i>Saccharum</i> spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo de lignina e razão S/G entre os tratamentos. As medias foram comparados pelo teste de Wilcoxon, $p \leq 0.05$ . As barras verticais indicam o erro padrão das medias produto de 3-5 replicatas biológicas.....	132
Figura 31 – Seções transversais de diferentes regiões do internódio 7° em linhagens transgênicas de <i>Shmyb85</i> <i>Saccharum</i> spp. cultivar híbrido SP-3280 submetidas ao reagente Fluoroglucinol e de Maüle para detecção de lignina total, S e G, respectivamente. Colunas da esquerda: Região periférica (Rind); Colunas da direita: Região central do colmo (Pith). e: epiderme; f: fibras; fp: parênquima fundamental; mx: metaxilema; ph: Floema; px: protoxilema; vb: feixe vascular. Barras = 50 $\mu$ m.....	133
Figura 32 – Seções transversais de diferentes regiões do internódio 7° em linhagens transgênicas de <i>Shmyb58/63</i> <i>Saccharum</i> spp. cultivar híbrido SP-3280 submetidas ao reagente Fluoroglucinol e de Maüle para detecção de lignina total, S e G,	

respectivamente. Colunas da esquerda: Região periférica (Rind); Colunas da direita: Região central do colmo (Pith). e: epiderme; f: fibras; fp: parênquima fundamental; mx: metaxilema; ph: Floema; px: protoxilema; vb: feixe vascular. Barras = 50  $\mu$ m..... 134

Figura 33 – Secções transversais de diferentes regiões do internódio 7° em controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280 submetidas ao reagente Fluoroglucinol e de Maüle para detecção de lignina total. Colunas da esquerda: Região periférica (Rind); Colunas da direita: Região central do colmo (Pith). e: epiderme; f: fibras; fp: parênquima fundamental; mx: metaxilema; ph: Floema; px: protoxilema; vb: feixe vascular. Barras= 50 $\mu$ m..... 135

Figura 34 – A e B) Conteúdo de celulose C e D) hemicelulose e E e F) pectina no 5° internódio em linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo de lignina e razão S/G entre os tratamentos. As medias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das medias produto de 3-5 replicatas biológicas..... 137

Figura 35 – Porcentagem de sacarificação no 5° internódio em linhagens transgênicas de *Shmyb85* (A), *Shmyb58/63* (B) e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo de lignina e razão S/G entre os tratamentos. As medias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das medias produto de 3-5 replicatas biológicas..... 139

Figura 36 – Conteúdo de fenóis totais no 5° internódio em linhagens transgênicas de *Shmyb85* (A), *Shmyb58/63* (B) e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo de lignina e razão S/G entre os tratamentos. As medias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das medias produto de 3-5 replicatas biológicas..... 139

Figura 37 – A e B) Unidades DUALEX de Flavonoides e C e D) de Antocianinas na folha +3 em linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo das substâncias flavonóides e antocianinas. As medias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das medias produto de 3-10 replicatas biológicas.....141

Figura 38 – Expressão relativa (RT-qPCR) de genes TFs (A): *ShMYB85*; *ShMYB58/63* e da rota biossintética da Lignina (B) em folhas injuriadas e não injuriadas do controle

(WT) de *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas na expressão relativa do gene entre os tratamentos. As medias foram comparados pelo t-test,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das medias de quintuplicatas biológicas independentes..... 143

Figura 39 – Coeficiente da indução de expressão relativa (folhas injuriadas/folhas não injuriadas) de *ShMYB85*, *ShMYB58/63* e *ShCOMT* em linhagens transgênicas de *Shmyb85* e *Shmyb58/63*. Os círculos vermelhos representam a queda paralela do coeficiente em *ShMYB85* e *ShCOMT* nas linhagens transgênicas de *Shmyb85*. Os círculos azuis representam a queda paralela do coeficiente em *ShMYB58/63* e *ShCOMT* nas linhagens transgênicas *Shmyb58/63*. WT e EV são considerados como controles + e - respectivamente..... 144

Figura 40 – Coeficiente da indução de expressão relativa (folhas injuriadas/folhas não injuriadas) de *ShMYB85*, *ShMYB58/63* e coeficiente de indução da lignificação em distintas linhagens transgênicas de *Shmyb85* e *Shmyb58/63*. Os círculos vermelhos representam a queda paralela do coeficiente de expressão de *ShMYB85* e lignificação em linhagens transgênicas de *Shmyb85*. Os círculos azuis representam a queda paralela do coeficiente de expressão de *ShMYB58/63* e lignificação em linhagens transgênicas de *Shmyb58/63*. WT e EV são considerados como controles + e - respectivamente.....145

Figura 41 – Caracterização estrutural da lignina por 2D-RMN. Os espectros 2D-HSQC NMR (em DMSO-  $d_6$ ) são apresentados para as regiões aromáticas/insaturadas ( $\delta C/\delta H$  90–148/5.7–7.9) de paredes celulares das linhagens de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) em *Saccharum* spp. cultivar híbrido SP-3280. As principais estruturas encontradas são as unidades: *p*-hidroxifenil (H), guaiacil (G) e siringil (S), e os *p*-hidroxicinamatos (*p*-cumarato, *p*CA e ferulato, FA). As estruturas são coloridas para coincidir com os contornos atribuídos nos espectros de NMR. Valores integrados para cada monômero de lignina (H, G, S), *p*CA e FA (de acordo com a soma da unidades totais de lignina e carboidratos). As atribuições de sinal estão listadas na Tabela 10.....150

Figura 42 – Caracterização estrutural de carboidratos por 2D-RMN. Regiões anoméricas de carboidratos ( $\delta C/\delta H$  89-110 / 3,7-5,8) dos espectros de 2D-RMN HSQC 2D (em DMSO-  $d_6$ ) de paredes celulares das linhagens de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) em *Saccharum* spp. cultivar híbrido SP-3280. As estruturas são coloridas para coincidir com os contornos atribuídos nos espectros de NMR. As atribuições dos sinais de carboidratos estão listadas na Tabela 12..... 151

Figura 43 – Expressão relativa (RT-qPCR) e fold-change de genes TFs (A e C): *ShMYB85* e (B e D) *ShMYB58/63*, respectivamente no 5° internódio de linhagens de transgênicas de *Shmyb85* e *Shmyb58/63 Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas na expressão relativa dos genes entre os tratamentos. As médias foram comparadas pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das médias de quintuplicatas biológicas independentes.....154

Figura 44 – Expressão relativa (RT-qPCR) de genes representativos da rota biossintética de lignina, respectivamente, no 5° internódio de linhagens de transgênicas de *Shmyb85* (A) e *Shmyb58/63* (B) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas na expressão relativa de um gene particular entre os tratamentos. As médias foram comparadas pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das médias de quintuplicatas biológicas independentes.....155

## LISTA DE TABELAS

Tabela 1 - Principais elementos cis encontrados na região regulatória de <i>ShCAD8</i> nas espécies de <i>Saccharum</i> segundo PLANTCARE.....	72
Tabela 2 - Principais elementos cis encontrados na região regulatória de <i>ShCOMT</i> nas espécies de <i>Saccharum</i> segundo PLANTCARE.....	72
Tabela 3 - Principais elementos cis encontrados na região regulatória de <i>ShF5H</i> nas espécies de <i>Saccharum</i> segundo PLANT CARE.....	73
Tabela 4 - Fatores de Transcrição que interagem com elementos cis presentes na regulatória do gene <i>ShCAD8</i> nas espécies de <i>Saccharum</i> (predição segundo TFDB e PLANT PAN 3.0).....	77
Tabela 5 - - Fatores de Transcrição que interagem com elementos cis presentes na regulatória do gene <i>ShCOMT</i> nas espécies de <i>Saccharum</i> (predição segundo TFDB e PLANT PAN 2.0).....	77
Tabela 6 - Fatores de Transcrição que interagem com elementos cis presentes na regulatória do gene <i>ShF5H</i> nas espécies de <i>Saccharum</i> (predição segundo TFDB e PLANT PAN 2.0).....	78
Tabela 7 - Lista de Primers desenhados para a construção do pUBI_hp_ <i>Shmyb85</i> e pUBI_hp_ <i>Shmyb58/63</i> .....	112
Tabela 8 - Sequências dos pares de primers utilizados na amplificação por RT-qPCR e tamanho do amplicon de genes da rota biossintética da lignina e TFs em linhagens transgênicas: <i>Shmyb85</i> , <i>Shmyb58/63</i> , EV e WT de <i>Saccharum</i> spp. cultivar híbrido SP-3280.....	135
Tabela 9A - Perfil de compostos fenólicos em linhagens transgênicas de <i>Shmyb85</i> e controles (EV e WT) <i>Saccharum</i> spp. cultivar híbrido SP-3280. Letras diferentes indicam diferenças para um composto fenólico entre distintos genótipos. As medias foram comparados pelo teste de Wilcoxon, $p \leq 0.05$ . O erro padrão e produto da média de 3-5 replicatas biológicas.....	140
Tabela 9B - Perfil de compostos fenólicos no em linhagens transgênicas de <i>Shmyb58/63</i> e controles (EV e WT) <i>Saccharum</i> spp. cultivar híbrido SP-3280. Letras diferentes indicam diferenças para um composto fenólico entre distintos genótipos. As medias foram comparados pelo teste de Wilcoxon, $p \leq 0.05$ . O erro padrão e produto da média de 3-5 replicatas biológicas.....	140
Tabela 10 - Atribuições dos sinais de correlação de lignina nos espectros 2D HSQC em linhagens transgênicas de <i>Shmyb85</i> e controles (EV e WT) de <i>Saccharum</i> spp.	

cultivar híbrido SP-3280. Os sinais foram atribuídos por comparação com a literatura (del Río et al., 2012, 2015)..... 148

Tabela 11A - Análise 2D-RMN semiquantitativa em linhagens transgênicas de *Shmyb85* e controles (EV e WT) de *Saccharum* spp. cultivar híbrido SP-3280. A composição da amostra representa a abundância relativa de unidades de lignina (H, G, S), tricina (T) e p-hidroxicinamatos (pCA, FA), de acordo com a soma das unidades totais de lignina e carboidratos\*. Letras diferentes representam diferenças para um composto fenólico entre distintos genótipos. As medias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . O erro padrão e produto da média de de 3 replicatas biológicas..... 149

Tabela 11B - Análise 2D-RMN semiquantitativa em linhagens transgênicas de *Shmyb58/63* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. A composição da amostra representa a abundância relativa de unidades de lignina (H, G, S), tricina (T) e p-hidroxicinamatos (pCA, FA), de acordo com a soma das unidades totais de lignina e carboidratos\*. Letras diferentes indicam diferenças para um composto fenólico entre distintos genótipos. As medias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . O erro padrão e produto da média de 3 replicatas biológicas..... 149

Tabela 12 - Atribuições dos sinais de correlação para os carbonos anoméricos nos espectros 2D HSQC dos diferentes carboidratos identificados presentes em linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) de *Saccharum* spp. cultivar híbrido SP-3280..... 152

## SUMARIO

1 Introdução .....	26
1.1 Etanol de primeira e de segunda geração.....	26
1.2 O Gênero <i>Saccharum</i> , cana-de-açúcar e cana energia .....	27
Cana-de-açúcar .....	27
O gênero <i>Saccharum</i> .....	28
Cana Energia .....	30
1.3 Lignina.....	32
1.4 Biossíntese de Lignina .....	33
1.5 Regulação da deposição e biossíntese de Lignina: Fatores de transcrição .....	37
2 Justificativa.....	44
3 Objetivos .....	50
Objetivo Geral.....	50
Objetivos específicos:.....	50
4 Capítulos da Tese de Doutorado: Identificação e caracterização de fatores de transcrição que regulam a biossíntese de lignina em <i>Saccharum</i> spp. cultivar híbrido SP-3280 .....	51
CAPITULO I. Identificação de fatores de transcrição que regulam a biossíntese e deposição de lignina em <i>S. spontaneum</i> e <i>S. officinarum</i> .....	53
CAPÍTULO I. Identificação de fatores de transcrição que regulam a biossíntese e deposição de lignina em <i>S. spontaneum</i> e <i>S. officinarum</i> .....	54
1 Materiais e Métodos .....	54
1.1 Obtenção do Material vegetal.....	54
1.2 Isolamento, clonagem e sequenciamento da região regulatória de genes da biossíntese de lignina em espécies de <i>Saccharum</i> .....	55
Extração de DNA genômico.....	55
Isolamento dos promotores: <i>pShCAD8</i> , <i>pShCOMT</i> e <i>pShF5H</i> .....	56

Clonagem e sequenciamento dos promotores <i>pShCAD8</i> , <i>pShCOMT</i> e <i>pShF5H</i> .....	59
1.3 Análise <i>in silico</i> dos promotores dos genes da biossíntese da Lignina: <i>ShCAD8</i> , <i>ShCOMT</i> e <i>ShF5H</i> .....	59
1.4 Identificação e análise de expressão de sequências ortólogas de genes da biossíntese da lignina e de fatores de transcrição (TFs) em espécies de <i>Saccharum</i> .....	60
Extração de RNA total e síntese de cDNA.....	60
Identificação de sequências ortólogas de genes da biossíntese da lignina em espécies de <i>Saccharum</i> .....	61
Identificação de sequências ortólogas de fatores de transcrição (TFs) e análise filogenética em cana-de-açúcar.....	61
Perfil de expressão de genes da biossíntese da lignina e de fatores de transcrição (TFs) em espécies de <i>Saccharum</i> .....	62
1.5 Determinação de Lignina Acetyl Bromide (AcBr).....	63
1.6 Determinação da razão S/G.....	64
1.7 Correlação e Análise de Rede Bayesiana.....	65
1.8 Isolamento e ensaio de transativação de protoplastos em espécies de <i>Saccharum</i> .....	66
Isolamento de Fatores de Transcrição: <i>ShMYB85</i> e <i>ShMYB58/63</i> .....	66
Clonagem dos Promotores: ( <i>pShCAD8</i> , <i>pShF5H</i> e <i>pShCOMT</i> ) e Fatores de Transcrição ( <i>ShMYB85</i> e <i>ShMYB58/63</i> ).....	67
Ensaio de Transativação de Protoplastos.....	68
1.9 Análises estatísticas.....	68
2 Resultados e Discussão.....	69
2.1 Isolamento e análise <i>in silico</i> de elementos <i>cis</i> da região regulatória dos genes <i>ShCAD8</i> , <i>ShCOMT</i> e <i>ShF5H</i> em espécies de <i>Saccharum</i> .....	69
2.2 Identificação de TFs <i>in silico</i> , na região regulatória dos genes <i>ShCAD8</i> e <i>ShCOMT</i> e <i>ShF5H</i> em espécies de <i>Saccharum</i> , busca de sequências ortólogas de TFs e análise filogenética em cana-de-açúcar <i>Saccharum spp.</i> .....	75

NAC:ShNST1 e ShVNI2.....	79
ShWRKY12.....	80
ShSHN1.....	81
ShKNAT7.....	82
ShMYB46/83.....	83
MYBs:Ativadores ou repressores específicos dos polímeros da SCW.....	84
ShMYB31 e ShMYB42.....	84
ShMYB85, ShMYB58/63, ShMYB61 e ShMYB52/54.....	85
2.3 Perfil de expressão dos genes e fatores de transcrição relacionados à biossíntese de lignina em espécies de <i>Saccharum</i> .....	87
Perfil de expressão de genes da rota biossintética de lignina em espécies de <i>Saccharum</i> .....	87
Perfil de expressão de fatores de transcrição (TFs) e a sua relação com a biossíntese e deposição de lignina em espécies de <i>Saccharum</i> .....	90
2.4 Análise de componente Principal (PCA).....	97
2.5 Modulação da deposição e de genes da rota da biossíntese de lignina por Fatores de transcrição em espécies de <i>Saccharum</i> .....	99
2.6 Transativação de Protoplastos.....	103
3. Conclusões.....	107
CAPÍTULO II. Caracterização funcional de <i>Shmyb85</i> e <i>Shmyb58/63</i> em <i>Saccharum spp.SP-3280</i> .....	109
1 Material e Métodos.....	109
1.1 Construção de pGVG_hp_ <i>Shmyb85</i> (s+as+intron) e pGVG_ <i>Shmyb58/63</i> (s+as+intron).....	109
Escolha de Sequência para formar RNAi no silenciamento de <i>ShMYB85</i> e <i>ShMYB58/63</i> .....	109
Estratégias de Clonagem e Desenho de Primers.....	111
1.2 Transformação do híbrido SP-3280 de cana-de-açúcar ( <i>Saccharum spp.</i> ) para a obtenção de linhagens transgênicas silenciadas ( <i>Shmyb85</i> e <i>Shmyb58/63</i> ) em <i>ShMYB85</i> e <i>ShMYB58/63</i> .....	118

1.3 Material vegetal e condições de crescimento.....	119
1.4 Injúria (Indução da lignificação) e amostragem.....	120
1.5 Caracterização funcional de linhagens transgênicas silenciadas ( <i>Shmyb85</i> e <i>Shmyb58/63</i> ) para <i>ShMYB85</i> e <i>ShMYB58/63</i> .....	121
Análises histoquímicas.....	121
Análises bioquímicas.....	124
Determinação de Lignina Acetyl Bromide (AcBr).....	124
Determinação da razão S/G.....	124
Determinação dos polissacarídeos da parede celular.....	124
Determinação da porcentagem de sacarificação .....	125
Determinação de compostos fenólicos solúveis totais .....	125
Análise por espectroscopia 2D-NMR.....	126
Análise de Expressão por RT-qPCR .....	127
Análises estatísticas .....	128
2. Resultados .....	129
2.1 Caracterização funcional de linhagens transgênicas silenciadas ( <i>Shmyb85</i> e <i>Shmyb58/63</i> ) para <i>ShMYB85</i> e <i>ShMYB58/63</i> .....	129
Determinação do Conteúdo de Lignina Acetil bromide e Razão S/G.....	129
Análises histoquímicas.....	130
Determinação dos polissacarídeos da parede celular.....	136
Determinação da porcentagem de sacarificação.....	138
Determinação e perfil de compostos fenólicos solúveis totais.....	138
Determinação de Flavonoides e Antocianinas.....	141
Indução da Lignificação (injúria).....	142
Análise por espectroscopia 2DNMR.....	146
Expressão de Fatores de Transcrição: <i>ShMYB85</i> e <i>ShMYB58/63</i> .....	152
Expressão de genes representativos da rota biossintética da lignina.....	152
3. Discussão.....	156

4. Conclusão.....	164
5. Referencias.....	165
6. Material suplementar.....	188
7. Anexos.....	221

# 1 INTRODUÇÃO

## 1.1 Etanol de primeira e de segunda geração

Os rápidos crescimentos populacionais, bem como a industrialização global, exigem fornecimento contínuo e constante de energia ( Li et al., 2008; Yuan et al., 2008). A redução das reservas de petróleo, o aquecimento global devido à emissão de gases de efeito estufa e outros problemas associados com o uso de combustíveis fósseis tem motivado a pesquisa dirigida ao desenvolvimento de fontes alternativas de energia, de preferência renovável, para a crescente demanda por energia em todo o mundo (Hisano et al., 2009; Waclawovsky et al., 2010). O principal biocombustível alternativo é o bioetanol, com uma produção anual que tem aumentado de 17 a 86 x10<sup>6</sup> m<sup>3</sup>, no período de 2000 a 2011 (Ren, 2012), seguido do biodiesel, com uma produção anual de 21.4 x 10<sup>6</sup> m<sup>3</sup> em 2011.

O bioetanol pode ser produzido a partir do amido e da sacarose, denominado etanol de primeira geração (E1G), e biomassa lignocelulósica, chamado de etanol de segunda geração (E2G). O etanol do amido é produzido principalmente a partir do milho e, da sacarose, principalmente da cana-de-açúcar. Mais 10 bilhões de galões de E1G são produzidos no mundo a cada ano, através de processos industriais especializados, incluindo a hidrólise e fermentação (Goldemberg, 2007; Rass-Hansen et al., 2007). Embora a produção de E1G represente a opção mais conveniente e tecnicamente avançada, alguns tem sugerido que, a longo prazo, pode ocorrer uma competição entre as áreas destinadas ao plantio da cana-de-açúcar e culturas alimentícias (De Oliveira et al., 2005; Farrell et al., 2006). Para evitar esta situação no futuro, especialmente em locais que não dispõem de um clima favorável ou uma suficiente extensão territorial para cultivo, torna-se necessário investir no desenvolvimento de tecnologias para a produção de E2G. A tecnologia do E2G envolve o uso da biomassa vegetal, que pode servir como matéria-prima renovável para a produção de energia, podendo ser transformada em biocombustíveis através de um processo chamado sacarificação. Assim, a biomassa remanescente de uma cultura, a biomassa lignocelulósica, como por exemplo o bagaço de cana-de-açúcar, seria uma grande fonte de açúcares para a produção etanol (Carroll & Somerville, 2009; Simmons et al., 2008; Somerville et al., 2010). Atualmente, apenas insignificantes quantidades de E2G são produzidos em unidades industriais ao redor

do mundo, indicando que ainda esta tecnologia não está comercialmente factível, sendo que as limitações são tanto econômicas e técnicas, estas últimas decorrentes de problemas na otimização do processo (Lennartsson et al., 2014). O principal desafio técnico está relacionado à recalcitrância da lignina, isto é, os polissacarídeos da parede celular não são prontamente disponíveis ao processo de sacarificação porque a lignina liga-se à celulose e à hemicelulose, protegendo-as fisicamente (Anderson & Akin, 2008; Buckeridge et al., 2008). Esta situação faz com o acesso à celulose por enzimas hidrolíticas seja dificultado assim como o tempo de degradação da biomassa seja longo, elevando o custo do processo (Soccol et al., 2010). Para resolver este problema é necessário submeter à matéria prima a tratamentos químicos prévios, removendo a lignina, reduzindo a cristalinidade da celulose e aumentando a porosidade do material, facilitando assim o processo de sacarificação. Embora os pré-tratamentos químicos existentes sejam relativamente eficientes para decompor a parede celular, existe a conseqüente geração de resíduos químicos que, além de serem potencialmente poluentes, produzem produtos que inibem a fermentação dos açúcares liberados, tais como os furfurais (Buckeridge et al., 2010). Alternativamente, para superar o problema do pré-tratamento com produtos químicos, tem sido proposto o desenvolvimento de plantas com menor recalcitrância, seja por melhoramento genético ou pela produção de plantas geneticamente modificadas, com alteração no conteúdo ou a composição de lignina através da super-expressão ou silenciamento envolvidos na biossíntese, polimerização e regulação deste polímero (Araújo et al., 2019; Handakumbura & Hazen, 2012; Poovaiah et al., 2014; Ralph et al., 2006; Chen and Dixon, 2007a).

## **1.2 O Gênero *Saccharum*, cana-de-açúcar e cana energia**

### **Cana-de-açúcar**

A cana-de-açúcar (*Saccharum* spp.) é altamente eficiente na conversão de energia solar em fotoassimilados, podendo acumular até 18% de sacarose nos colmos (Moore, 1995), sendo a matéria prima para a produção de açúcar e bioetanol (de Souza et al., 2014; Jung et al., 2012). O substrato rico em sacarose obtido pelo processamento dos colmos é usado para a fermentação alcoólica e produção de E1G (de Souza et al., 2014; Jung et al., 2012). A biomassa obtida após a extração do caldo rico em sacarose, o bagaço, é constituída principalmente por celulose (39%),

hemicelulose (25%) e lignina (23%) e apresenta um enorme potencial para a produção de E2G (Carroll & Somerville, 2009; Rezende et al., 2011; Szczerbowski et al., 2014). Todos esses componentes são geneticamente controlados e laboriosos de serem estudados em cana-de-açúcar, devido ao genoma complexo dos híbridos oriundos dos eventos de hibridação e poliploidização entre as espécies de *Saccharum*, principalmente de *S. officinarum* e *S. spontaneum*, contribuindo 90% e 10%, respectivamente para o genótipo híbrido de cana-de-açúcar (Ming et al., 2006).

### **O gênero *Saccharum***

O gênero *Saccharum* compreende mais de 10 espécies, uma vez que passou a englobar espécies do gênero *Erianthus*, já que são botanicamente próximos (Cheavegatti-Gianotto et al., 2011; Grivet et al., 2006) (Figura 1). Até o momento alguns estudos sistemáticos relacionado à biossíntese da lignina e a bioquímica da parede celular tem sido feito com espécies de *Saccharum* (Ferreira et al., 2016; Llerena et al., 2019), e com híbridos comerciais disponíveis (Bottcher et al., 2013; Santos et al., 2015 a,b). Algumas espécies do gênero apresentam características comuns quanto ao acúmulo de sacarose e produção de fibras, tais como *S. spontaneum*, *S. officinarum*, *S. robustum* e *S. barberi* (Figura 2) *S. officinarum* e *S. spontaneum* são contrastantes em relação ao teor de sacarose e de fibra, sendo que a primeira acumula mais sacarose mas tem menor teor de fibras. *S. officinarum* é a única espécie dentro do gênero *Saccharum* cujo número cromossômico não é variável entre os indivíduos (Price, 1963) e acredita-se que ela é originada a partir de *S. robustum*. Por outro lado, *S. spontaneum* é uma espécie complexa, altamente polimórfica, e a mais primitiva das espécies do gênero *Saccharum*. Possui uma imensa variabilidade genética que pode contribuir com características de interesse agrônomo aos programas de melhoramento genético (Naidu & Sreenivasan, 1987) que visam o desenvolvimento de variedades comerciais com grande potencial para a produção de biomassa (Roach, 1978). Diversas evidências moleculares indicam que *S. spontaneum* é muito diferente geneticamente em relação às outras espécies de *Saccharum* (Burnquist et al., 1992; Vijayan Nair et al., 1999). De maneira similar a *S. spontaneum*, plantas de *S. robustum* possuem colmos ricos em fibras e pobres em sacarose, e embora as plantas sejam extremamente vigorosas, são susceptíveis a estresses abióticos e bióticos (Stevenson, 1965). Embora haja interesse potencial na utilização de *S. robustum* em programas de melhoramento genético devido ao seu

grande vigor, há poucas citações da utilização da espécie, exceto nos híbridos produzidos pelo programa de melhoramento no Havaí (Naidu & Sreenivasan, 1987; Roach & Daniels, 1987). A espécie *S. barberi* surgiu da hibridação natural de *S. officinarum* com *S. spontaneum* (D'Hont et al., 2002). Esta espécie já foi cultivada e possui um conteúdo moderado de sacarose, com resistência a estresses e alto conteúdo de fibras em relação a *S. officinarum*. Atualmente não existe grande interesse no seu uso no melhoramento genético, principalmente devido à dificuldade de florescimento e à esterilidade.

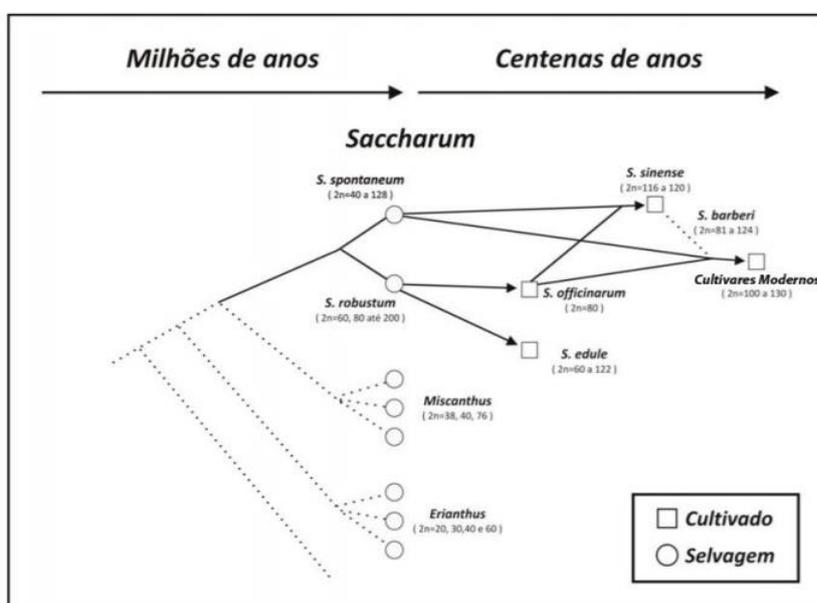


Figura 1 – A provável história evolutiva das espécies de *Saccharum* (Grivet et al., 2006).

***S. officinarum* (2n=80)**

- Colmos largos com alto conteúdo de sacarose e baixo conteúdo de fibra
- Baixa resistência a estresses e pouco perfilhamento
- Número cromossômico estável entre indivíduos

***S. barberi* (2n=81-124)**

- Colmos largos com moderado conteúdo de fibra e sacarose.
- Resistência moderada a estresses bióticos e abióticos
- Dificuldade no florescimento-esterilidade

***S. spontaneum* (2n= 36-128)**

- Colmos estreitos com alto conteúdo de fibra e baixo de sacarose .
- Espécie altamente polimórfica, variável e complexa.
- Grande capacidade de perfilhamento e resistência a estresses.

***S. robustum* (2n=60-80 ate 200)**

- Colmos vigorosos com baixo conteúdo de sacarose e alto de fibras.
- Resistência a estresses bióticos e abióticos.
- Pouco perfilhamento

Figura 2 – Principais espécies do gênero *Saccharum* que apresentam características comuns, principalmente quanto ao acúmulo de sacarose e produção de fibras, tais como *S. barberi* - *S. officinarum* e *S. spontaneum* - *S. robustum*, respectivamente, para o seu uso com fins de bioenergia.

### **Cana Energia**

A maioria dos programas de melhoramento genético em cana-de-açúcar utilizam o retrocruzamento dos híbridos com espécies ancestrais de *S. officinarum*, devido ao alto teor de sacarose e baixo conteúdo de fibras (Carvalho-Netto et al., 2014). Mais recentemente tem se dado atenção ao que se conhece por Cana Energia, uma forma distinta de cultura de cana-de-açúcar destinada para a produção de biomassa, ficando a produção de sacarose em segundo plano (Carvalho-Netto et al., 2014), com foco no aumento de alelos de *S. spontaneum* (alto teor de fibra e baixo de sacarose) nos híbridos produzidos. O melhoramento genético para o desenvolvimento da cana energia, conferindo alta produtividade de biomassa, digestibilidade da parede celular, excepcional crescimento em áreas com pouca administração de nutrientes, perfilhamento e resistência a doenças, também pode ser aproveitado a partir do potencial existente na ampla diversidade genética das espécies ancestrais do gênero *Saccharum* e outros gêneros relacionados, tais como *Erianthus*, *Miscanthus*, *Narenga* e *Sclerostachya*. Recentes esforços do melhoramento estão tomando distintas vias, indicando que pode-se desenvolver dois tipos de cana energia. A cana energia de tipo

I, que possui característica de cultura “dual”, com conteúdo de sacarose próximo ou comparável às variedades híbridas produtores de sacarose, mas com muito mais fibra, e a cana energia de tipo II, desenvolvida apenas para a produção de biomassa, com alto teor de fibra, na qual a quantidade de sacarose é moderada (Tew & Cobill, 2008). A cana energia pode ser utilizada para obtenção de calor, vapor, eletricidade e biocombustíveis (etanol de primeira e segunda geração) o que é favorecido pelo seu maior teor de lignina, que possui maior poder calorífico que os outros componentes da parede celular (Van Heiningen, 2006;Tew & Cobill, 2008). Entretanto, o maior teor de lignina traz desvantagens quando se considera o seu uso para a produção de E2G

Entre as perspectivas futuras em relação à cana energia temos:

- Por causa da elevada produtividade de biomassa, a cana energia tem o potencial de produzir maiores quantidades de açúcares celulósicos por hectarea que as variedades híbridas de cana-de-açúcar, convencionais. Caso isto se cumpra, a cana energia pode substituir, em um futuro próximo, uma parte significativa das áreas destinadas para a cultura da cana-de-açúcar, mesmo aquelas destinadas para a indústria de E2G. Isso só será possível através da superação dos desafios no processamento da cana energia e o fornecimento de um conjunto de variedades especificamente desenvolvidos para este propósito, é dizer etanol de primeira ou segunda geração, vapor ou eletricidade e/ou açúcares celulósicos para a produção bioquímica.
- Por ser mais rústica, a cana energia pode ser plantada em terra com pouca necessidade de nutrientes, ou áreas de solo e climas piores do que aqueles reservados para a produção de alimentos ou cana-de-açúcar, convencional. A cultura de cana energia exige menor uso de fertilizantes, herbicidas e pesticidas.
- A cana energia pode produzir mais perfilhos e possui um coeficiente de multiplicação superior (1:30 ou mais) contra uma taxa menor (1:10) encontrada nas variedades híbridas atuais.
- A safra de cana energia será importante para a contenção da erosão do solo e na recuperação de solos degradados, pelo seu vigoroso sistema radicular na forma de rizoma, com raízes fortes e fasciculadas.

### 1.3 Lignina

A lignina é um heteropolímero complexo, formado pelo acoplamento oxidativo combinatório de três álcoois que são sintetizados no citoplasma das células vegetais: *p*-cumaril, coniferil e sinapil. Estes álcoois diferem no grau de metoxilação (Boerjan et al., 2003) sendo transportados do citoplasma para o apoplasto, onde são oxidados por peroxidases e/ou laccases, a fim de serem incorporados no polímero pré-formado (Wang et al., 2013). Após a incorporação, estes monolignóis passam a ser denominados *p*-hidroxifenil (H), guaiacil (G) e siringil (S), respectivamente, e a sua contribuição para a composição monomérica da lignina varia significativamente entre os tipos de células, tecidos e plantas (Bonawitz & Chapple, 2010; Cesarino et al., 2012). Outros fenilpropanoides também podem ser incorporados à estrutura da lignina, em diferentes níveis (Raes et al., 2003). Uma recente caracterização estrutural de paredes celulares de monocotiledôneas mostrou que o flavonóide tricina faz parte da lignina nativa (del Río et al., 2015; Lan et al., 2015), sendo este monômero atuante na formação do sítio de nucleação para início da biossíntese de lignina (Cesarino et al., 2016; Eloy et al., 2016; Lan et al., 2015)

A lignina presente nas gimnospermas é composta por unidades G e pequenas quantidades de H, enquanto que em angiospermas são compostas por unidades G, S e traços de H. Em monocotiledôneas, tanto as unidades S e G se apresentam em níveis similares e a quantidade de unidades H é maior que em eudicotiledôneas (Vanholme et al., 2010). O conteúdo e composição da lignina em monocotiledôneas é regulada diferentemente em paredes celulares primárias e secundárias, diferentes tecidos e diversos estresses (Grabber et al., 2004; Moura et al., 2010). Uma característica geral em internódios jovens de gramíneas é que a lignificação começa antes do término do processo de alongamento (Bottcher et al., 2013; Chen et al., 2002; Jung & Casler, 2006; Shen et al., 2009). Os tecidos lignificados são o protoxilema e metaxilema, possuindo principalmente unidades H e G, com poucas unidades S (Grabber, 2005; H. G. Jung & Engels, 2002). O acúmulo de lignina cresce com a maturidade, onde a deposição deste polímero é encontrada na epiderme, hipoderme, fibras do esclerênquima, xilemáticas e parênquima de armazenamento que são enriquecidos com lignina S (Casu et al., 2007; Cesarino et al., 2012; Jacobsen et al., 1992). Quantidades significativas de ácidos hidroxicinâmicos (ferulato e *p*-coumarato) são encontrados associados com paredes celulares de gramíneas (Molinari et al.,

2013; Ralph et al., 1995). Ambos hidroxicinamatos, especialmente o ferulato, permitem o entrelaçamento entre lignina e hemiceluloses, enquanto que *p*-coumarato principalmente acila as unidades S e a sua acumulação está associada com a lignificação e com a maturidade da parede celular.

#### 1.4 Biossíntese de Lignina

O estudo da biossíntese e deposição da lignina na parede celular vegetal é um dos principais focos de pesquisa na área biocombustíveis. A lignina é amplamente estudada em plantas modelos como *Arabidopsis thaliana*, *Nicotiana tabacum* e *Populus trichocarpa* (Zhong & Ye, 2009). Várias espécies de plantas foram modificadas geneticamente com a finalidade de reduzir ou alterar a composição ou o teor da lignina (Vanholme et al., 2010, 2012). Modificações genéticas no conteúdo e composição de lignina dependem da posição da enzima na via biossintética e do perfil de expressão - super-expresssão e/ou silenciamento - de genes envolvidos no processo de lignificação (Poovaiah et al., 2014) (Figura 3). Alterações nas expressões de *C3H*, *HCT*, ou *4CL* levam a mudanças quantitativas nos níveis de lignina, enquanto que a regulação de *F5H* e *COMT* acarreta alterações na relação S/G e, por conseguinte, no tipo de lignina (Chen & Dixon, 2007; Li et al., 2003; Ralph et al., 2006). A recente identificação de outro gene da biossíntese da lignina, *Caffeoyl Shikimate Esterase* (*CSE*), adiciona outro passo nesta rota metabólica que pode ser manipulado (Ha et al., 2016; Vanholme et al., 2013). Além disso, a relação S/G e das ligações inter-monoméricas no polímero da lignina são características importantes para prever o grau e natureza da condensação do polímero e, conseqüentemente, sobre a recalcitrância da biomassa vegetal (Kiyota et al., 2012). As vias biossintéticas dos monômeros da lignina são relativamente conservadas entre as distintas espécies e estudos de eudicotiledôneas seriam válidos para serem aplicadas em monocotiledôneas para etanol de 2G (Fu et al., 2011; Jung & Engels, 2002; Zhou et al., 2009). Embora os estudos com genes da rota biossintética de lignina e qualidade de biomassa estiveram focados em eudicotiledôneas (por exemplo, *Alfalfa* e *Populus*) e plantas modelo como *Arabidopsis* e *Nicotiana* (Zhou et al., 2009), poucos foram executados em monocotiledôneas. O estudo sistemático da biossíntese de lignina em cana-de-açúcar tem sido amplamente explorados na identificação e caracterização dos genes da via metabólica (Ramos et al., 2001; Papini-Terzi et al., 2005; Bottcher et al., 2013; Cesarino et al., 2013; Santos et al., 2015a,b; Vicentini et al., 2015;

Ferreira et al., 2016) porém muitos aspectos do metabolismo deste polímero permanecem mal compreendidos ou são desconhecidos nesta cultura (Cesarino et al., 2016; Bewg, 2015)

As enzimas CAD (*cinnamyl alcohol dehydrogenase*), F5H (*ferulate 5-hydroxylase*) e COMT (*caffeic acid O-methyltransferase*) são chaves no conteúdo e a composição da lignina, por estarem posicionadas no final da rota de biossíntese de lignina. O passo final na biossíntese dos monolignóis é catalisado pela CAD e o seu envolvimento na biossíntese de lignina tem sido estudado em diferentes espécies: *Populus deltoides* X *Populus nigra*, *Eucalyptus camaldulensis*, *Panicum virgatum*, *Zea mays* e *Brachipodium distachyon* (Bouvier d'Yvoire et al., 2013; Chabannes et al., 2001; Fornalé et al., 2012; Jackson et al., 2008; Lapierre et al., 2004; Ralph et al., 2001; Saathoff et al., 2011; Valério et al., 2003). A redução na atividade de CAD resultou em uma redução do conteúdo de lignina sem afetar a produção de biomassa (Poovaiah et al., 2014). Modificação genética da F5H e COMT provocam alterações na razão S/G na composição da lignina (Li et al., 2003; Ralph et al., 2006; Chen & Dixon, 2007b). O silenciamento de COMT em cana-de-açúcar levou a uma redução significativa da subunidade S da lignina (Jung et al., 2012; Wu et al., 2019; Bewg et al., 2016) Em coníferas, a subunidade S não é presente, mas quando transformadas com F5H a subunidade S é produzida e quando co-transformada com COMT obteve-se sua incorporação no polímero da lignina em elementos traqueais (Wagner et al., 2015). O mutante de *Arabidopsis fah1*, deficiente na atividade F5H, apresenta teores residuais de unidades S, de forma contrária super-expressão de *F5H* no mutante *fah1* resulta em baixos teores de deposição de Lignina G, mas altos teores de S comparando com o WT (Franke et al., 2000; Meyer et al., 1998; Stewart et al., 2009). O ortólogo de *AtF5H: PoF5H (Populus trichocarpa)* foi capaz de complementar o fenótipo mutante de *fah1* e incrementou o teor de subunidades S no WT (Sibout et al., 2002). Similares trabalhos alterando a razão S/G por super-expressão de *F5H* para *Populus* são os descritos por (Franke et al., 2000; Huntley et al., 2003; Jiang et al., 2020; Stewart et al., 2009). De outro lado um análise da parede celular de Tabaco e Alfalfa super-expressando *F5H*, revelou um significativo aumento da razão S/G, respectivamente (Franke et al., 2000; Reddy et al., 2005). A função de F5H permanece bastante elusiva em monocotiledôneas e desconhecida em cana-de-açúcar, porém, alguns esforços foram feitos em monocotiledôneas em ordem a identificar,

caracterizar e validar a super-expressão da *F5H* neste grupo de plantas. Em linhagens transgênicas de COMT-RNAi de *Panicum virgatum*, a super-expressão de *PvF5H*, diminuiu as unidades G, aumentando as subunidades 5-OH, enquanto a deficiência das subunidades S foi parcialmente ou totalmente restaurada dependendo do grau da super-expressão de COMT (Wu et al., 2019) Em arroz a super-expressão de *OsCAld5H1* (*CYP84A5*) resultou em ligninas enriquecidas em unidades S como foi revelado por análise estrutural NMR (Takeda et al., 2017). A super-expresssão (*SbF5H*) em sorgo, aumentou tanto a Lignina S e proporção da razão S/G, enquanto o crescimento e desenvolvimento da planta permaneceram relativamente inalterados (Tetreault et al., 2020)

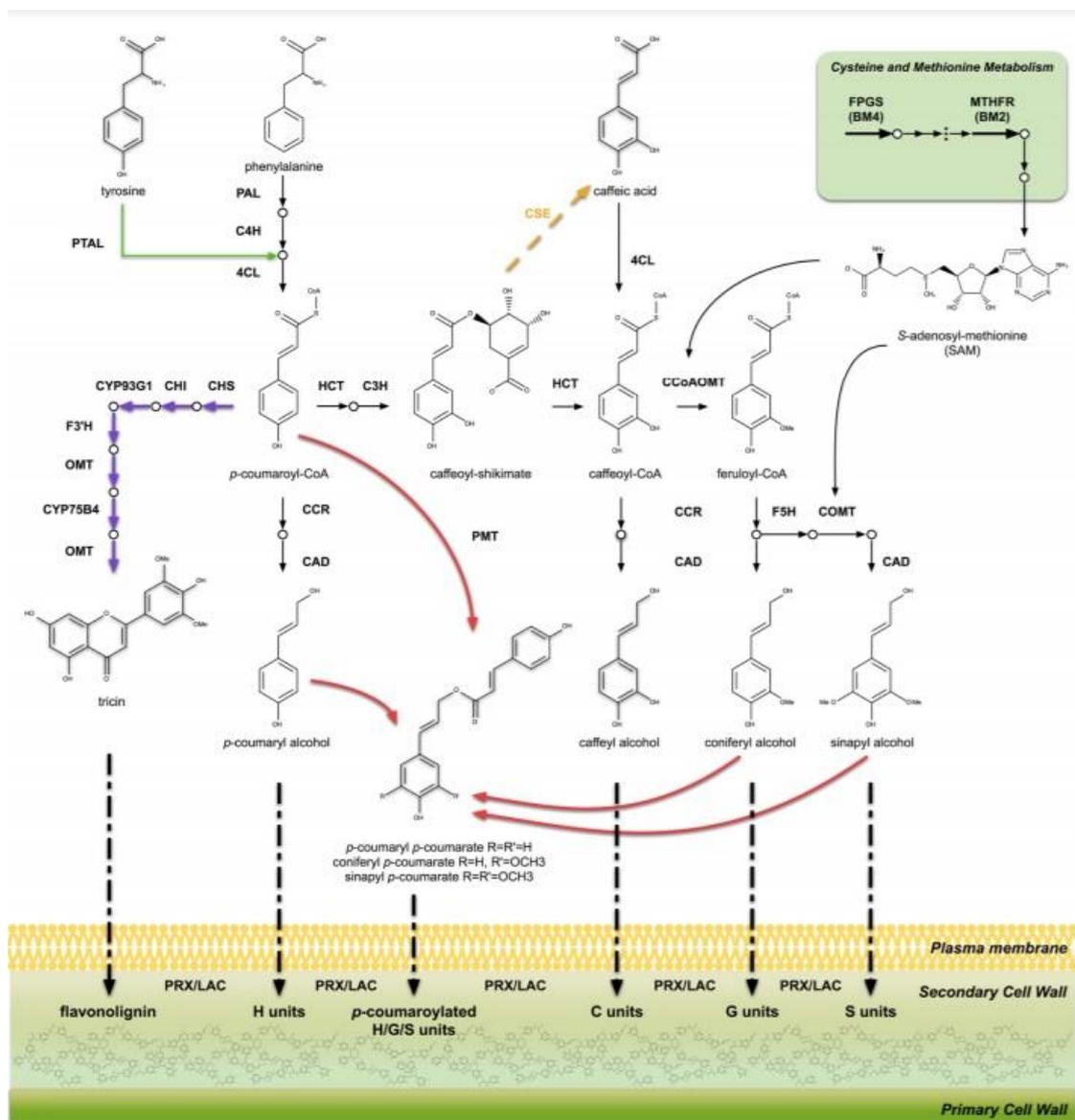


Figura 3 – Representação esquemática da biossíntese da lignina, adaptada de Cesarino et al (2006), apresentando os mais recentes descobrimentos em gramíneas representadas pelas setas coloridas. A seta que representa a conversão de cafeoil-chiquimato em ácido cafeico está tracejada porque a função da *CSE* (cafeoyl shikimate esterase), não foi confirmada em gramíneas por meio da genética reversa. *PAL* phenylalanine ammonia-lyase, *PTAL* bifunctional phenylalanine tyrosine ammonia-lyase, *C4H* cinnamate 4-hydroxylase, *4CL* 4-coumarate:CoA ligase, *C3H* p-coumarate 3-hydroxylase, *HCT* p-hydroxycinnamoyl-CoA:quinic/shikimate hydroxycinnamoyltransferase, *CSE* (cafeoyl shikimate esterase), *CCoAOMT* caffeoyl-CoA O-methyltransferase, *CCR* cinnamoyl-CoA reductase, *CAD* cinnamyl alcohol dehydrogenase, *COMT* caffeic acid O-methyltransferase, *F5H* ferulate 5-hydroxylase, *PMT* p-coumaroyl-CoA:monolignol transferase, *CHS* chalcone synthase, *CHI* chalcone isomerase, *F3'H* flavonoid 3'-hydroxylase, *OMT* O-methyltransferase, *FPGS* folypolyglutamate synthase, *MTHFR* methylenetetrahydrofolate reductase, *PRX* peroxidase, *LAC* laccase.

## 1.5 Regulação da deposição e biossíntese de Lignina: Fatores de transcrição

A coordenação da deposição da SCW e por tanto da lignina é concomitantemente regulada por fatores de transcrição (TF), modulando a expressão das enzimas catalíticas e alterando a composição e conteúdo desse componente na SCW (Figura 2). Uma complexa rede hierárquica, extensa e de vários níveis de Fatores de Transcrição (TFs) (Figura 4), ativam ou reprimem genes da biossíntese da lignina, ligando-se aos promotores diretamente via SNBE (Elemento de ligação NAC de parede secundária) e/ou SMRE (Elemento responsivo a MYB de parede secundária) (Zhong et al., 2010b; Zhong & Ye, 2012). Os estudos associados à interação entre TFs e elementos cis de promotores regulando a expressão de genes relacionados a SCW em cana-de-açúcar até o momento não tem sido descritos, desde que o isolamento e caracterização de promotores ativos em cana-de-açúcar de maneira geral é muito laboriosa, uma vez que o genoma complexo desta monocotiledônea, altamente poliploide e aneuploide, contendo sequências inativas (Ming et al., 2006). Por outro lado, tal complexidade dificulta a montagem de um genoma de referência e demanda metodologias para análise de dados específicas, configurando-se um desafio para pesquisas genômicas (Hoang et al., 2015), sendo assim que a disponibilidade de promotores ativos para cana-de-açúcar para efeitos de modificação biotecnológica é reduzida em comparação a outras espécies vegetais de interesse agrícola (Kinkema et al., 2014).

Muitos fatores de transcrição envolvidos diretamente ou indiretamente na regulação da biossíntese da lignina tem sido identificados e caracterizados, principalmente em *Arabidopsis* (Gray et al., 2012; McCann & Carpita, 2008; Vogel, 2008; Zhao & Dixon, 2011), porém a informação no que respeita a gramíneas é ainda limitada (Bhatia et al., 2019) e muito pouco compreendida em cana-de-açúcar. (Brito et al., 2015; Martins et al., 2018; Poovaiah et al., 2016). Assim, distinta composição de SCW permite inferir distinto controle transcricional da Lignificação entre eudicotiledôneas e monocotiledôneas (Handakumbura & Hazen, 2012). TFs Ortólogos de eudicotiledôneas regulando a SCW foram identificados no genoma de gramíneas e o perfil de expressão tem ajudado a identificar possíveis reguladores candidatos implicados na deposição do SCW em algumas gramíneas. Até o momento poucos TF

foram caracterizados funcionalmente como máster switch, capaz de ativar a biossíntese os principais componentes da SCW em gramíneas. Uma das primeiras tentativas foi usar sistemas heterólogos para avaliar a função de um grupo de TFs NAC e MYB de arroz, milho e outras gramíneas na regulação da deposição da SCW (Cesarino et al., 2016).

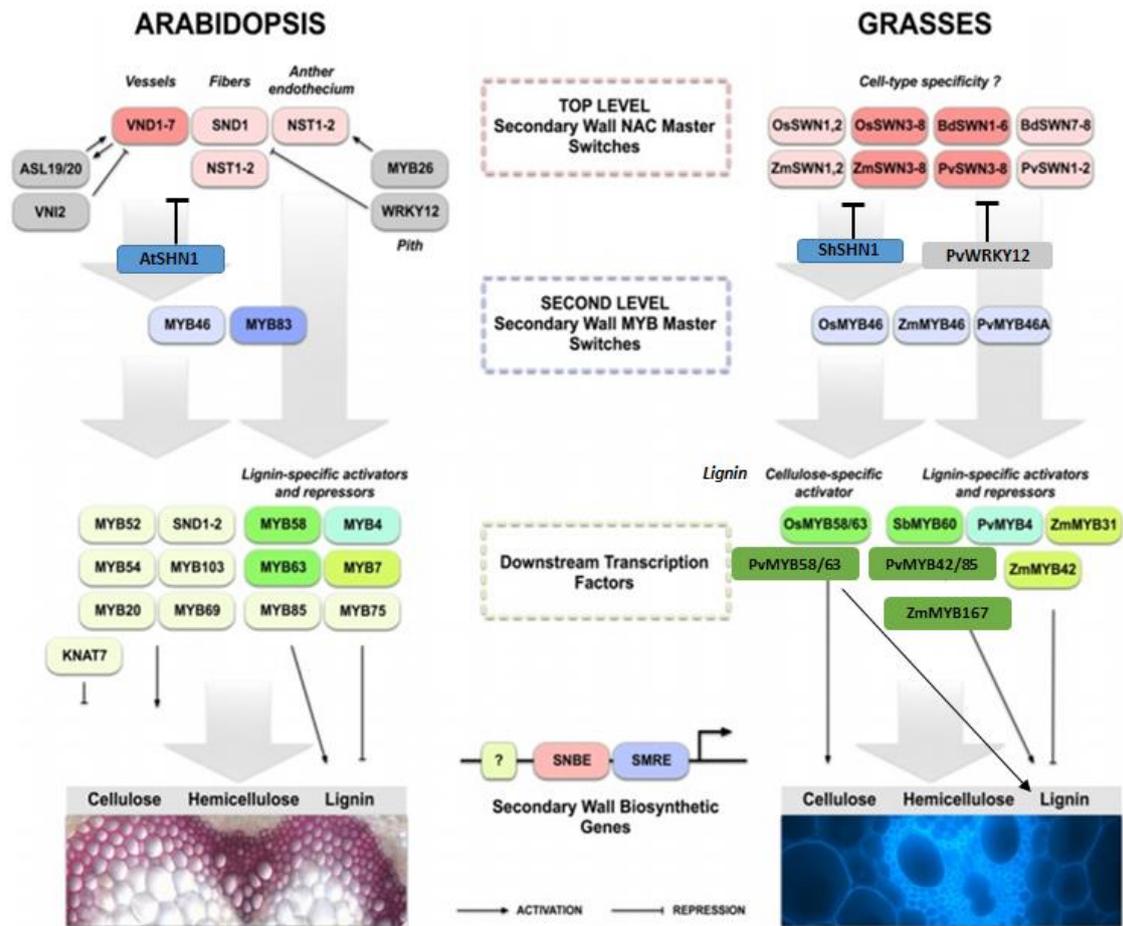


Figura 4 – Redes transcricionais hierárquica que controlam a deposição da parede celular secundária (SCW) (e lignina) em Arabidopsis e gramíneas. Os níveis hierárquicos são destacados com cores diferentes. Os fatores de Transcrição (TFs) ortólogos entre Arabidopsis e gramíneas são denotados pela mesma cor. Em Arabidopsis, os *master switches*, NAC (NAM, ATAF1/2 e CUC2) ativam todo o programa de formação da parede celular secundária em uma maneira específica ao tipo de célula, enquanto nas gramíneas a atividade específica dos NACs para um tipo de célula é ainda pouco conhecida. TFs *upstream* e *downstream* aos NACs em gramíneas permanecem ainda pouco estudados em comparação a Arabidopsis. Adaptado de Cesarino et al. (2016).

Em *Arabidopsis*, o TF AtSHN (SHINE/WAX INDUCER (SHN/WIN)), sub-clado da família do fator de transcrição AP2/ERF é considerado um máster switch e foi envolvido primeiramente na regulação da biossíntese de cutina/cera, tolerância e seca (Gray et al., 2012). Mas quando em *O. sativa*, o *AtSHN2* foi super-expresso constitutivamente, resultou no aumento do conteúdo de celulose (35%), diminuição de lignina (45%) e alteração na razão S/G (Ambavaram et al., 2011). Por outro lado, a super-expresssão de *ShSHN1* de cana-de-açúcar em *O. sativa*, provocou um aumento de biomassa(91-140%), conteúdo de celulose (10-22%) e eficiência de sacarificação(5-53%), assim como uma redução do conteúdo de lignina (17-35%) e aumento da razão S/G( 53-106%)(Martins et al., 2018).

No nível superior da rede hierárquica paralelamente a SHN (SHINE/WAX INDUCER (SHN/WIN)), os TFs da família NAC (NAM, ATAF1/2 e CUC2, denominados também SWNs - SECONDARY WALL NACs) são reguladores *master switches*, pois ativam diretamente a expressão de um segundo nível de *master switches* de TFs que correspondem à família de MYBs, sendo que ambos em coordenação ativam o nível inferior da rede de TFs (a maioria deles da família MYB), que constitui uma bateria de TFs que envolvem ativadores e repressores transcricionais específicos da biossíntese da Lignina. Esta hierarquia é responsável pela coordenada expressão de genes relacionados a biossíntese da parede celular secundária (SCW) em fibras e/ou vasos (Cassan-Wang et al., 2013; Zhong et al., 2010a). Em fibras de *Arabidopsis* a regulação é controlada por NST1, NST2 e NST3/SND1. Em vasos xilemáticos (diferenciação de protoxilema e metaxilema e programação de morte celular), o processo é regulado por VND6 e VND7 (Kubo et al., 2005; Wang & Dixon, 2012; Yamaguchi et al., 2008). Os SWNs reconhecem regiões palindrômicas *cis* nos promotores de 19 pb (SNBE) que permitem a regulação via TFs (Zhong et al., 2010b). A super-expresssão de qualquer destes SWNs leva a lignificação ectópica de células que normalmente possuem paredes celulares primárias (Grima-Pettenati et al., 2012). SND1 e NST1 tem uma função redundante na ativação da parede celular em fibras de *Arabidopsis* (Zhong et al., 2006; Mitsuda et al., 2007; Zhong et al., 2007b) e o duplo “knockout” destes TFs ou repressão de SND1 resultam na não formação de SCW em fibras interfasciculares e vascular.

Em gramíneas, foi descrito que em *O. sativa*, OsSND1 é responsável pela deposição de distintos componentes da parede celular (Hirano et al., 2013). Alguns OsSWNs são capazes de complementar *snd1/nst1* em *Arabidopsis* no espessamento da SCW, e quando super-expressos resultam na deposição ectópica de lignina, xilano e celulose (Zhong et al., 2011). Em *B. distachyon*, BdSWNs induz a formação ectópica de parede celular em tabaco; BdSWN5, um ortólogo de AtVND7, pode acelerar o desenvolvimento do xilema e a deposição de ectópicas paredes celulares em distintos órgãos de *B. distachyon*, assim como também a super-expressão de genes relacionados a celulose syntase (*BdCES4*), uma protease específica do xilema (*BdXCP1*) e um ortólogo de *AtMYB46* (*BdMYB1*) (Valdivia et al., 2013). Em switchgrass (*P. Virgatum*), distintos PvSWNs foram caracterizados e são altamente expressos em colmos associados às células do esclerênquima (Zhong et al., 2015). Em estudos de cana-de-açúcar foram utilizados os genes *4CL* e *PAL* como iscas para identificação de 18 TFS co-expressos, sendo *ScNAC83* e *ScNAC36* altamente expressos em *S. spontaneum* e associados à deposição de lignina (Ferreira et al., 2016). Em variedade híbridas de cana-de-acucar contrastantes para o conteúdo de lignina, os genes *ShVNI2* e *ShNST1-2* apresentaram maior expressão no genótipo de maior conteúdo de lignina (IACSP04-529) tanto na medula e córtex de colmo intermediários e maduros (Brito et al., 2015).

O par redundante funcional de TFs: MYB46 e MYB83 tem sido também identificados como importantes reguladores da biossíntese da parede celular e hierarquicamente abaixo dos SWNs na regulação dos fenilpropanóides envolvidos com lignina (Ko et al., 2009; McCarthy et al., 2009). MYB46/MYB83 são alvos de NST3/SND1 e tem função redundante na formação da parede celular. MYB46 e MYB83 ativam genes da biossíntese da celulose, lignina e xilano e causa ectópica formação de SCW. Ambos os genes são alvos dos *master switches* : VND6 e VND7 (Wang & Dixon, 2012). Plantas *myb46/myb83* resultam em fenótipo mais severo que os caracterizados em *nst1/nst3* (Ohashi-Ito et al., 2010; Yamaguchi et al., 2011). Em gramíneas, os genes ortólogos de arroz, milho e *Panicum* : *OsMYB46*, *ZmMYB46* e *PvMYB46A*, respectivamente, são capazes de complementar mutantes de *myb46/myb83* de *Arabidopsis*, revelando uma conservação no controle da síntese da parede celular em eudicotiledôneas e monocotiledôneas para estes máster switches (Zhong, et al., 2011; Zhong et al., 2015)

A identificação de elementos cis-reguladores na maioria dos promotores dos genes da via Biosintética dos monolignóis tem aberto o caminho para a compreensão da regulação transcricional da biossíntese da lignina (Zhao & Dixon, 2011). Ensaio de EMSA (“*electrophoretic mobility shift assay*”) permitiram identificar elementos *cis* associados a TFs da família MYB. Esses elementos *cis* (ricos em AC) estão presentes na maioria dos promotores dos genes da biossíntese da lignina (Zhao & Dixon, 2011). *C4H* e *COMT* são regulados especificamente por AtMYB58 (Zhou et al., 2009). *F5H* em *Arabidopsis* não contém elementos *cis* AC e é regulada indiretamente por MYB58 e MYB63 através de SND1 (Zhao et al., 2010; Zhou et al., 2009). Em *Arabidopsis* os TFs MYB58, MYB63 e MYB85 são ativadores que estão especificamente envolvidos na regulação causando a deposição ectópica de lignina, enquanto MYB4 é repressor de *C4H*, inibindo a biossíntese de compostos derivados de fenilpropanóides (Zhong et al., 2008; Zhou et al., 2009). A caracterização funcional dos correspondentes ortólogos em *Populus* (Zhong et al., 2011b) e *Pinus* (Bomal et al., 2008) sugere que o seu papel na regulação da biossíntese da lignina é conservada entre diferentes espécies. Outros TFs como KNAT7, SND2, SND3, MYB32, MYB52, MYB54, MYB61, MYB69 e MYB103 também em *Arabidopsis*, estão posicionados no último nível da rede, provavelmente envolvidos no ajuste fino da regulação transcricional da deposição da SCW, dentro deste último KNAT e MYB32 é um repressor e os demais são ativadores (Newman et al., 2004; Preston et al., 2004; Zhong & Ye, 2015). Um TF que não pertence a família MYB que promove a biossíntese de lignina é NtLIM, e tem sido descrito que reconhece elementos AC dos promotores de *PAL*, *4CL* e *CAD*. O silenciamento de *NtLIM* promoveu uma redução de 30% do conteúdo de lignina, podendo ser um alvo no controle de conteúdo de lignina em gramíneas (Kawaoka et al., 2000; Kawaoka & Ebinuma, 2001). Em variedades híbridas cana-de-açúcar apresentando distintos conteúdos de lignina foi avaliado a expressão da complexa família de MYBs (Brito et al., 2015).

Os supostos ortólogos destes TFs em cana-de-açúcar, não tem sido identificados ou funcionalmente caracterizados até agora (Cesarino et al., 2016). Em comparação aos máster switches NACs e MYBs, um significativo número de TFs principalmente da família MYB, ubicados no nível inferior da rede hierárquica que ativam ou reprimem especificamente genes da SCW, incluindo aquele da biossíntese da lignina, tem sido caracterizados amplamente em *Arabidopsis* e outras eudicotiledôneas modelo em

ordem de validar a sua função, porém os esforços ainda são limitados ou desconhecidos para várias culturas bioenergia como cana-de-açúcar. Baseados na sequência de homologia de repressores MYB de *A. thaliana*, 5 TFs de milho foram identificados: ZmMYB31, ZmMYB42, ZmMYB2, ZmMYB8 e ZmMYB39 como candidatos para direta repressão de *ZmCOMT* (Fornalé et al., 2006). O ZmMYB31 é o primeiro regulador da lignina caracterizado em monocotiledôneas (Fornalé et al., 2010). Quando *ZmMYB31* foram super-expressos em *A. thaliana* causaram uma diminuição nos níveis de expressão de genes associados à biossíntese da lignina e redução no conteúdo de lignina, sem alterar a composição deste polímero, mas observando um fenótipo severo no crescimento, além disso, em destaque fazendo uso da imunoprecipitação de cromatina (ChIP) se demonstrou que ZmMYB31 interage in vivo, com os promotores de *ZmCOMT* e *ZmF5H* a través do elemento cis ACCT/AACC (Fornalé et al., 2010). A regulatória função de MYB31 e MYB32 como repressores também tem sido descrito em *S. bicolor* e *O. sativa* (Agarwal et al., 2016). *ZmMYB42* causou uma redução nos níveis de lignina S em *A. thaliana* (Sonbol et al., 2009). Recentemente, *ZmMYB31* e *ZmMYB42* foram super-expressos em cana-de-açúcar e ambos reprimem os genes da via biossintética da lignina (Poovaiah et al., 2016). Em *P. virgatum*, a super-expressão de *PvMYB4* (altamente expresso em feixes vasculares e ortólogo de *AtMYB4*), resultou em uma redução do conteúdo de lignina e altera a composição deste polímero (Shen et al., 2012), interagindo com elementos cis AC, presentes nos promotores dos genes da biossíntese de lignina, aquela superexpressão em tabaco e *Panicum*, levou a diminuição de tamanho, redução do conteúdo fenólico, de lignina e baixa abundância de ligações ester de p-coumárico na parede celular, Além disso houve uma forte queda na expressão dos genes da lignina, mas sem afetar a expressão de genes relacionados no braço biosintético dos flavonoides (Shen et al., 2012). Em *O. sativa*, a super-expressão e a sobexpressão de *OsMYB58/63*, *OsMYB42/85*, *OsMYB103* e *OsMYB55/61* provocaram um aumento e queda na expressão de *OsCAD2*, demonstrando que estes TFs podem ter uma função associada na deposição da SCW (Hirano et al., 2013)

Em *Sorghum bicolor* foi identificado *SbMYB60* um co-ortólogo de *AtMYB58* e *AtMYB63*, ativadores específicos da deposição de lignina de *A. thaliana*. A super-expressão de *SbMYB60* aumentou a expressão de vários genes da via biossintética dos monolignóis, acrescentando o conteúdo de lignina, compostos aromáticos, e

alterando os teores de polissacarídeos da parede celular (Scully et al., 2016). Foi demonstrado usando ensaio de transativação de protoplastos de arroz, que: OsMYB58/63 e OsMYB42/85 ativaram fortemente a expressão o gene da cellulose synthase *OsCesA7*, sendo mais evidente esta indução para OsMYB58/63 (Noda et al., 2015). De outro lado, a super-expresssão de *PvMYB58/63A* e *PvMYB42/85A* induziu um aumento da razão S/G, da expressão dos genes de lignina incluindo *F5H* e o conteúdo de lignina acetyl bromide, sendo que Só *PvMYB58/63A* provocou aumento dos genes associados a cellulose synthase e xylan synthase, como também de genes da biossíntese de flavonoides. (Rao et al., 2019) O descrito acima em relação a regulação da SCW por parte de MYB58/63 em *S. bicolor*, *O. sativa* e *P.virgatum* contrasta com a função de AtMYB58/63 (ativador específico da lignificação) desde que este TF não controla a deposição de celulose e hemicelulose na SCW, sugerindo que o clado MYB58/63 em gramíneas podem funcionar com amplos ativadores da biosíntese de lignina, celulose e hemicelulose, em vez de possuir uma regulação específica de lignina atribuída para AtMYB58/63. Foi descrito que a super-expresssão de *ZmMYB167* de milho, ortólogo de *AtMYB85* provocou em *Brachypodium distachyon*, aumento do conteúdo de lignina, dos monômeros S, ácido *p*-coumárico e com alteração no crescimento e redução da sacarificação. Em milho houve também aumento do conteúdo de lignina, e ácido *p*-coumárico, além do ácido ferúlico, mas não houve alteração da composição de lignina, sacarificação, nem do crescimento. Em ambas espécies tampouco se evidenciou alteração da composição de carboidratos (Bhatia et al., 2019).

O TF WRKY12 que não pertence a família MYB e que se localiza no nível superior da rede hierárquica é um repressor da deposição da SCW e o silenciamento do ortólogo AtWKRY2 ocasionou espessamento da parede celular em células parenquimáticas da medula com ectópica deposição de lignina, xilano e celulose em eudicotiledôneas (*Medicago* e *Arabidopsis*) (Wang et al., 2010). Expressão ectópica de *PtWRKY19* de *Populus trichocarpa*, ortólogo de AtWRKY12, no mutante *Atwrky12*, resgatou suficientemente o fenótipo em células parenquimáticas da medula (Yang et al., 2016). De outro lado o silenciamento dos ortólogos de AtWRKY12 em *Medicago sativa*, *Panicum virgatum* e *Zea mays* produziu tanto aumento e/ou deposição ectópica de células lignificadas no parênquima medular, como da biomassa e a sua densidade, a exceção de *P. virgatum*, onde houve redução de biomassa, sugerindo o papel deste TF como repressor da lignificação tanto em monocotiledôneas e eudicotiledôneas

(Gallego-Giraldo et al., 2016). Outro TF que promove a biossíntese de lignina é NtLIM que reconhece elementos AC dos promotores de *PAL*, *4CL* e *CAD*. O silenciamento de *NtLIM* promoveu uma redução de 30% do conteúdo de lignina, podendo ser um alvo no controle de conteúdo de lignina em gramíneas (Kawaoka et al., 2000; Kawaoka & Ebinuma, 2001).

Tem sido demonstrado que a função regulatória da SCW entre eudicotiledôneas e gramíneas é conservada por parte de alguns TFs, Porém os padrões de ligação de um TF em particular ao promotor de um gene alvo e a auto-regulação (ligação ao seu próprio promotor) tem vindo a mostrar a ser específicos da espécie indicando uma potencial sub-funcionalização explicando a divergência funcional da regulação entre sorgo, milho, arroz e *Arabidopsis* tal como o exemplo que foi mencionado acima para o caso de MYB58/63 que pode regular diferencialmente a deposição de dos polímeros segundo a espécie (Agarwal et al. 2016). Esta observação destaca a importância de uma avaliação mais aprofundada da contribuição e do papel específico de distintos fatores de transcrição na complexa rede transcricional do controle da deposição da SCW e por tanto de lignina em gramíneas como cana-de-açúcar. Os estudos em sistemas heterólogos tem mostrado informações importantes, mas são necessárias mais identificações e caracterizações de TFs em gramíneas bioenergia, como a cana-de-açúcar, visando a produção e o desenvolvimento do E2G.

## 2 Justificativa

Este projeto é baseado nos dados obtidos previamente na dissertação de mestrado do aluno Juan Pablo Portilla Llerena. Nela foram estudadas as espécies *S. officinarum*, *S. barberi*, *S. robustum* e *S. spontaneum*, sendo observadas diferenças quanto aos componentes da parede celular, açúcares solúveis, como também na quantidade e tipo de lignina. Os dados são coerentes no sentido de que separam as duas canas que tem sido indicadas como cana energia, *S. spontaneum* e *S. robustum*, que acumulam mais fibra, daquelas que acumulam mais açúcares solúveis na forma de sacarose, *S. officinarum* e *S. barberi*. Além disso, *S. spontaneum* e *S. robustum* acumulam mais lignina insolúvel e possuem uma razão S/G menor do que as outras (Figuras 5 e 6).

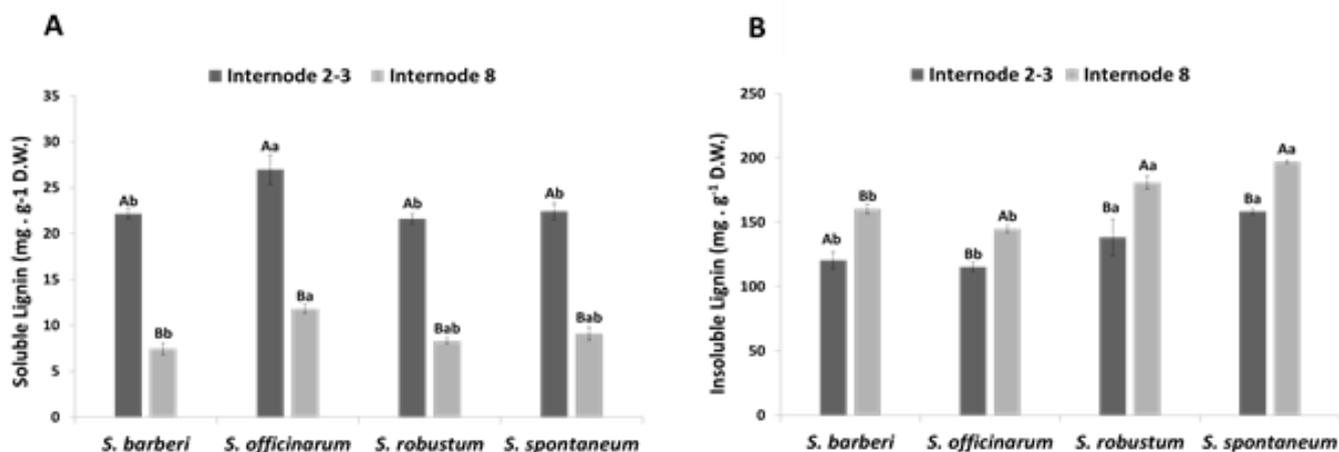


Figura 5 – Conteúdo de (A) lignina solúvel e (B) insolúvel nos internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significantes entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie. Letras minúsculas diferentes indicam diferenças significantes entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes.

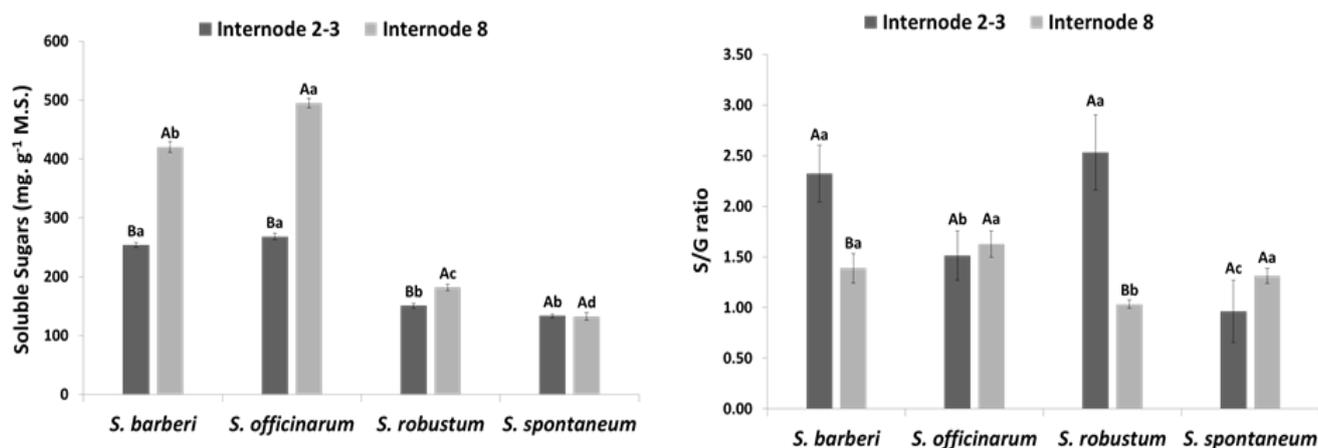


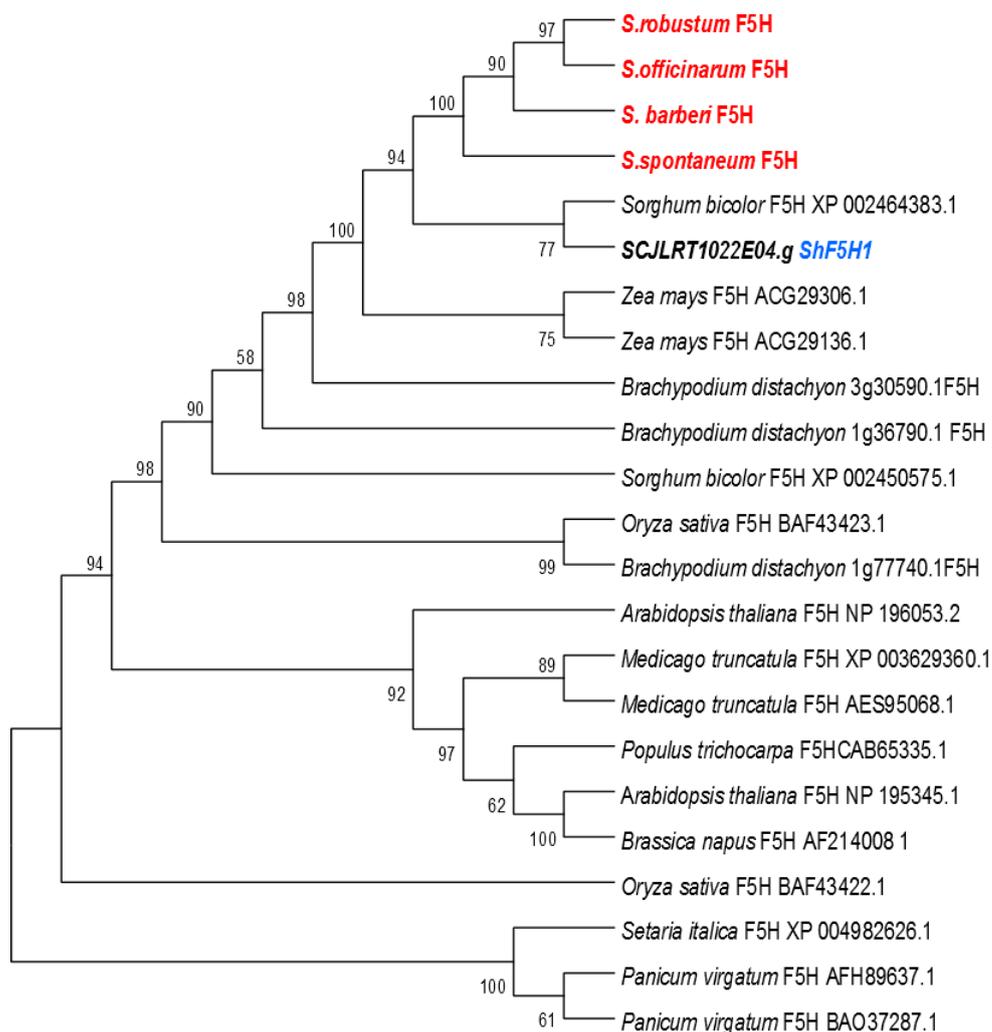
Figura 6 – Conteúdo de açúcares solúveis e razão S/G nos internódios de espécies de *Saccharum*. Letras maiúsculas diferentes denotam diferenças significantes entre internódios de diferente estágio de desenvolvimento dentro de uma mesma espécie. Letras minúsculas diferentes indicam diferenças significantes entre internódios de um mesmo estágio de desenvolvimento das diferentes espécies. As médias foram

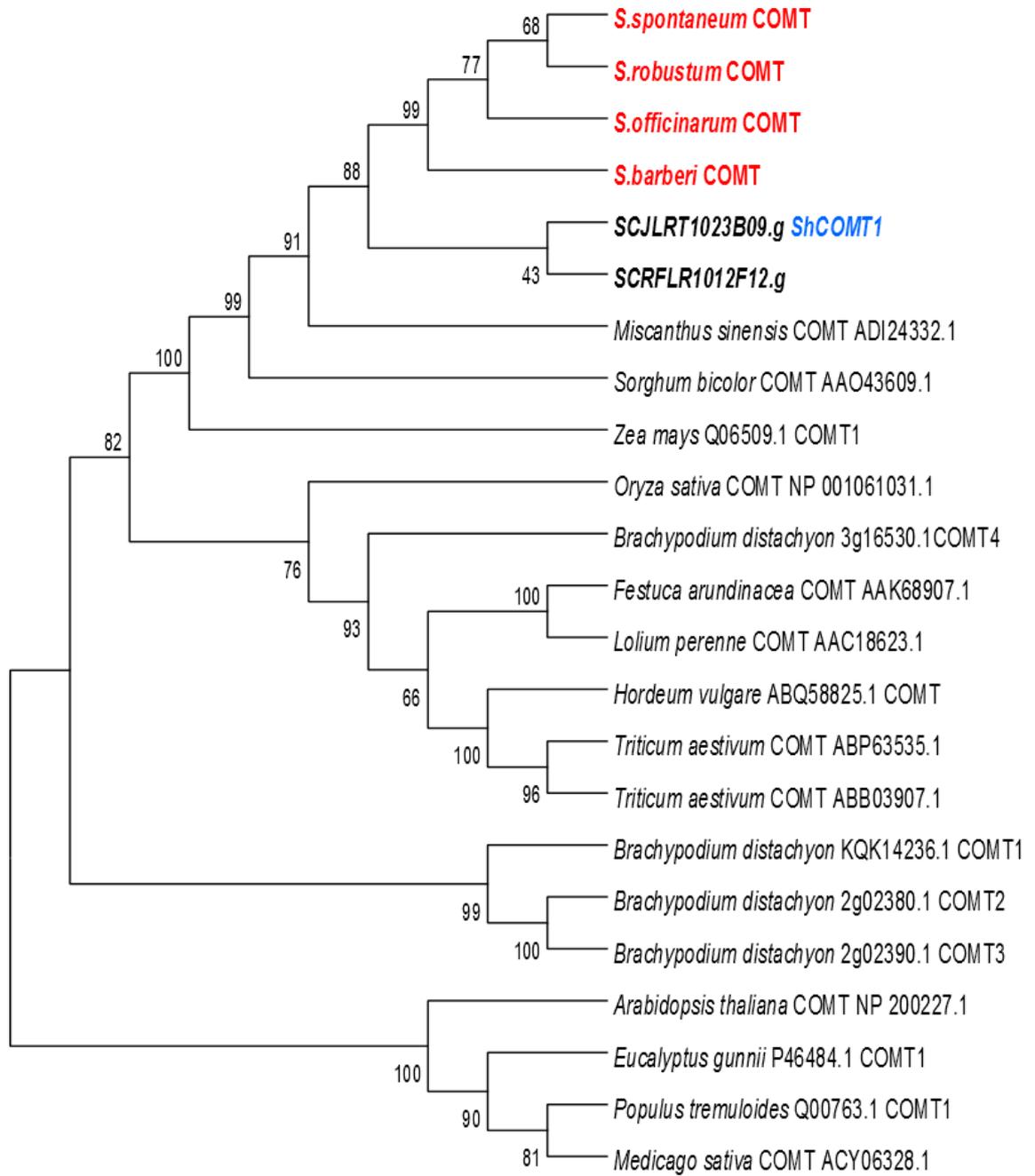
comparadas pelo teste de Tukey,  $P \leq 0,05$ . As barras verticais indicam o erro padrão das médias de quintuplicatas independentes.

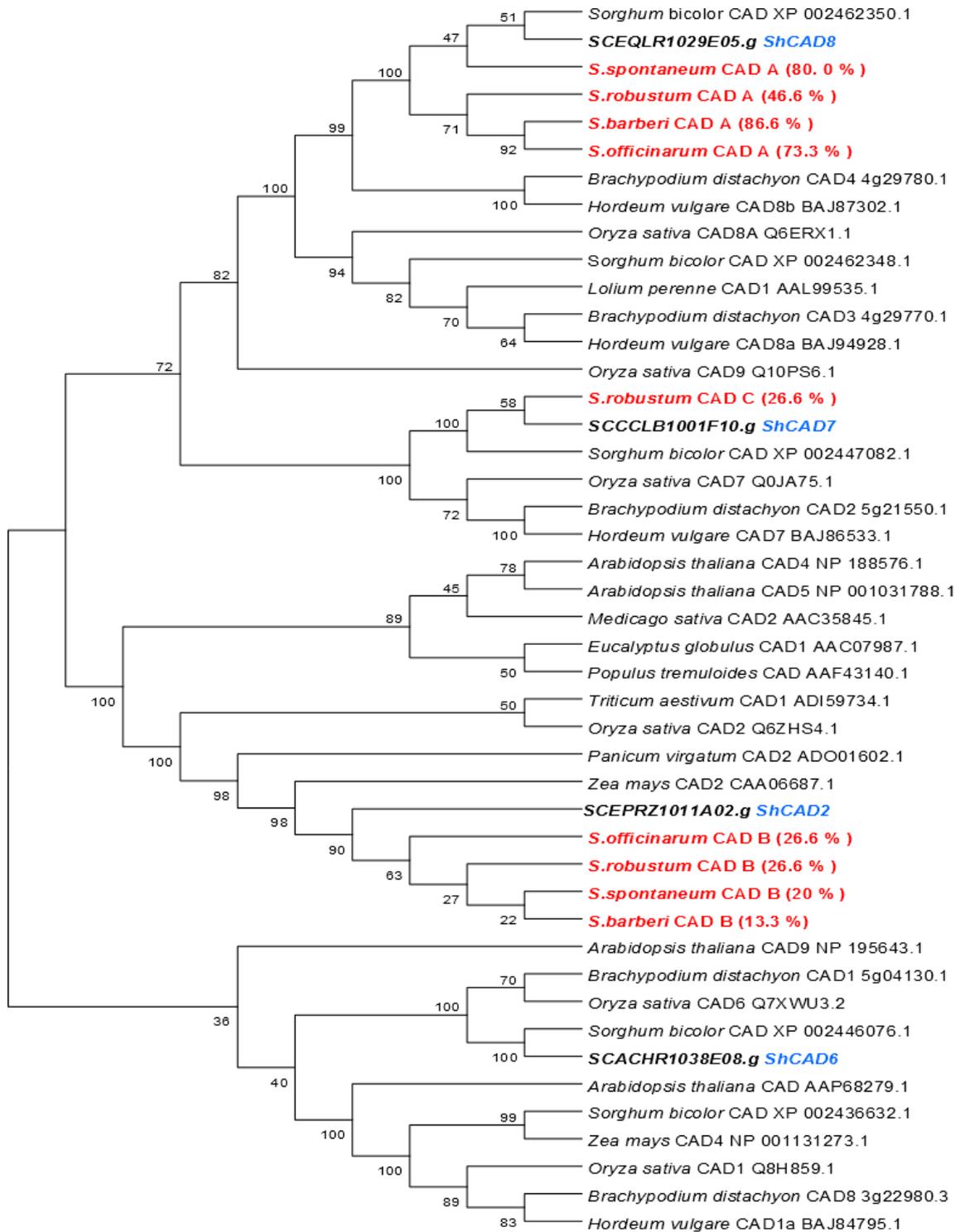
Nessa dissertação de mestrado foram identificados e isolados todos os genes da via de síntese de lignina nas quatro espécies de *Saccharum*. Foram alinhadas as sequências de dois híbridos de cana-de-açúcar identificados em nosso laboratório e de outras gramíneas e usando primers de regiões conservadas foram feitos RT-PCR usando cDNA produzido a partir de RNA extraído da mistura do internódio jovem (2°+3°) e maduro (8°) de plantas de cada espécie, com objetivo de identificar todas as isoformas possíveis. Os primers foram desenhados nas extremidades dos genes e eventualmente outros primers internos foram desenhados. As bandas produzidas foram retiradas de gel, sequenciadas e identificadas por alinhamento com as sequências conhecidas de cana-de-açúcar. No mínimo trinta colônias, contendo o inserto (gene da via de biossíntese de lignina), para cada gene das 4 espécies em estudo, foram sequenciadas no sentido “forward” e “reverse”. Sabendo com que frequência apareceu determinada sequência, foi possível inferir quais dos genes eram mais expressos nas quatro espécies. Inesperadamente os resultados indicaram uma grande conservação em termos de sequências entre as quatro espécies e praticamente os mesmos daqueles identificados por Bottcher et al (2013) em híbridos comerciais. Foram identificados 13 genes sendo que dos três genes de interesse neste projeto, *F5H*, *COMT* e *CAD*, foram identificados um, um e três genes, respectivamente. Entre os três genes *CAD* identificados (*ShCAD2*, *ShCAD7* e *ShCAD8*), dois deles (*ShCAD2* e *ShCAD8*) foram encontrados nas espécies *S. spontaneum*, *S. barberi* e *S. officinarum*, sendo *ShCAD8* o mais expresso, com 80%, 86% e 73% das sequências identificadas no sequenciamento. Em *S. robustum* este gene também foi o mais expresso, com 46,6% das sequências, com 26,6% para os dois outros genes. Na Figura 7 mostramos as árvores filogenéticas destas sequências. Os alinhamentos dessas sequências mostram que elas são praticamente idênticas àquelas identificadas por Bottcher et al. (2013), como homologia acima de 98%. Assim, sendo as mesmas sequências, mas com níveis de expressão diferentes, espera-se o controle da biossíntese de lignina em cana-de-açúcar esteja a cargo de genes controladores, ou seja, de Fatores de Transcrição (TFs) e que de alguma forma estes estariam influenciando nas variações no conteúdo e composição de lignina na parede celular.

Figura 7 – Análises filogenéticas de sequências de aminoácidos das enzimas da via biossintética dos monolignóis, obtidas da base de dados do NCBI, GeneBank, Phytozme e homólogos identificados neste estudo para as espécies de *Saccharum*. As anotações em azul mostram os genes (e respectivas SASs) identificados por Bottcher et al (2013).

### **ShF5H**



**ShCOMT**

**ShCAD8**

### 3 Objetivos

#### Objetivo Geral

Identificar e caracterizar os fatores de transcrição que regulam as enzimas da biossíntese da lignina em plantas de *Saccharum* spp. visando a alteração e modulação da deposição de lignina na parede celular.

#### Objetivos específicos:

1. Isolar a região regulatória dos genes *ShCAD8*, *ShF5H* e *ShCOMT* e identificar *in silico* prováveis elementos de ligação cis em *S. officinarum* e *S. spontaneum*.
2. Identificar fatores de transcrição (TFs) *in silico* que se ligam a elementos cis da região regulatória dos genes *ShCAD8*, *ShF5H* e *ShCOMT* visando uma melhor compreensão da rede transcricional em *S. officinarum* e *S. spontaneum*.
3. Avaliar e correlacionar o Perfil de expressão dos genes da rota biossintética da lignina, TFs (identificados *in silico*), conteúdo e composição de lignina em diferentes tecidos quanto a lignificação visando a escolha de TFs candidatos para silenciamento e caracterização funcional.
4. Verificação da interação entre *ShMYB85* e *ShMYB58/63* e região regulatória dos genes *CAD*, *F5H* e *COMT* através de ensaio de transativação de protoplastos em *S. officinarum* e *S. spontaneum*
5. Caracterizar funcionalmente *ShMYB85* e *ShMYB58/63* selecionados para silenciamento através da transformação do cultivar híbrido modelo de cana-de-açúcar SP-3280.

#### **4 Capítulos da Tese de Doutorado: Identificação e caracterização de fatores de transcrição que regulam a biossíntese de lignina em *Saccharum* spp. cultivar híbrido SP-3280**

Para atender o objetivo geral, o trabalho foi segmentado em dois capítulos: No primeiro capítulo envolve a identificação de fatores de transcrição que regulem a biossíntese e deposição de lignina em *S. spontaneum* e *S. officinarum*. De forma resumida para atingir isto, preliminarmente foi isolado os promotores de genes representativos da rota biossintética da lignina: *ShCAD8*, *ShF5H* e *ShCOMT* em ambas as espécies e as sequências destes foram analisadas *in silico* com o intuito de verificar o repertório de fatores de transcrição (TFs) que se ligam a elementos *cis*. Os parâmetros correspondentes ao perfil de expressão do repertório de TFs ortólogos em cana-de-açúcar encontrados nos promotores e alguns caracterizados na literatura relacionados à deposição da SCW, a expressão dos genes da biossíntese da lignina, o conteúdo e composição deste polímero em distintos tecidos, enquanto a lignina foram submetidos a uma análise de componente principal (PCA), correlação de Pearson e de redes Bayesianas visando encontrar evidências de possíveis relações entre TFs candidatos associados à deposição de lignina. Para verificar isto, foram feitas interações entre os TFs candidatos e os promotores de genes representativos: *ShCAD8*, *ShF5H* e *ShCOMT* nas espécies de *Saccharum* a partir de um ensaio *in vitro* de transativação de protoplastos BY-2 de *Nicotiana tabacum*.

No segundo capítulo refere-se essencialmente à caracterização funcional de linhagens transgênicas silenciadas de *ShMYB85* (*Shmyb85*) e *ShMYB58/63* (*Shmyb58/63*) (que foram escolhidos como candidatos a partir das análises do primeiro capítulo) em cana-de-açúcar cultivar híbrido SP-3280, aqui se aborda a montagem de plasmídeos contendo as construções para silenciamento por RNAi *hairpin* de *ShMYB85* e *ShMYB58/63*, para a transformação em calos embriogênicos de cana-de-açúcar e a obtenção de linhagens transgênicas *Shmyb85* e *Shmyb58/63* de cana-de-açúcar cultivar híbrido SP-3280, silenciadas para *ShMYB85* e *ShMYB58/63*. Adicionalmente este capítulo informa sobre a caracterização funcional das diferentes linhagens transgênicas de *Shmyb85* e *Shmyb58/63* através de diversos testes histoquímicos, determinações de polissacarídeos de parede celular, conteúdo e composição de lignina, sacarificação, perfil e conteúdo de fenóis totais, flavonoides

e antocianinas, assim como também da expressão de genes da rota biossintética da lignina e os TFs: *ShMYB85* e *ShMYB58/63*.

**CAPITULO I. Identificação de fatores de transcrição  
que regulam a biossíntese e deposição de lignina em  
*S. spontaneum* e *S. officinarum***

---

# **CAPÍTULO I. Identificação de fatores de transcrição que regulam a biossíntese e deposição de lignina em *S. spontaneum* e *S. officinarum***

## **1 Materiais e Métodos**

### **1.1 Obtenção do Material vegetal**

Plantas das espécies *S. spontaneum*, *S. officinarum*, cultivadas na Casa de Vegetação do Departamento de Biologia Vegetal do Instituto de Biologia, Universidade Estadual de Campinas, provenientes de clones obtidos por cultura de tecidos fornecidos do Centro de Cana Instituto Agrônomo de Campinas, Ribeirão Preto, São Paulo, Brasil foram utilizadas neste estudo. Os clones foram semeados em bandejas de plástico contendo vermiculita e mantidos em casa de vegetação, com irrigação diária e sem controle de temperatura ou fotoperíodo. As plântulas crescidas foram transplantadas para vasos de 50 L (3 plantas por vaso) contendo substrato orgânico comercial (nitrogênio = 0,5%, umidade = 50%, carbono orgânico = 15%, pH = 6, C/N = 130 mmol/Kg, relação CIC/C = 8%, Genefertil-Genesol) e mantidos sob irrigação por aspersão (30 min, duas vezes por dia) em estufa por um período aproximadamente de um ano. Passado este período, os vasos foram transferidos para fora da estufa. O substrato dos vasos foi trocado e as plantas foram deixadas crescer por um período de mais 4 meses, sob irrigação por aspersão (30 min, três vezes por dia). A coleta foi feita considerando 3 repetições (3 vasos) dos caules das diferentes espécies, aparentemente saudáveis, do mesmo estágio fisiológico e de tamanho uniforme. Para as análises bioquímicas (conteúdo de lignina e razão S/G), expressão relativa dos genes da rota biossintética da lignina (qRT-PCR) e fatores de transcrição (TFs) foram usados 7 órgãos: folha jovem (YL;+2), folha madura (ML;+4), córtex do internódio 3° (R3), córtex do internódio 5° (R5), medula do internódio 3° (P3), medula do internódio 5° (P5) e raiz (R). Uma lâmina de aço foi usada para separar a córtex (Cesarino et al. 2012). Para o isolamento dos promotores dos genes representativos da rota biossintética de lignina: *ShCAD8*, *ShCOMT*, e *ShF5H* foi usada a folha jovem (+2). Após serem lavados com água da torneira o material vegetal foi colocado em saco de polietileno e, em seguida, foram colocados em recipientes térmicos contendo nitrogênio líquido, para posterior liofilização. O material liofilizado foi macerado em mortar previamente esterilizados por autoclave e resfriado com nitrogênio líquido, e

transferidos para tubos Falcon esterilizados, e armazenados em freezer a  $-80^{\circ}\text{C}$  até o seu uso.

## **1.2 Isolamento, clonagem e sequenciamento da região regulatória de Genes da Biossíntese de Lignina em espécies de *Saccharum***

### **Extração de DNA genômico**

Em um microtubo de 2 mL (resfriado em nitrogênio) contendo aproximadamente 1/3 do seu volume com material vegetal (folha +4), foi adicionado 550  $\mu\text{L}$  de tampão de homogeneização e 450  $\mu\text{L}$  de tampão mix (aquecido a  $65^{\circ}\text{C}$ ). O conjunto foi misturado por inversão e colocado em banho maria a  $65^{\circ}\text{C}$  por 1h, agitando a cada 10 min. Se deixou resfriar as amostras a temperatura ambiente. Foi adicionado um volume igual a 1 mL de Fenol (equilibrado a  $\text{pH}=7.9$ ):clorofórmio:álcool isoamílico (25:24:1), agitando por inversão. Os microtubos contendo a mistura homogeneizada, foram centrifugados a 12000 rpm por 5 min a  $4^{\circ}\text{C}$ . O sobrenadante, se transferiu a um novo microtubo de 2 mL e foi acrescentado um volume igual de isopropanol gelado, seguido de 200  $\mu\text{L}$  de NaCl 5M, agitando por inversão cuidadosamente e logo incubando as amostras a  $-20^{\circ}\text{C}$  por 1h e 30 min. As amostras foram centrifugadas a 12000 rpm por 5 min a  $4^{\circ}\text{C}$ . Foi descartado o sobrenadante e o *pellet* foi lavado com 500  $\mu\text{L}$  etanol ao 70% gelado por duas vezes (em cada passo os microtubos foram centrifugados nas mesmas condições acima mencionadas e o etanol foi descartado cuidadosamente). Os microtubos foram deixados na capela por aproximadamente 30 min para evaporar o etanol. Uma vez o *pellet* seco, este foi solubilizado em 50  $\mu\text{L}$  de buffer TE+RNAseA (10 $\mu\text{g}/\mu\text{L}$ ), incubando os microtubos a  $37^{\circ}\text{C}$  por pelo menos 1 hora. A quantificação total de DNA genômico foi feita em um espectrofotômetro Nanodrop 2000C a 260 nm e a pureza foi avaliada pela razão  $A_{260}/A_{280} \sim 2$  e a razão  $A_{260}/A_{230} \sim 1.8$ . A integridade foi verificada em gel de agarose ao 0.8% com brometo de etídio e visualizada sob luz UV em fotodocumentador (Gel Doc 2000, BIORAD). O DNA genômico não apresentou “smear” e foi maior 10kb quando comparando com o marcador molecular (DNA 1 kb *ladder from New England Biolabs*).

## Isolamento dos promotores: *pShCAD8*, *pShCOMT* e *pShF5H*

### Promotor *pShCAD8*:

1. Para isolar o promotor *pShCAD8* nas espécies de *Saccharum* foi realizada uma caminhada genômica consecutiva e sobreposta baseada no protocolo do *GenomeWalker™ Universal Kit (Clontech)*, mas com as seguintes modificações mencionadas abaixo.
2. Para avaliação a pureza de DNA genômico isolado, em um microtubo de 0.6 mL, foi feita uma digestão de 5 µL de DNA (100 µg/ µL) com 1.6 µL de Dral (10units/µL), 2 µL de 10X Buffer Dral, em um total de 20 µL de reação a 37°C overnight. 5 µL da digestão e 0.5 µL do DNA não digerido foram verificadas por eletroforese em gel agarose ao 0.8% com brometo de etídio. Somente o DNA digerido apresentou “smear”.
3. Em cada microtubo de 1.5 mL, foi digerido 25 µL de DNA genômico (100 µg/ µL) com 10 µL de Buffer10X e 8 µL(10units/µL) de: Dral, EcoRV, PvuII e StuI construindo as bibliotecas (DL1, DL2, DL3, e DL4; DL=*DNA library*), em uma reação total de 100 µL a 37°C por 2h, agitando a baixa velocidade entre 5-10 segundos, e logo retornando os microtubos a 37°C *overnight*. 5µL da digestão foram verificadas por eletroforese em gel agarose ao 0.8% com brometo de etídeo para visualizar se a digestão foi completa. As bibliotecas digeridas foram purificadas com o kit *NucleoSpin® Gel and PCR Clean-up Macherey-Nagel*. Uma digestão PvuII de DNA genômico humano foi usado como um controle positivo.
4. Para cada microtubo de 0.6 mL foi transferido 4.8 µL das DL1, DL2, DL3 e DL4, purificadas incluindo o controle (Pvu II) e foi adicionado 1.9 µL do *GenomeWalker Adaptor*, 25 µM (para que os *primers* externos (*AP1* e *AP2*) incluídos no kit consigam anelar durante a PCR com os *primers* específicos de *ShCAD8*), sendo catalisada a ligação de adaptadores com 0.5 µL T4 DNA Ligase (6 units/µL) e 0.8 µL de Buffer 10X a 16°C *overnight* no termociclador Veriti 96-Well Thermal Cycler (AB Applied Biosystems), fazendo volume total de reação de 8 µL. Para parar a reação, os microtubos foram incubados a 70 °C por 5 min e foi adicionado 32 µL de buffer TE (10/1, pH :7,5) fazendo um volume total no microtubo de 40 µL.
5. DL1, DL2, DL3 e DL4 foram utilizadas separadamente como *template* para um primeiro *round* de PCR, tendo como *primer* senso externo o *AP1* (incluído no kit) e

como antisense o *primer* externo do gene *ShCAD8* (GSP1-*ShCAD8*). O produto do primeiro *round* de PCR foi diluído dez vezes e utilizado como *template* para um segundo *round* de PCR (*nested* PCR) tendo como *primer* senso o do kit, AP2, e como antisense o primer interno do *ShCAD8* (GSP2-*ShCAD8*).

6. O primeiro round de PCR envolveu essencialmente três passos: A) 7 ciclos, cada ciclo consistiu em uma desnaturação de 94°/ 25 segundos e anelamento de 72°C /3 min. B) 32 ciclos, cada ciclo uma desnaturação de 94°C /25 segundos e anelamento de 67°C/3 mim e C) uma extensão final a 67°C/7 mim. O segundo round de PCR (*nested* PCR) consistiu também de três passos similares, mas a diferença foi que A, consistiu 5 ciclos e B ,20 ciclos.
7. As concentrações finais dos componentes da reação de PCR para 25 µL foram: 1X *Advantage 2 PCR Buffer* (Clontech), dNTP (10mM each), 1X *Advantage 2 Poymerase* (Clontech), AP1 (0.2 µM), AP2 (0.2 µM), GSP1-*ShCAD8* (0.2 µM), GSP2-*ShCAD8* (0.2 µM), 1µL de DL (mais o controle Pvull) e água deionizada (MilliQ). 5 µL do produto do primeiro foram verificadas por eletroforese em gel agarose ao 1% com brometo de etídeo e visualizadas em um foto-documentador Gel Doc, 2000 BIORAD. So as bandas mais intensas e acima de 1kb do primeiro *round* passaram por um segundo round de PCR, para logo ser analisadas e processadas para posteriores aplicações.

### **Promotores: *pShCOMT* e *pShF5H***

1. Para isolar os promotores: *pShCOMT* e *pShF5H* em ambas espécies, foi usado os bancos de dados: A) CTBE; cultivar *Saccharum* SP-3280(Riaño-Pachón & Mattiello, 2017) B) Phytozome12.0; *Sorghum bicolor* v3.1.1 e adicionalmente para o promotor da *ShF5H*, o C) The Sugarcane genome project , cultivar de *S. spontaneum* GXS87-16 ([https://www.ncbi.nlm.nih.gov/assembly/GCA\\_900500655.1/](https://www.ncbi.nlm.nih.gov/assembly/GCA_900500655.1/)). Como isca usamos o primeiro exon, sendo a busca feita usando o algoritmo blastn (Altschul et al., 1990) e se limitou só a contigs que retornaram com *e-value*=0 e identidade no menor a 98%. Para o *pShF5H* foram selecionados os contigs: SCSP803280\_000004297(A), Sb\_ *F5H*\_Chr01:17681872...17686866 (B) e o ctg8755 (C). Para o *pShCOMT*, os contigs foram: SCSP803280\_000024077 (A) e Sb\_ *COMT*\_Chr07:4724504...4726892 (B).

2. Os contigs foram alinhados no Clustal Omega (Sievers et al., 2011) e se desenhou primers em regiões conservadas *upstream* e dentro da região do *primer* exon dos genes usando o Primer 3 (Rozen & Skaletsky, 2000).
3. Se desenvolveu PCR touchdown para isolar os promotores de ambas espécies. A PCR touchdown para o caso do pShF5H envolveu 2 passos: A) 10 ciclos, cada ciclo com desnaturação a 98°C/20s, anelamento de 30s desde 73°C até 63°C e extensão a 72°C/90s. B) 17 ciclos, com desnaturação 98°C/20s, anelamento de 63°C/30s e extensão a 72 °C/90s. Houve uma desnaturação inicial de 98°C/2 min e uma extensão final, 72°C/7 min.
4. As concentrações finais para 25 µL de reação padrão de PCR foram: 200 µM de cada dNTP, 1,5 mM de MgCl<sub>2</sub>, 0.005U de iProof High Fidelity DNA Polimerase, Buffer 1X iProof HF, 0.2 µM de cada primer, 0.36 µL(1.08 ng) DNA, formamida 2% e água livre de nucleases.
5. 3 µL produto da PCR amplificado, correspondente ao tamanho esperado foram separadas em gel de agarose 1 % contendo brometo de etídio e visualizadas em um foto-documentador Gel Doc, 2000 BIORAD, o restante do produto de PCR foi purificado com o kit *PureLink Quick Gel Extraction and PCR Purification Combo Kit*.
6. A PCR touchdown para o caso de promotor da *ShCOMT* envolveu 2 passos: A) 10 ciclos, cada ciclo: desnaturação a 98°C/10s, anelamento de 30s desde 60°C até 55°C e extensão a 72°C/2:20 min. B) 25 ciclos, cada ciclo: desnaturação 98°C/10s, anelamento de 55°C e extensão a 72 °C/2:20 min. Houve uma desnaturação inicial de 98°C/2 min e uma extensão final 72°C/10 min.
7. As concentrações finais para 25 µL de reação padrão de PCR foram: 200 µM de cada dNTP, 1,5 mM de MgCl<sub>2</sub>, 0.005U de *iProof High Fidelity DNA Polimerase*, *Buffer 1X iProof HF*, 0.5 µM de cada primer, 1.5µl DNA (75 ng), formamida 2% e água livre de nucleases.
8. 3 µL produto da PCR amplificado, correspondente ao tamanho esperado foram separadas em gel de agarose 1 % contendo brometo de etídeo e visualizadas em um foto-documentador Gel Doc, 2000 BIORAD, o restante do produto de PCR foi purificado com o kit *PureLink Quick Gel Extraction and PCR Purification Combo*

*Kit*. Os produtos de PCR purificados serviram como template para um segundo *round* de PCR com as mesmas condições acima descritas.

### **Clonagem e sequenciamento dos promotores *pShCAD8*, *pShCOMT* e *pShF5H***

O promotor do gene *ShCAD8* e aqueles do *ShF5H* e *ShCOMT* foram clonados no *pGEM-T Easy Promega* e *pCR 2.1-TOPO invitrogen*, respectivamente. Os plasmídeos recombinantes foram introduzidos em *E. coli* cepa DH5 $\alpha$  por transformação termocompetente (*Novagen*), as bactérias foram plaqueadas em meio LB semi-sólido contendo 100  $\mu\text{g mL}^{-1}$  de ampicilina, 0,5 mM de IPTG e 80  $\mu\text{g mL}^{-1}$  de XGal e incubadas a 37°C *overnight*. Cerca de 25 colônias transformadas de coloração branca foram selecionadas para a mini-preparação de plasmídeos pelo método de lise alcalina. A presença do insertos *pShCAD8* e *pShCOMT*, nos plasmídeos, foi confirmada por uma reação de restrição com EcoRI (Promega), que corta o vetor em ambas as extremidades do inserto, liberando o fragmento inserido, visualizado em gel de agarose. Para o *pShF5H*, a presença do inserto no vetor, foi verificado por PCR com primers M13, desde que este promotor contém sítios de restrição endógenos para EcoRI. Os insertos foram verificados em gel de agarose ao 1% com brometo de etídio e visualizada sob luz UV em fotodocumentador (Gel Doc 2000, BIORAD). Para confirmar a identidade dos insertos presentes nos plasmídeos em ambas as espécies de *Saccharum*, estas foram enviados para sequenciamento Sanger no Laboratório Central de Tecnologias de Alto Desempenho em Ciências da Vida (LaCTAD), UNICAMP, em sequenciador ABI 3730XL DNA Analyzer, usando *primers* flanqueadores do vetor (M13).

### **1.3 Análise *in silico* dos promotores dos genes da biossíntese da Lignina: *ShCAD8*, *ShCOMT* e *ShF5H***

A sequências obtidas a partir do sequenciamento Sanger foram manualmente processadas e montadas usando o programa BioEdit (Hall, 1999), e alinhadas usando Clustal Omega (Sievers et al. 2011) sendo submetidas preliminarmente ao banco de dados do CTBE (Riaño-Pachón & Mattiello, 2017) usando o algoritmo blastn (Altschul et al., 1990) para verificar a existência da região do primeiro exon onde foi desenhado os *primers* reversos, com o intuito de ter uma ideia de funcionalidade da região regulatória isolada. O tamanho do primeiro exon nos distintos genes em estudo nas espécies de *Saccharum*, foram inferidos a partir da

conservação do tamanho desta região entre distintas monocotiledôneas usando a base de dados de Phytozome v12.1 (<https://phytozome.jgi.doe.gov>). Para descartar a existência de possíveis ORF encontradas na região regulatória isolada foi usada a ferramenta ExPASy Translate Tool (<https://web.expasy.org>). Elementos *cis* putativos de forma geral foram identificados usando os Programa PlantCARE (Lescot et al., 2002) e em particular para os SNBE, SMRE e AC foi usado o programa ExactSearch (Gunasekara et al., 2016), os Programas de predição de sítios de ligação de fatores de transcrição sobre sequencias regulatórias: plant transcription factor database PlantTFDB (Jin et al., 2017), PlantRegMap (Tian et al., 2020) e PLANT PAN 3.0 (Chow et al., 2019) foram usados para identificar *in silico* potenciais fatores de transcrição que se liguem a elementos *cis* das regiões regulatórias dos genes envolvidos a biossíntese da Lignina: *ShCAD8*, *ShCOMT* e *ShF5H*.

#### **1.4 Identificação e análise de expressão de sequencias ortólogas de genes da biossíntese da lignina e de fatores de transcrição (TFs) em espécies de *Saccharum***

##### **Extração de RNA total e síntese de cDNA**

A extração de RNA total em raízes de *S. spontaneum* e *S. officinarum* foi feita de acordo com o método descrito por Chang et al. (1993). A extração de RNA total de 6 tecidos (Folha Nova(+2), Folha Velha(+4), córtex do internódio 3°, córtex do internódio 5°, medula do internódio 3° e medula do internódio 5°)) nas espécies de *Saccharum* foi feita de acordo com o método descrito por Porto et al. (2011) com as seguintes modificações : Em um microtubo de 2 mL foi adicionado aproximadamente 1/3 do seu volume com material vegetal macerado e adicionou-se 1,5 mL do Reagente de Guanidina-Tiocianato y Fenol, PVPP e 15 ul de  $\beta$ -mercaptoetanol . O microtubo foi agitado por 30 s e incubado à temperatura ambiente por 30 min. Em seguida o material foi centrifugado a 12.000 xg por 10 min à 4°C, sendo a partir de então mantido em gelo. O sobrenadante foi recuperado, a ele adicionado 0,5 mL de clorofórmio, seguido de agitação em vórtex por 15 s e centrifugação à 12.000 xg a 4°C, por 15 min. Recuperou-se 1 mL do sobrenadante e a ele foram adicionados 0,25 mL de isopropanol e 0,25 mL de 10 M LiCl (solução com DEPC 0.1%). O extrato foi mantido “overnight” à -20°C para precipitação do RNA. Após esse período o conteúdo foi novamente centrifugado à 12.000 xg a 4°C por 15 min, e o *pellet*

formado foi lavado duas vezes com 1 mL de etanol 70% (solução com DEPC 0.1%), com centrifugação a 12.000 xg por 5 min entre as lavagens. O pellet foi seco à temperatura ambiente e, posteriormente, solubilizado em água deionizada (MilliQ) autoclavada, definindo-se para as amostras 40 µl. A quantificação do RNA total foi feita em espectrofotômetro Nanodrop Thermo Scientific 2000C a 260 nm e a pureza verificada pela razão  $A_{260}/A_{280}$ , que deve ser próxima de 2,0 e a razão  $A_{260}/A_{230}$  não menor que 1.8. A integridade do RNA foi verificada por eletroforese em gel agarose 1.5% com brometo de etídio e visualizada sob luz UV em fotodocumentador (Gel Doc 2000, BIORAD). As amostras de RNA total foram tratadas usando o kit *DNAse I Amplification Grade* (Sigma) e submetidas à transcrição reversa usando o kit *RevertedAid First Strand cDNA Synthesis* conforme instruções do fabricante (*Thermo Scientific*), um total de 3µg de RNA foi utilizada para a síntese de cDNA primeira fita.

### **Identificação de sequencias ortólogas de genes da biossíntese da lignina em espécies de *Saccharum***

As sequencias dos 10 unigenes identificados por Llerena et al. (2019) em *S. spontaneum* e *S. officinarum*: 1 *C4H* (*cinnamate 4-hydroxylase*), 1 *4CL* (*4-hydroxycinnamoyl CoA: ligase*), 1 *HCT* (*hydroxycinnamoyl-CoA: shikimate/quinic acid p-hydroxycinnamoyl-transferase*), 1 *F5H* (*ferulate 5-hydroxylase*), 1 *C3H* (*p-coumaroylshikimate 3'-hydroxylase*), 2 *CCoAOMT* (*caffeoyl CoA 3-O-methyltransferase*), 1 *CCR* (*cinnamoyl CoA reductase*), 1 *COMT* (*caffeate O-methyltransferase*) e 2 *CAD* (*cinnamyl alcohol dehydrogenase*), foram usadas para avaliar o seu perfil de expressão em distintos tecidos. Como dois genes foram isolados para *CCoAOMT* e *CAD* eles foram identificados como A e B. *CCoAOMT A*, *B* e *CAD A,B* são homólogos de *ShCCoAOMT 2, 1* e *ShCAD8, 2* respectivamente (Bottcher et al.2013).

### **Identificação sequencias ortólogas de fatores de transcrição (TFs) e análise filogenética em cana-de-açúcar.**

O repertório encontrado de TFs que se ligam *in silico* aos promotores isolados dos genes *ShCAD8*, *ShCOMT* e *ShF5H* e outros TFs descritos na literatura regulando a deposição de lignina foram considerados para a busca das sequencias dos TFs ortólogos em cana-de-açúcar. Assim sequencias de TFs de plantas modelo de

gramíneas (*Z. mays*, *S. bicolor*, *P. virgatum* e *B. distachyon*) e *Arabidopsis* serviram de “baits” usando as ferramentas blastn e tblastx (Altschul et al., 1997), limitando-se a busca de CDS ou EST com  $e\text{-value cut-off} < e^{-100}$ , nos bancos disponíveis do SUCEST (<http://sucest-fun.org/>), CTBE (Riaño-Pachón & Mattiello, 2017), Sugarcane Genome Hub (Garsmeur et al., 2018), Phytozome (<https://phytozome.jgi.doe.gov>), NCBI (<https://www.ncbi.nlm.nih.gov>) e RNAseq derivado de dois genótipos de cana-de-açúcar contrastantes para lignina (Vicentini et al., 2015). As sequências obtidas dos TFs máster switches e ativadores ou repressores do nível inferior da rede hierárquica dos bancos de dados diversos foram usadas para análises filogenéticas, com o intuito identificar homólogos dos TFs envolvidos na regulação da biossíntese de lignina na cana-de-açúcar. As árvores filogenéticas foram construídas com sequências de aminoácidos de cana-de-açúcar, as gramíneas mencionadas acima e as eudicotiledôneas *A. thaliana* e *P. trichocarpa*. O alinhamento múltiplo de sequências de aminoácidos foi realizado com o programa ClustalW (Thompson et al., 2002). As análises e as árvores filogenéticas foram realizadas e visualizadas, respectivamente através do programa MEGA versão 5.02 (Tamura et al., 2011) e as relações evolutivas foram inferidas usando o algoritmo Neighbor-joining com Bootstrap para 1000 repetições. “Gap regions” foram excluídas por ajuste manual.

### **Perfil de expressão de genes da biossíntese da lignina e de fatores de transcrição (TFs) em espécies de *Saccharum***

De forma geral, as sequências dos genes da rota biossintética da lignina isoladas em espécies de *Saccharum* por Llerena et al. 2019, assim como as sequências dos fatores de transcrição ortólogos identificados aqui para cana-de-açúcar, foram utilizadas para desenho de *primers* específicos, para avaliar a expressão gênica por qRT-PCR quantitativo (qPCR). Especificamente para as sequências ortólogas dos TFs, estas foram alinhadas com os ortólogos de *S. bicolor* fazendo uso do programa Clustal Omega (Sievers et al., 2011) e a partir do alinhamento, foram desenhados *primers* para RT-qPCR em regiões não conservadas ((fora do domínio MYB, NAC, WRKY, AP2 e KNOTTED HOMEODOMAIN (KNOX)), perto do final da CDS, 3'UTR ou 5'UTR, desde que o domínio proteico dos genes desta famílias de TFs são muito conservadas. Para o desenho do *primers* tanto para o genes da lignina como para os TFs foi usado o programa Primer 3 (Rozen & Skaletsky, 2000),

considerando a  $T_m \sim 60^\circ\text{C}$ , a % GC ( 60), tamanho dos primers foi entre 18-23 pb e o tamanho de amplicon fluando entre de 90 a 180 pb. A possível presença de estruturas internas secundárias e potencial formação de dímeros nos *primers* foi verificado pelo *software* OligoAnalyzer 3.1(Owczarzy et al., 2008). A especificidade do primer para cada um dos *templates* foi avaliada com a ferramenta Primer-Blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast>). As sequências dos *primers*, e o tamanho esperado do amplicon são fornecidos nas Tabelas S3A e S3B. A curva de eficiência dos primers foi determinada pelo software Step One Plus Software v2.3 Life Technologies. As amostras de cDNA diluídas 12.5 X para 7 tecidos (Folha Nova, Folha Velha, córtex do internódio 3°, córtex do internódio 5°, medula do internódio 3°, medula do internódio 5° e raiz) das espécies *S. officinarum* e *S. spontaneum* foram usadas nesta análise. As reações de qRT-PCR (processadas em 3 repetições biológicas e 3 replicatas técnicas) foram preparadas com o *iTaq™ universal SYBR® Green supermix BIORAD* e analisadas em um StepOnePlus™ Real-Time PCR System, usando o método: 95°C por 3 min seguidos por 40 ciclos de 95°C por 10 s e 60°C por 30 s. A especificidade dos produtos amplificados foi avaliada pela análise da curva de dissociação gerada pelo equipamento. Para a normalização interna adequada, foi utilizado o gene constitutivo *gliceraldeído 3-fosfato desidrogenase (GAPDH)* e *UBIQUITIN*, que foram identificados como os mais estável por Bottcher et al. (2013). A expressão relativa foi apresentada utilizando-se a fórmula:  $2^{-\Delta Ct}$ , onde  $\Delta Ct = (Ct_{tag} - Ct_{ref})$ , sendo o Ct definido como o número de ciclos necessários para a taxa de amplificação do gene alvo se tornar exponencial, tag = gene alvo, e ref = gene referência. Este método é derivado do método  $2^{-\Delta\Delta Ct}$  (Livak & Schmittgen, 2001).

### **1.5 Determinação de Lignina Acetyl Bromide (AcBr)**

A determinação do conteúdo de Lignina Acetyl Bromide foi realizada de acordo com a metodologia descrita por (Hatfield & Fukushima, 2005) em duas etapas com as seguintes modificações: Na primeira etapa em um microtubo de 2ml (pesado preliminarmente) foi colocado 20 mg de biomassa vegetal finamente macerado se procedeu a isolar a parede celular, lavando esta sequencialmente com 1 ml de: água deionizada (MilliQ), etanol, clorofórmio e acetona, incubando a 98°C, 76°C, 59°C e 54°C, respectivamente, em cada passo de incubação, os microtubos foram agitados a 750 rpm por exatamente 30 min fazendo uso de um Thermomixer C

Eppendorf e centrifugando por 5 min a 14000 rpm eliminando cuidadosamente o sobrenadante e mantendo o pellet. O pellet foi seco em vácuo usando um Concentrator Plus Eppendorf durante 45 min. O microtubo contendo a parede celular isolada foi pesado. A diferença de peso do microtubo com a amostra lavada menos o microtubo é o peso de parede celular isolada. Em uma segunda etapa, ao microtubo contendo a parede celular isolada foi adicionado 250  $\mu$ L de uma solução fresca de acetyl bromide (25% AcBr em Ácido Acético) cuidadosamente para evitar “splashing” da amostra, os microtubos foram incubados a 50°C (reação) sem agitar e passado este tempo foi deixado mais uma hora em vórtex por 15 min, posteriormente os microtubos foram colocados em gelo e centrifugados por 15 min a 14000 rpm. Em um novo microtubo de 2ml foi adicionado: 100  $\mu$ L do sobrenadante da reação, 400  $\mu$ L de NaOH 2M e 75  $\mu$ L de hidroxilamine hidrocloreto fresco, e a mistura agitada. Para os brancos foi considerado 100  $\mu$ L de 25 de AcBr em ácido acético. Foi Adicionado finalmente 1.425 ml de ácido acético e os microtubos foram agitados por inversão. Foram feitas leituras a 280 nm fazendo uso de espectrofotômetro Thermo Scientific™ NanoDrop 2000. Aqui usamos o fator de correção de 23.35. O conteúdo de Lignina Acetyl bromide foi expresso em % de parede celular.

### **1.6 Determinação da razão S/G**

A determinação da razão S/G foi feito de acordo com Kiyota *et al.* (2012). Amostras de 100 mg foram hidrolisadas com uma solução de 2,0 ml de NaOH 4M em tubos de ensaio com tampa de rosca e vedados com fita teflon. Os tubos foram aquecidos durante 24 h a 95°C em um sistema de aquecimento do tipo dry-block. Os tubos foram resfriados em temperatura ambiente e as amostras foram acidificadas com aproximadamente 1,6 ml de HCl 6M. O tubo foi agitado para garantir a neutralização da base. Após a acidificação, o pH foi conferido com fita de papel indicador. Em seguida a amostra foi centrifugada a 13000 rpm por 5 min e uma alíquota de 500  $\mu$ L do sobrenadante foi transferida para outro recipiente (microtubo de 2,0 mL) e extraída duas vezes com 1,0 mL de acetato de etila, juntando as duas frações da fase orgânica. O solvente foi evaporado sob fluxo de N<sub>2</sub> e o conteúdo foi dissolvido em 1,0 mL de água deionizada (MilliQ). Em seguida as amostras foram analisadas por UHPLC-MS. Se utilizou um sistema de UHPLC acoplado com um espectrômetro de massas do tipo triplo quadruplo equipado com fonte de ionização ESI (modelo

ACQUITY, Waters Corp, Manchester, UK). O software MassLynx V.4.1 (Waters Corp. Manchester, UK) foi utilizado para aquisição e tratamento dos dados. Os compostos provenientes da hidrólise alcalina foram separados usando uma coluna C8 (50 mm x 2,1 mm; 1,7  $\mu$ m). Como fase móvel se usará água (A) e acetonitrila (B), seguindo o gradiente: 90% A até 4,5 minutos, mudando para 100% B em 8,0 minutos. Essa condição foi mantida constante até 9,0 minutos, e em seguida retornou-se para condição inicial estabilizando em 10,0 minutos. O fluxo se manteve constante durante a corrida em 0,300 mL/min. O volume de injeção foi de 5,0  $\mu$ L. Os espectros de massas foram adquiridos em modo SIM para os íons m/z 121, 151 e 181, em modo negativo sendo a voltagem do capilar -3,0 KV e do cone -52, -26 e -30 V para H, G e S, respectivamente. A temperatura da fonte e de dessolvatação foi de 150 e 350 °C, respectivamente.

### **1.7 Correlação e Análise de Rede Bayesiana**

Os dados de expressão (RT-qPCR) e os dados metabólicos foram usados em uma análise de correlação de Pearson para identificar uma possível relação entre a expressão dos TFs identificados aqui e a expressão de genes da biossíntese de lignina (Llerena et al.2019), razão S/G e conteúdo de lignina em espécies de *Saccharum*. A análise foi realizada no R software (Team, R 2013). A análise da rede bayesiana foi realizada usando o software BNFinder (Wilczyński & Dojer, 2009). Um total de 23 pontos de dados foram coletados da rede e correspondem a TFs, genes da biossíntese de lignina, de amostras que pertencem a 7 tecidos ((Folha Nova (+2), Folha Velha (+4), córtex do internódio 3°, córtex do internódio 5°, medula do internódio 3°, medula do internódio 5° e raiz), enquanto 10 pontos de dados pertencem aos genótipos e tecidos e dados bioquímicos. Nas redes obtidas, foi assumido que a expressão de genes pode depender de condições (características genéticas, morfológicas e bioquímicas), mas não o contrário. Além disso, a orientação das interações regulatórias foi inferida. Foi considerado os dados quantitativos de PCR como variáveis contínuas e outros dados (genótipo, tipo de tecido, razão S/G, conteúdo de lignina) como variáveis discretas. O critério de pontuação de equivalência Bayesiana-Dirichlet foi usado para compreender a estrutura da rede composta por essas variáveis discretas e contínuas. Neste estudo, as interações representaram o fato de que uma variável depende de alguma outra variável, e o tipo de interação pode ser uma correlação positiva ou negativa entre

variáveis. A topologia da rede foi devidamente reconstruída usando o Cytoscape (Smoot et al., 2011).

## **1.8 Isolamento e ensaio de transativação de protoplastos em espécies de *Saccharum***

### **Isolamento de Fatores de Transcrição: *ShMYB85* e *ShMYB58/63***

1. Sequências já caracterizadas de *S. bicolor*, *Zea Mays*, *B. distachyon*, *P. virgatum* e *Arabidopsis* serviram como iscas usando as ferramentas blastn e tblastx (Altschul & Madden, 1997), para isolar os ortólogos de *ShMYB85* e *ShMYB58/63* nas espécies de *Saccharum*, usando os banco de dados do SUCEST (<http://sucest-fun.org/>), CTBE (Riaño-Pachón & Mattiello, 2017), Sugarcane Genome Hub (Garsmeur et al., 2018), Phytozome (<https://phytozome.jgi.doe.gov>), NCBI (<https://www.ncbi.nlm.nih.gov>) e RNAseq derivado de dois genótipos de cana-de-açúcar contrastantes para lignina (Vicentini et al., 2015)
2. As sequências codificantes completas (CDS) dos TFs ortólogos em cana-de-açúcar: *ShMYB85* (Sh\_250M06\_t000170), *ShMYB58/63* (Sh\_227K15\_t000030) foram obtidos do Sugarcane Genome Hub (Garsmeur et al., 2018) e serviram para desenhar primers com sítios *attB* parcial e total (para clonagem no vector Invitrogen pDONR221 Gateway) com o intuito de amplificar toda a CDS completa destes TFs. *ShMYB85*, *ShMYB58/63* foram amplificadas a partir de um *pool* de cDNA correspondente aos tecidos da medula e córtex do internódio 5 de *S. spontaneum*
3. Foi feito dois *rounds* de PCR. O primeiro, consistiu de 35 ciclos: cada ciclo: desnaturação a 98°C/20s, anelamento de 55°C/30s e extensão a 72°C/1 min. Houve uma desnaturação inicial de 98°C/2 min e uma extensão final 72°C/5 min. O segundo envolveu 2 passos: A) 5 ciclos, cada ciclo: desnaturação a 98°C/15s, anelamento de 45°C/30s e extensão a 72°C/1 min. B) 20 ciclos, cada ciclo: desnaturação a 98°C/15s, anelamento de 57°C/30s e extensão a 72°C/1 min. Houve uma desnaturação inicial de 98°C/2 min e uma extensão final 72°C/10 min.
4. As concentrações finais para 25 µL de reação padrão de PCR tanto do primeiro e segundo *round* foram: 200 µM de cada dNTP, 0.005U de *iProof High Fidelity DNA Polimerase*, *Buffer 1X iProof HF* (1,5 mM de MgCl<sub>2</sub>), 0.2 µM de cada primer

(*attB* parcial e total), 1µl cDNA (do 5to internódio de *S. spontaneum*) , formamida 2% e água livre de nucleases.

5. 3 µL produto do segundo round PCR amplificado, correspondente ao tamanho esperado foram separadas em gel de agarose 1.5 % contendo brometo de etídeo e visualizadas em um foto-documentador Gel-Doc 2000 BIORAD, o restante do produto de PCR foi purificado com o kit *PureLink Quick Gel Extraction and PCR Purification Combo Kit*.

### **Clonagem dos Promotores: (*pShCAD8*, *pShF5H* e *pShCOMT*) e Fatores de Transcrição (*ShMYB85* e *ShMYB58/63*)**

1. Como mencionado anteriormente o promotor do gene *ShCAD8* e aqueles do *ShF5H* e *ShCOMT* de ambas as espécies foram clonados no *pGEM-T Easy Promega* e *pCR 2.1-TOPO invitrogen*, respectivamente, e a presença do inserto no vector e identidade destes foi verificado por PCR com primers M13 e sequenciamento Sanger, respetivamente. Com exceção do promotor da *ShF5H* em *S. spontaneum*, os promotores incluídos nestes vectores foram amplificados usando primers específicos com sítios *attB*, nas mesmas condições como se descreve acima para os TFs em estudo.
2. Todos os promotores e TFs (com exceção do *pShF5H* de *S. spontaneum*) foram clonados no vetor *Invitrogen pDONR221 Gateway*, usando *Invitrogen Gateway BP Clonase II Enzyme mix*. O *pShF5H* de *S. spontaneum* foi clonado no vetor *invitrogen pENTR/D-TOPO* usando os mesmos *primers* para *S. officinarum* sem sítios *attB* sendo que o primer reverso foi desenhado na sentido 5´ com os nucleotídeos CCAA e o primer reverse sem adicionar nenhum nucleotídeo. A presença dos insertos nos vetores foi verificada por PCR com *primers* universais M13.
3. Os vetores de entrada positivos para os Promotores e TFs foram recombinados nos vetores de destino: *pGWL7* e *p2GW7*, para gerar o vetor repórter dirigindo a expressão da *Renilla* luciferase (rLUC) e o vetor efetor, em que um constitutivo promotor CaMV 35S impulsiona a expressão *ShMYB85*, *ShMYB58/63* respectivamente usando *Invitrogen LR Clonase II Plus Enzyme mix*. A presença do inserto nos vetores de entrada e a identidade foram verificado por PCR e sequenciamento Sanger usando *primers attB*, respectivamente.

## Ensaio de Transativação de Protoplastos

1. O isolamento e transformação de protoplastos BY2 de *N. tabacum* foi feito de acordo com a metodologia descrita por Lin et al. (2018).
2. O ensaio de transativação foi desenvolvido com 6 replicatas (cada uma contendo 100 µL de solução protoplasto (500 protoplastos por µL) para cada combinação do promotor e TFs, usando o Kit *Dual Luciferase Promega* seguindo as instruções do fabricante
3. Foi necessário um total de 4µg de cada vetor efetor (contendo os TFs), vector repórter (contendo os promotores) e o vector de normalização contendo o promotor CaMV 35S dirigindo a expressão da *Renilla* luciferase (rLUC).
4. A transfecção dos protoplastos foi feita usando uma solução de Ca<sup>2+</sup> / PEG, e as células foram incubadas por 24 h no escuro, antes da lise. A transativação dos promotores, *pShCAD8*, *pShF5H* e *pShCOMT* por ShMYB85 e ShMYB58/63, respectivamente, foi estimado, calculando a razão das atividades luciferase firefly/renilla (fLUC/rLUC)
5. Como controle negativo, um vector p2GW7 que contém a sequência codificante do gene reporter β-GLUCURONIDASA (GUS) (em lugar dos TFs) foi co-transformada com o vector pGWL7 que contém o promotor candidato.

### 1.9 Análises estatísticas

Para as análises bioquímicas (conteúdo de lignina e razão S/G) e análises de expressão (RT-qPCR) foi realizada uma análise de variância (ANOVA) simples, e quando houve diferenças entre as médias a comparação foi feita pelo teste de Tuckey ( $p < 0,05$ ). Para o ensaio de transativação de protoplastos BY-2 em *N. tabacum* as diferenças significativas foram determinadas pelo t-test ( $p < 0,05$ ). Para as análises bioquímicas foram feitas 3 replicatas biológicas e no que respeita a de expressão (RT-qPCR) dos TFs ortólogos e genes da biossíntese de lignina em espécies de *Saccharum* foram feitas 3 replicatas biológicas com três replicatas técnicas. Os resultados das análises bioquímicas e do ensaio de transativação foram expressas pela média ± erro padrão. Para expressão génica os resultados foram expressados pela média ± desvio padrão. Para o controle da transferência do erro no cálculo de expressão génica foi utilizado um modelo linear de acumulo de erro ( $\sigma_{\Delta Ct}^2 = \sigma_{Ct,ref}^2 + \sigma_{Ct}^2$ ) no cálculo do valor de  $\Delta Ct$  e não linear ( $\sigma_{2^{-\Delta Ct}}^2 = \left(\frac{d[2^{-\Delta Ct}]}{d[\Delta Ct]}\right)^2 \sigma_{\Delta Ct}^2$ ) no cálculo do valor

de  $2^{-\Delta Ct}$  (Brown & Mac Berthouex, 2002). Foi utilizada uma análise de componente principal (PCA) usando o software R (Team, R 2013) para simplificar as tendências e verificar possíveis agrupamentos entre a expressão de TFs, genes de biossíntese de lignina e parâmetros metabólicos (razão S/G e conteúdo de lignina) para cada uma das espécies de *Saccharum*, e revelar as variáveis mais importantes para cada agrupamento em apenas dois componentes (dimensões) que resumem todas as variáveis avaliadas (Lever et al., 2017) usando o software R (Team, R 2013), sendo aplicado o método de análise de fatores para dados mistos (FAMD) (Pagès, 2014).

## 2 Resultados e Discussão

### 2.1 Isolamento e análise *in silico* de elementos *cis* da região regulatória dos genes *ShCAD8* e *ShCOMT* e *ShF5H* em espécies de *Saccharum*

Inicialmente, no que respeita a *pShCAD8*, foi amplificado com sucesso um fragmento de DNA genômico de 1880 pb (DL1) para *S. officinarum* e 2212 pb (DL3) para *S. spontaneum*, dos quais 1619 pb e 1950 pb corresponderam só a região regulatória (promotor + região 5'UTR), e 261 pb ao primeiro exon respectivamente. É importante deixar claro que para *S. spontaneum* foi isolado adicionalmente uma sequência regulatória de 1150 pb (DL1) que possui 100 % de similaridade na homologia de sequência com DL3, porém como DL3 possui maior extensão de pb foi considerada para análises posteriores. Para *pShCOMT*, tanto para *S. officinarum* como para *S. spontaneum* foi obtido um fragmento de DNA de 1428 pb correspondendo à parte regulatória 1319 pb para ambas as espécies e 109 pb ao primeiro exon. Para *pShF5H* tanto para *S. officinarum* como para *S. spontaneum* foi obtido um fragmento de 1774 sendo que 1701 pb corresponde só à região regulatória e 73 pb ao primer exon (Figura 8). A análise do fragmento amplificado para *pShCAD8* das espécies de *Saccharum* que inclui 261 pb do primeiro exon e o restante para a região 5'UTR e promotora, revela uma sequência que compreende 620 pb de alta similaridade (96%) entre as espécies em direção *upstream* do codão inicial ATG (Figura S1). Em destaque para o caso dos promotores *pShCOMT* e *pShF5H*, a totalidade da região regulatória isolada possui ~ 95% de similaridade entre as espécies em estudo ao contrário no encontrado para *pShCAD8* (61%) (Figura S1). As regiões regulatórias destes genes foram analisadas na ocorrência de elementos *cis* acima do códon inicial ATG, identificando distintos elementos *cis* nas sequências de *pShCAD8*, *pShCOMT* e *pShF5H* nas

espécies de *Saccharum* (Tabela S1). Um distinto número e frequência dos elementos *cis* característicos: 5UTR Py-rich stretch (elemento *cis* conferindo elevados níveis de transcrição) e TATA box foi encontrado (Figura S1). De outro lado analisando o sentido *forward* e *reverse* da região regulatória para *pShCAD8*, podemos observar a presença de 5 tipos de sequencias SNBE (Secondary wall NAC binding elements) para *S. officinarum* e 3 para *S. spontaneum*. No que respeita para as sequencias SMRE (secondary wall MYB-responsive element) podemos assinalar que *S. officinarum* apresenta 1 tipo enquanto foram 3 para *S. spontaneum* (Tabela S1A). Para o caso de *pShCOMT* (Tabela S1B), a mesma quantidade e tipos de SMREs foram encontrados entre a especies. A região regulatória de *pShF5H*, possui 1 SNBE e 2 SMREs para *S. officinarum* e para *S. spontaneum*, 2 SNBEs e 2 SMREs (Tabela S1C). Assim, elementos *cis* AC II e III em promotores que regulam genes que participam na biossíntese da lignina em tecidos xilemáticos foram encontrados nos *pShCAD8*, *pShCOMT* e *pShF5H* indistintamente nas espécies de *Saccharum*. Não foi encontrado elementos *cis* AC do tipo I na região regulatória de *pShCAD8*, *pShCOMT* e *pShF5H* e em destaque se observa que para *pShCOMT* não encontramos sequencias SNBE. A presença *in silico* dos motivos presentes nas regiões regulatórias aqui, foi baseado nas SNBE e SMRE encontradas em *A. thaliana*, usando como isca, no algoritmo desenvolvido por Gunasekara et al. (2016) , as sequências consenso palindrômica imperfeita caracterizadas funcionalmente de 19 pb SNBE: (T/A)NN(C/T)(T/C/G)TNNNNNNA(A/C)GN(A/C/T)(A/T) (Zhong et al., 2010b) e a sequência de 7 pb SMRE: ACC(A/T)A(A/C)(T/C) (Zhong & Ye, 2012). A ausência destes motivos aqui, pode ser devido a que nas regiões regulatórias isoladas existam uma ou mais bases degeneradas, indicando isto uma expansão destas sequencias consenso nas espécies de *Saccharum* ou uma falta de base de dados robusta e atualizada que reconheça esses possíveis SNBE em *pShCOMT*. Para corroborar isto optamos por verificar as sequencias regulatórias de *pShCOMT* em outros algoritmos de predição também baseados em *A. thaliana* (Jin et al., 2016; Tian et al., 2020; Chow et al., 2019) encontrando SNBE no *pShCOMT* que se ligam aos NACs: NST1/SND1 e VNI2 (TTGCGTCAGGGTCAAGT e TGACGCAAACCTAACGCAATA, respectivamente). Porém é importante assinalar que no promotor da *caffeoyl shikimate esterase* (CSE), foi encontrado *in silico* um total de 7 tipos de SNBE: 5 em sentido *forward* e 2 em sentido *reverse*, mas quando foi verificado esta possível interação através de um ensaio de transativação nenhum dos NACs avaliados

impulsou a ativação do promotor deste gene (Vargas et al., 2016). Por tanto a presença desses elementos *cis* de ligação *in silico* a TFs não implica necessariamente que eles são elementos funcionais (Zhong et al., 2010b), mas são um possível indicativo de uma possível interação, mais ainda em gramíneas, onde não existe uma identificação e caracterização funcional desses motivos *cis*. Consideramos que é urgente cada vez mais esforços na identificação de elementos *cis* em regiões regulatórias de genes da biossíntese de lignina e sua atualização em base de dados em *A. thaliana* (melhor organismo modelo caracterizado) e principalmente em gramíneas com fins de bioenergia. De outro lado é bem conhecido que a expressão dos genes da biossíntese da lignina é afetada por muitos processos biológicos do metabolismo intermediário e por fatores ambientais (Moura et al., 2010). Aqui para *pShCAD8*, *pShCOMT* e *pShF5H*, outros elementos *cis* foram encontrados na região regulatória relacionados a resposta hormonal por ácido abscísico (ABRE), processos anaeróbicos (ARE), processos de variação circadiana e responsivos à luz (AT-1 motif; BOX-4; G-BOX; GATA motif; MRE e GT1) indução por estresse hídrico (MBS), baixa temperatura (LTR), metabolismo da giberelina (P-box), metil-jasmonato na defesa contra patógenos (CGTCA-motif; TCA elements e TGACG-motif) e auxinas (TGA box; TGA elements), estes elementos estão descritos nas Tabelas 1-3.

Tabela 1 - Principais elementos *cis* encontrados na região regulatória de *ShCAD8* nas espécies de *Saccharum* segundo PLANT CARE

<b>Função</b>	<b>Elemento Cis</b>	<b>Posição em <i>S. spontaneum</i></b>	<b>Posição em <i>S. officinarum</i></b>
Respostas ao ácido abscísico	ABRE	+664	- 826
Desenvolvimento da parede celular	MYB	+871; -1285	-615;-752;-756
	MYB recognition site	+479	+920;-1079
	MYB-like	—	-756
Respostas induzidas pela anaeróbia	ARE	-718; +1740	-436;-689
Respostas à Luz	AT1-motif	-1120	—
	Box 4	-1117;-1199	+195
	G- box	-663; +952	+826
	GATA- motif	-1851	-1519
	Pc-CMA2c	-1307	—
	GT1	—	+429
Respostas envolvidas com Metil-Jasmonato	CGTCA-motif	+221; +590; +500;	+1034
	TGACG-motif	+1363; -427; -662; -512	—
Resposta a baixa temperatura	LTR	-248;+730;-346	—
Resposta a seca	MBS	-227;+1408;-409	—
Resposta envolvida a giberelina	P-box	-577	—
Resposta envolvida com ácido salicílico	TCA-elements	-680	-968
Respostas envolvidas a auxina	TGA-box	-1360	-1301
	TGA element	—	-956

Tabela 2 - Principais elementos *cis* encontrados na região regulatória de *ShCOMT* nas espécies de *Saccharum* segundo PLANT CARE

<b>Função</b>	<b>Elemento Cis</b>	<b>Posição em <i>S. spontaneum</i></b>	<b>Posição em <i>S. officinarum</i></b>
Respostas ao ácido abscísico	ABRE	+446; +646	-598;-399
Respostas à Luz	Box-4	+354;-934	+305;-905
	G-box	+646	+598;+370
	GT1-motif	-983	-974
	I-box	+249	+203
	MRE	+919	+890
Respostas envolvida com Metil-Jasmonato	CGTCA-motif	+432;+1355;-912	+385;+1317;-883
Respostas envolvida a giberelina	P-box	—	+270

Tabela 3 - Principais elementos *cis* encontrados na região regulatória de *ShF5H* nas espécies de *Saccharum* segundo PLANT CARE

Função	Elemento Cis	Posição em <i>S. spontaneum</i>	Posição em <i>S. officinarum</i>
Respostas ao ácido abscísico	ABRE	-735	-722
Desenvolvimento da parede celular	MYB	-709;-916;-658;-1079;	-696;-1064;-901
	MYB recognition site	+1079;+709	-696;-1064
Respostas induzidas pela anaeróbia	MYB-like	-916	-901
	ARE	-----	+1176
Respostas à Luz	AT1-motif	-----	-----
	Box 4	+65;-1096;+298;- 1113	+307;+1081;+287;+1098
	G- box	+735	+722
	GATA-motif	-426	-415
	Pc-CMA2c	-----	-----
Respostas envolvidas com Metil-Jasmonato	GT1-motif	+1064	+1049
	CGTCA-motif	+318	+307
	TGACG-motif	-318	-307
Resposta a baja temperatura	LTR	-----	-----
Resposta a seca	MBS	+706; +1141	+693;+1126
Resposta envolvida a giberelina	P-box	-----	-----
Resposta envolvida com ácido salicílico	TCA-elements	-1601	-1611
Respostas envolvidas a auxina	TGA-box	-----	+1605
	TGA element	+1595	-----

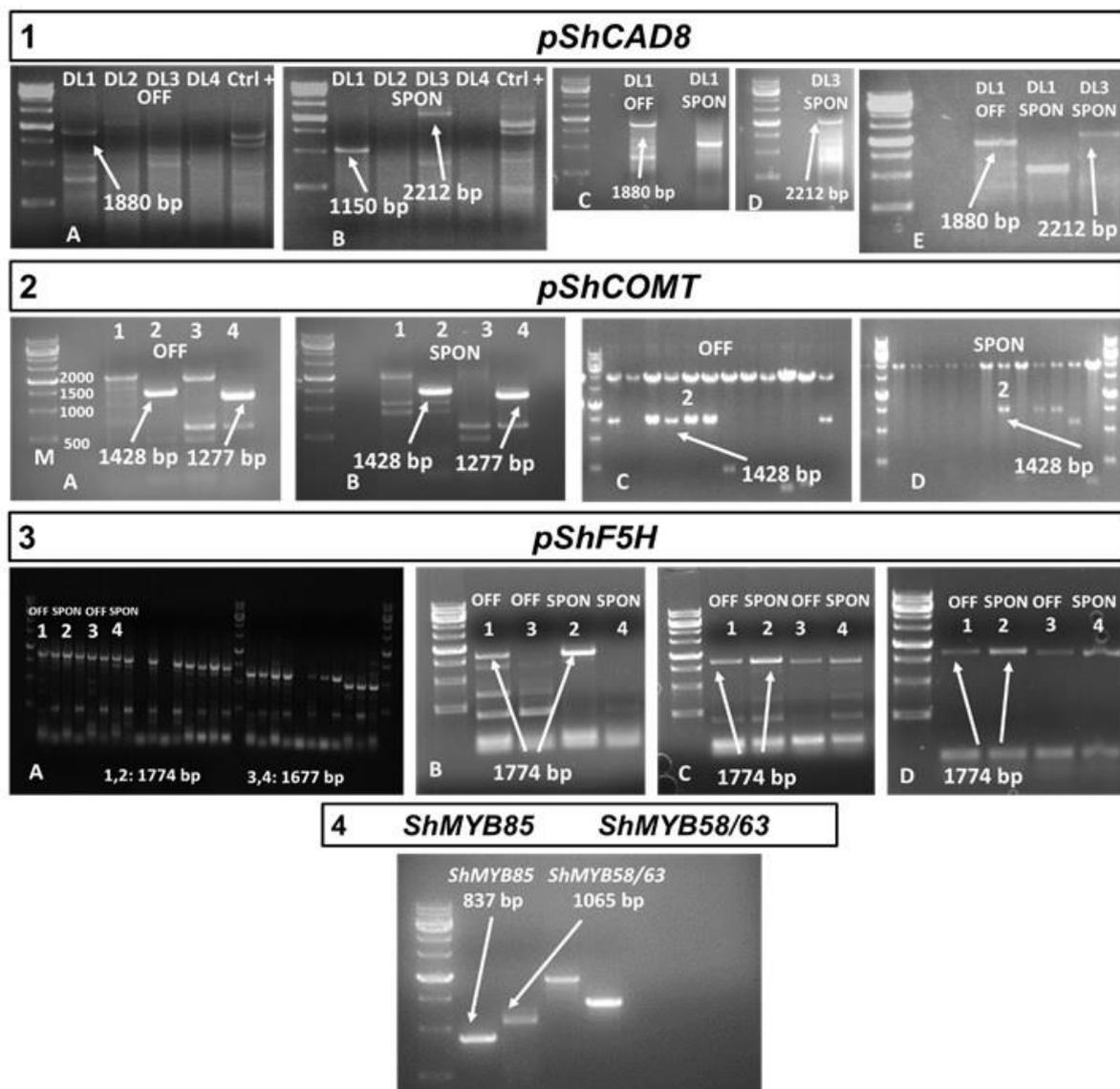


Figura 8 – Otimização do isolamento de promotores e fatores de transcrição em espécies de *Saccharum*. 1: *pShCAD8*: (1A-B) 1<sup>o</sup> round PCR das bibliotecas DL1, DL2, DL3, DL4 e Ctrl+(PvuII). (1C-D) 2<sup>o</sup> round PCR das bibliotecas DL1(1880 bp) em *S. officinarum* (OFF) e DL1(1150 bp) e DL3(2212 bp) em *S. spontaneum* (SPON). (1E) Amplificação de DL1 (OFF), DL1 e DL3 (SPON) a partir do 2<sup>o</sup> round de PCR isolado de gel de agarose ao 1 %. 2: *pShCOMT*: 2(A-B) 2<sup>o</sup> Round de PCR usando 4 combinações de primers (1-4) a partir de DNA genômico de OFF e SPON, 2 (1428 bp) e 4 (1277 bp) apresentaram as bandas mais intensas. 2(C-D) Avaliação de minipreps por digestão com EcoRI para verificar a presença do produto de PCR da combinação 2 em OFF e SPON. 3: *pShF5H*: (3A) Bandas de distintos tamanhos obtidas a partir de distintas combinações de primers usando DNA gnômico. 1-3 e 2-4 representam as duas bandas de maior tamanho em OFF e SPON, respectivamente. (3B-D) Série de otimizações para a obtenção de bandas específicas de maior tamanho (1774 bp) mudando parâmetros de PCR como descritos no texto em OFF e SPON. 4 Amplificação de *ShMYB85* e *ShMYB58/63* usando como *template* o cDNA do 5<sup>to</sup> internódio de SPON. As bandas visualizadas para *pShCAD8*, *pShCOMT*, *pShF5H*, *ShMYB85* e *ShMYB58/63* foram separadas em gel de agarose ao 1% contendo brometo de etídio. M= Marcador Molecular; 1kb Ladder, Sinapse Inc - M1181.

## **2.2 Identificação de TFs *in silico*, na região regulatória dos genes *ShCAD8* e *ShCOMT* e *ShF5H* em espécies de *Saccharum*, busca de sequencias ortológicas de TFs e análise filogenética em cana-de-açúcar *Saccharum spp.***

O repertório encontrado de TFs que se ligam *in silico* aos promotores isolados dos genes *ShCAD8*, *ShCOMT* e *ShF5H* correspondem a *master switches* e ativadores ou repressores posicionados no nível inferior da rede hierárquica transcricional que regula a SCW é apresentado na (Tabela 4-6). Outros TFs descritos na literatura regulando especificamente a deposição de lignina que não foram encontrados a interagir *in silico* com elementos *cis* nas regiões regulatórias dos genes em estudo, também foram considerados na busca das sequencias dos TFs ortólogos em cana-de-açúcar. Desde que os algoritmos usados aqui para a identificação dos TF que se ligam *in silico* nas regiões regulatórias estão baseados em *A. thaliana* e *Zea mays* (Agarwal et al., 2016; Fornalé et al., 2010; Gray et al., 2012; Zhao & Dixon, 2011; Sonbol et al., 2009) usamos as sequencias de TFs caracterizadas funcionalmente destas espécies como iscas para a sua busca em banco de dados, principalmente no SUCEST (<http://sucest-fun.org>), Sugarcane Lignin RNA-seq Database (Vicentini et al., 2015), CTBE (Riaño-Pachón & Mattiello, 2017), Sugarcane Genome Hub (Garsmeur et al., 2018), Phytozome (<https://phytozome.jgi.doe.gov>), NCBI (<https://www.ncbi.nlm.nih.gov>). 10 TFs em cana-de-açúcar foram identificadas como homólogos das sequencias de *Arabidopsis* e 2 para *Z. mays*. No que respeita a *Arabidopsis*, as sequencias homólogas de cana-de-açúcar foram confirmados adicionalmente por uma busca reversa no banco de dados *Arabidopsis* Information Resource (TAIR) usando a ferramenta BLASTX (Altschul et al. 1990). Indistintamente da região regulatória do gene e da espécie, foram identificados sítios de ligação para AtVND6 e AtVND7, porém não foi possível montar sequencias homólogas o suficientemente extensas, a partir das distintas bases de dados em cana-de-açúcar existentes para a sua inclusão nas análises filogenéticas e de expressão génica. É importante mencionar que para obter as CDS completas de alguns dos TFs homólogos de cana-de-açúcar aqui identificados, foi necessário montar mais de uma ESTs a partir de um banco de dados ou de diferentes. As CDS completas de todos os TFs identificados em cana-de-açúcar, os códigos das EST(clones) dos bancos de dados e do seus ortólogos usados na análise filogenética aqui, são apresentados na

figura S2 e tabela S2, respectivamente. Procurando pelos domínios específicos nas sequências dos homólogos identificados em cana-de-açúcar, os 12 TFs foram classificados em 4 famílias: MYB (MYBR2R3), NAC, AP2, WRKY e HOMEBOX (KNOTTED). Cada sequência dos TFs identificados em cana-de-açúcar das mencionadas famílias foram caracterizadas por análise filogenética mediante alinhamento com sequências aminoacídicas de TFs das melhores monocotiledôneas (*O. sativa*, *S. bicolor*, *B. distachyon*, *P. virgatum* e *Z. mays*) e eudicotiledôneas modelo (*A. thaliana* e *P. trichocarpa*) caracterizadas e relacionadas com a biossíntese de lignina e deposição da SCW. Os TFs de cana-de-açúcar identificados e os homólogos usados para a análise filogenéticas foram nomeados usando *Sh* (*Saccharum*) e as primeiras letras do gênero e espécie, respectivamente, mais o nome de seu respectivo homólogo caracterizado em *Arabidopsis* ou milho (ver Figura S2).

Como se observa na Figura 9, de forma geral, podemos assinalar, que as distintas famílias (clados) correspondentes aos TFs homólogos identificados para cana-de-açúcar se agruparam de forma similar como descrito em *Arabidopsis* (Zhong et al., 2010a; Zhong & Ye, 2009; Zhou et al., 2009). As sequências dos TFs de cana-de-açúcar para cada uma das famílias, se agregaram de forma próxima com ortólogos de monocotiledôneas em particular com aqueles *S. bicolor* e *Z. mays*. Apesar desta agrupação, os ortólogos de *A. thaliana* encaixaram no mesmo clado para cada família em particular, indicando uma possível conservação da função regulatória na deposição da parede celular entre sequências de monocotiledôneas e eudicotiledôneas (Handakumbura & Hazen, 2012).

A rede regulatória transcricional que compreende principalmente aos máster switches NACs WRKY, AP2/ERF e MYBs, e uma bateria de TFs da família MYBs e KNOX, posicionados no nível inferior da rede hierárquica ativando ou reprimindo genes da SCW, incluindo aqueles da biossíntese da lignina, tem sido caracterizados amplamente em *Arabidopsis* e outras eudicotiledôneas modelo em ordem de validar a sua função, porém os esforços ainda são limitados ou desconhecidos para várias culturas bioenergia como cana-de-açúcar.

Tabela 4 - Fatores de Transcrição que interagem com elementos *cis* presentes na regulatória do gene *ShCAD8* nas espécies de *Saccharum* (predição segundo TFDB e PLANT PAN 3.0)

<b>Fator de Transcrição (Ortólogo em Arabidopsis ou Milho)</b>	<b>Locus</b>	<b>Posição em <i>S. spontaneum</i></b>	<b>Posição em <i>S. officinarum</i></b>	<b>Função</b>
AtMYB46	AT5G12870	+557	-754	Ativador
AtMYB83	AT3G08500	+564	+752	Ativador
AtMYB52	AT1G17950	+533	-38	Ativador
AtMYB55	AT4G01680	-556	+982	Ativador
AtMYB61	AT1G09540	—	+990	Ativador
AtMYB63	AT1G79180	-562	—	Ativador
AtNST-1	AT2G46770	-1360	-1031	Ativador
AtWRKY-12	AT2G44745	-660	+1250	Represor
AtVND6	AT5G62380	-1360	—	Ativador
AtVNI2	AT5G13180	+222	—	Represor
AtKNAT7	AT1G62990	+257	—	Represor
ZmMYB31	GRMZM2G050305	-1204	—	Represor

Tabela 5 - Fatores de Transcrição que interagem com elementos *cis* presentes na regulatória do gene *ShCOMT* nas espécies de *Saccharum* (predição segundo TFDB e PLANT PAN 3.0)

<b>Fator de Transcrição (Ortólogo em Arabidopsis ou Milho)</b>	<b>Locus</b>	<b>Posição em <i>S. spontaneum</i> e <i>S. officinarum</i></b>	<b>Função</b>
AtMYB46	AT5G12870	+1189	Ativador
AtMYB61	AT1G09540	+854	Ativador
AtMYB83	AT3G08500	+1188	Ativador
AtMYB63	AT1G79180	+938	Ativador
AtMYB55	AT4G01680	-765	Ativador
AtMYB58	AT1G16490	+1158	Ativador
AtMYB46	AT5G12870	+1176	Ativador
AtNST-1	AT2G46770	-904	Ativador
AtWRKY12	AT2G44745	-462	Represor
AtVND6	AT5G62380	-911	Ativador
AtVNI2	AT5G13180	-924	Represor
ZmMYB31	GRMZM2G050305	-1156	Represor
ZmMYB31	GRMZM2G050305	-1204	Represor

Tabela 6 - Fatores de Transcrição que interagem com elementos *cis* presentes na regulatória do gene *ShF5H* nas espécies de *Saccharum* (predição segundo TFDB e PLANT PAN 3.0)

<b>Fator de Transcrição (Ortólogo em Arabidopsis ou Milho)</b>	<b>Locus</b>	<b>Posição em <i>S. spontaneum</i></b>	<b>Posição em <i>S. officinarum</i></b>	<b>Função</b>
AtMYB46	AT5G12870	-1326	-1393	Ativador
AtMYB83	AT3G08500	-----	-1629	Ativador
AtMYB52	AT1G17950	-1008	-1050	Ativador
AtMYB55	AT4G01680	-----	+72	Ativador
AtMYB61	AT1G09540	+1487	+1554	Ativador
AtMYB58	AT1G16490	-1485	-1552	Ativador
AtMYB63	AT1G79180	-1480;+5	+1547	Ativador
AtNST-1	AT2G46770	-----	+1490	Ativador
AtWRKY-12	AT2G44745	+473;+728	+519	Represor
AtVND6	AT5G62380	-----	-----	Ativador
AtVNI2	AT5G13180	-----	-----	Represor
AtKNAT7	AT1G62990	-1135	-1177	Represor
ZmMYB31	GRMZM2G050305	-1267	-1306	Represor



monocotiledôneas envolvidas com a regulação da deposição SCW, correspondentes a AtNST1/2/SND1 (AtNST1, AT2G46770.1) (Mitsuda et al., 2005), PtNST1 (PtrWND2A; Potri.014G104800.1) (Zhong et al., 2011), PvNST1 (PvSWN2B; Pavir.J21162) (Zhong et al., 2015), BdNST1 (BdSW8; Bradi3g13727) (Valdivia et al., 2013) e OsNST1 (OsSWN2; Os08g02300) (Zhong et al., 2011), indicando com isto que ShNST1 é um homólogo com potencial função de ativador *master switch* da deposição da SCW. De outro lado nós identificamos sequências não caracterizadas ortólogas de ShNST1 em *Z. mays* (ZmNST1; GRMZM2G092465) e *S. bicolor* (SbNST1; Sobic.007G018100) mais próximas a ShNST1 do que aquelas sequências identificadas de ZmSWN1 e SbNAM (Zhong et al., 2011; Wang et al., 2013).

Outro TF da família NAC, identificado em cana-de-açúcar foi o ShVNI2 (Sh\_212D06\_t000050) que se agrupou no mesmo clado com o AtVNI2. AtVNI2 interage com alta afinidade com VNDs, atuando como repressor transcricional da diferenciação em vasos xilemáticos, sendo *Arabidopsis* é o único organismo modelo onde até momento este TF foi identificado e caracterizado (Kubo et al., 2005; Yamaguchi et al., 2010). Para reforçar o hipotético papel de ShVNI2 como potencial repressor transcricional incluímos na análise filogenética sequências ortólogas não caracterizadas de eudicotiledôneas e monocotiledôneas modelo que retornaram usando como isca, o ShVNI2. PtVNI2 (Potri.001G061200.4), ZmVNI2 (GRMZM2G083347\_T01), BdVNI2 (Bradi4g07527.1), PvVNI2 (Pavir.Cb00769.1), OsVNI2 (Os12g29330.1) e SbVNI2 (Sobic.008G094700.1) se agruparam proximamente a AtVNI2, indicando com isto uma possível conservação funcional deste TF entre eudicotiledôneas e monocotiledôneas, incluindo a cana-de-açúcar.

### **ShWRKY12**

AtWRKY12 pertence ao grupo IIc da superfamília de TFs WRKY que está composta de três grupos (WRKY I, WRKYII e WRKY III), classificada assim, sobre a base do número do domínios WRKY e nas características dos seus motivos dedos de zinco (Eulgem et al., 2000; Wang et al., 2011). Tem sido descrito que AtWRKY12 se posiciona no nível superior da rede hierárquica de *A. thaliana* e é um repressor da deposição da SCW enquanto o silenciamento deste TF ocasionou espessamento da parede celular em células parenquimáticas da medula com ectópica deposição de lignina, xilano e celulose em eudicotiledôneas (*Medicago* e *Arabidopsis*) (Wang et al., 2010). Aqui nós identificamos que ortólogo ShWRKY12 (Sh\_226P19\_t000020) em

cana-de-açúcar se associou de forma próxima com sequências de eudicotiledôneas de AtWRKY12 (AT2G44745) (Wang et al., 2010) e PtWRKY12 (PtWRKY19;Potri.014G050000.1)(Yang et al., 2016) e de monocotiledôneas caracterizadas PvWRKY12 (PvWRKY; Pavir.Ga00648.1) e ZmWRKY12 (ZmWRKY; GRMZM2G123387\_T01) (Gallego-Giraldo et al., 2016;Rao et al., 2019) e aquelas não caracterizadas usadas aqui, correspondentes, a BdWRKY12 (Bradi5g17395.2) e SbWRKY12 (Sobic.006G166300.2). Expressão ectópica de *PtWRKY19* de *P. trichocarpa*, no mutante *Atwrky12*, resgatou suficientemente o fenótipo em células parenquimáticas da medula (Yang et al., 2016). O silenciamento de AtWRKY12 em *Medicago sativa*, *P. virgatum* e *Z. mays* produziu tanto aumento e/ou deposição ectópica de células lignificadas no parênquima medular, como da biomassa e a sua densidade, a exceção de *P. virgatum*, onde houve redução de biomassa, sugerindo o papel deste TF como repressor da lignificação tanto em monocotiledôneas e eudicotiledôneas (Gallego-Giraldo et al., 2016) incluindo portanto potencialmente a ShWRKY12 em cana-de-açúcar.

### **ShSHN1**

O SHINE (SHINE/WAX INDUCER(SHN/WIN)) tem sido considerado como regulador da biossíntese da SCW (Ambavaram et al., 2011; Liu et al., 2017). A família de TFs SHINE pertence à classe ETILENE RESPONSIVE FACTOR (ERF), que é um subgrupo de APETALA2 (AP2), uma das maiores superfamílias de proteínas regulatórias em plantas (Feng et al., 2005). Em *Arabidopsis*, o TF AtSHN2 é considerado um máster switch e foi descrito estar envolvido primeiramente na regulação da biossíntese de cutina/cera, tolerância e seca (Aharoni et al., 2004). Mas quando em *O. sativa*, o AtSHN2 foi super-expresso, resultou no aumento do conteúdo de celulose (35%), diminuição de lignina (45%) e alteração na razão S/G (Ambavaram et al., 2011). Aqui nós identificamos em cana-de-açúcar o homólogo ShSHN1 (Lócus 27034) que se associou proximamente a sequências caracterizadas de AtSHN (AtSHN2; At5g11190) (Aharoni et al., 2004) e ShSHN1 (MF537182) (Martins et al., 2018) e com aquelas identificadas mas não caracterizadas: OsSHN (OsWR1; Os02g10760) (Wang et al., 2012), PvSHN (PvERF002; Pavir.Aa02977.1) (Wuddineh et al., 2015), PtSHN (Potri.018G131400.1)(Liu et al., 2017), SbSHN (Sobic.004G084600.1), ZmSHN (GRMZM2G085678\_T01) e BdSHN (Bradi3g07450.1). A sequência identificada aqui, ShSHN1 foi 100% idêntica no que

respeita à homologia de sequência a ShSHN1 (Martins et al., 2018). A super-expressão de ShSHN1 de cana-de-açúcar em *O. sativa*, provocou um aumento de biomassa (91-140%), conteúdo de celulose (10-22%) e eficiência de sacarificação (5-53%), assim como uma redução do conteúdo de lignina (17-35%) e aumento da razão S/G (53-106%) (Martins et al., 2018). Para nosso conhecimento até o momento os únicos genes membros da família de SHN que estão associados com a regulação da SCW para as eudicotiledôneas são AtSHN2 (Aharoni et al., 2004) e PsnSHN2 (Liu et al., 2017) e para as monocotiledôneas ShSHN1 (Martins et al., 2018) e PvERF001 (Wuddineh et al., 2015). Potri.018G131400.1 foi filogeneticamente mais próxima do que PsnSH2 de *P. simonii* X *P. nigra* (Potri.018G028000.1) a ShSHN1. Apesar de que o estudo de PvERF001 foi focado para demonstrar a regulação da SCW, não se observou alteração do conteúdo de lignina nem da razão S/G, sendo a função de PvERF001 controversa, nos aqui observamos que PvERF002 foi mais próximo a ShSHN1 e ShSHN1 (Martins et al., 2018).

### **ShKNAT7**

Na classe II da família dos TFs Knotted-related homeobox (KNOX), o membro mais estudado é KNAT7. Aqui nos identificamos e observamos que o homólogo para cana-de-açúcar, ShKNAT7 (lcl|SP803280\_c97318\_g1\_i2+ ctg30250) se agrupou consistente com sequências caracterizadas de AtKNAT7 (AT1G62990) e PtKNAT (PoptrKNAT7; Potri.001G112200.1) (Li et al., 2012), OsKNAT7 (Os03g03164.1) (Wang et al., 2019; Yu, 2019) e outras sequências não caracterizadas de monocotiledôneas identificadas aqui: SbKNAT7 (Sobic.001G526200.1), ZmKNAT7 (GRMZM2G159431\_T01), PvKNAT7 (Pavir.J39882.1) e BdKNAT7 (Bradi1g76970.1). É importante destacar que a função de AtKNAT7 permanece ainda sem ficar clara, sendo que preliminarmente este TF tem sido descrito como repressor da deposição da SCW e que o silenciamento resulta em fibras interfasciculares e xilemáticas mais robustas, mas paradoxalmente no mesmo estudo, o mutante silenciado possui elementos de vasos com uma SCW fina (Li et al., 2011, 2012). Zhong et al. (2008), descreveram que houve redução da espessura da SCW em mutante silenciado por repressão dominante, mas de outro lado, foi encontrado que AtKNAT7 que afeta positivamente a biossíntese de xilanos (He et al., 2018). Estudos recentes demonstram que outro TF da Classe II KNOX: AtKNAT3 homólogo a AtKNAT7 é um ativador transcricional que conjuntamente com

AtKNAT7 podem formar um heterodímero, para promover a biossíntese da SCW em vasos xilemáticos, e de outro lado AtKNAT3 atua simultaneamente de forma antagônica com AtKNAT7 para influenciar a formação da parede secundária em fibras interfasciculares, indicando com isto que AtKNAT7 tem funções complexas com respeito a tipos de células específicos, assinalando que AtKNAT7 e AtKNAT3 regulam de forma sinérgica a biossíntese de lignina (Qin et al., 2020; Wang et al., 2020). Com respeito às monocotiledôneas o único homólogo para este tipo de plantas, caracterizado até o momento, é OsKNAT7, no qual foi verificado que possui uma função conservada de repressão da deposição da SCW em relação a AtKNAT7, mas foi evidenciando que a posição de OsKNAT na rede transcricional é de *máster switch* e interagindo com outros TFs, quando comparando com aquela de *A. thaliana*, onde AtKNAT7 se posiciona no nível inferior desta rede hierárquica (Wang et al., 2019; Yu, 2019).

### **ShMYB46/83**

O par redundante funcionalmente de TFs em *Arabidopsis* AtMYB46/83 (AtMYB46 e o seu parólogo AtMYB83), expressados especificamente em fibras e vasos, são considerados também *máster switches* e hierarquicamente abaixo dos SWNs, e tem sido também identificados como importantes reguladores da biossíntese da parede celular ativando genes da biossíntese da celulose, lignina e xilano causando ectópica deposição de SCW (Ko et al., 2009,2012; McCarthy et al., 2009; Wang & Dixon, 2012; Kim et al., 2013). Aqui nos identificamos em cana-de-açúcar o homólogo ShMYB46/83 (SCUTRZ3103F02 + lcl|SP803280\_c78017\_g1\_i1) que se associou mais proximamente com sequencias funcionalmente caraterizadas de eudicotiledôneas: AtMYB46/83 (AtMYB46, AT5G12870.1) (McCarthy et al., 2009) e PtMYB46/83 (Ptr20,Potri.009G061500.1) (McCarthy et al., 2010; Zhong et al., 2013) e monocotiledôneas caracterizadas: OsMYB46/83 (OsMYB46/83, Os12g33070), ZmMYB46/83 (ZmMYB46, GRMZM2G052606\_T01), PvMYB46/83 (PvMYB46B, Pavir.Ca02370.1) (Zhong, et al., 2011; Zhong et al., 2015) e não caracterizadas SbMYB46/83 (Sobic.008G112200.1) e BdMYB46/83 (Bradi4g06317.1). Em gramíneas, os genes ortólogos de arroz, milho e *Panicum*: OsMYB46, ZmMYB46 e PvMYB46A, respectivamente, são capazes de complementar mutantes de *myb46/myb83* de *Arabidopsis*, revelando uma conservação no controle da síntese da parede celular em eudicotiledôneas e monocotiledôneas e portanto potencialmente

em ShMYB46/83, identificado como homólogo para estes máster switches (Zhong, et al., 2011; Zhong et al., 2015).

### **MYBs: Ativadores ou repressores específicos dos polímeros da SCW**

Em *Arabidopsis*, uma bateria de TFs, controladas hierarquicamente pelas duas camadas de máster switches SWNs e MYB46/83, esta constituída principalmente pelos TFs : MYB52, MYB54, MYB58, MYB63, MYB85 MYB55, MYB61, MYB103, MYB69, MYB4, MYB32, SND2, SND3 e KNAT, envolve uma série de ativadores e repressores transcricionais específicos, posicionados no último nível da rede, desempenhando um ajuste fino da regulação transcricional da deposição da SCW, sendo que dentro deste último grupo de TFs, KNAT, MYB4 e MYB32 são considerados como repressores e os demais como ativadores (Cesarino et al., 2016; Rao & Dixon, 2018).

### **ShMYB31 e ShMYB42**

Os sintélogos ZmMYB31 (GRMZM2G050305\_T01) e ZmMYB42 (GRMZM2G419239\_T01) são os repressores melhormente caracterizados na deposição da lignina em *Z. mays*, e tem sido descrito a sua função em monocotiledôneas como *S. bicolor* e *O. sativa* (Agarwal et al., 2016; Fornalé et al., 2006, 2010; Sonbol et al., 2009) e em experimentos de expressão heteróloga em cana-de-açúcar (Poovaiah et al., 2016). Na árvore filogenética se observa que: ShMYB31 (Sh\_213J23\_t000050) e ShMYB42 (Sh\_241N24\_t000050) se agrupou consistentemente com ZmMYB31 e ZmMYB42 e outras sequencias de monocotiledôneas; SbMYB31 (Sobic.002G279100.1), OsMYB31(Os09g36730.1) e SbMYB42 (Sobic.007G177100.1) e OsMYB42(Os08g43550), no seus clados correspondentes, mas se observa que ortólogos em cana-de-açúcar estão mais próximos evolutivamente a AtMYB31 (AtMYB4;AT4G38620.1) do que AtMYB42 (AtMYB7; AT2G16720.1 (repressor da biossíntese de flavonoides e inducido baixo estresse por salinidade)) (Fornalé et al., 2014)) de *A. thaliana*. Tem sido descrito que ZmMYB31 e ZmMYB42 são filogeneticamente relacionados a AmMYB308, AmMYB330, AtMYB4 e AtMYB32 todos eles já caracterizados atuando como repressores da biossíntese de lignina, indicando com isto a potencial função de ShMYB31 e ShMYB42 como repressores da deposição de lignina em cana-de-açúcar (Fornalé et al., 2006; Jin et al., 2000; Preston et al., 2004; Tamagnone et al., 1998).

### **ShMYB85, ShMYB58/63, ShMYB61 e ShMYB52/54**

Em *Arabidopsis* os pares de TFs com redundância funcional correspondentes a: AtMYB85, AtMYB58/63, AtMY52/54 tem sido descritos como reguladores específicos de lignina, por que eles apresentam uma exclusiva ativação de todos os genes da rota biossintética de lignina com exceção de F5H (Zhao & Dixon, 2011; Zhong et al., 2008; Zhou et al., 2009). De outro lado AtMYB61, são também reguladores positivos da biossíntese da SCW e portanto de lignina em *Arabidopsis*, tendo dentro do seus alvos o TF repressor KNAT, a pectina metil metilesterase (AtPME) e AtCCoAOMT7, enzima chave que dirige o braço da biossíntese da Lignina G (Newman et al., 2004; Romano et al., 2012). OsMYB61 é capaz de modular o conteúdo de lignina nos feixes vasculares ativando genes da lignina como *CAD2*, e promover a biossíntese de celulose (Hirano et al., 2013; Huang et al., 2015).

Nós identificamos para cana-de-açúcar, que o ortólogo ShMYB58/63 (Sh\_227K15\_t000030) se associou com sequencias identificadas ou caracterizadas de eudicotiledôneas: AtMYB58/63 (AtMYB63; At1g79180) PtMYB58/63 (PtrMYB192; Potri.007G067600.1) (Zhong et al., 2011; Zhou et al., 2009) monocotiledôneas: ZmMYB58/63 (ZmMYB19; GRMZM5G833253\_T01)(Yang et al., 2017), OsMYB58/63 (Os02g46780) (Hirano et al., 2013; Noda et al., 2015), SbMYB58/63 (SbMYB60; Sobic.004G273800.1)(Scully et al., 2016), PvMYB58/63 (PvMYB58/63C; Pavir.Aa01159.1) (Rao et al., 2019) e não caracterizadas BdMYB58/63 (Bradi5g20130.1), evidência que indica ao homólogo ShMYB58/63 como potencial ativador específico da biossíntese de lignina em cana-de-açúcar.

Foi identificado em cana-de-açúcar, o homólogo ShMYB85 (Sh\_250M06\_t000170) como potencial ativador específico da deposição desde que se agrupou proximamente a sequências caracterizadas ou identificadas de eudicotiledôneas: AtMYB85 (AtMYB85,AT4G22680) (Zhong et al., 2008), PtMYB85 (PtrMYB92; Potri.001G118800.1) (Zhong et al., 2011) e monocotiledôneas: OsMYB85 (OsMYB42/85; Os09g36250.1) (Hirano et al., 2013), ZmMYB85 (Zm152; GRMZM2G104551\_T03) (Zhang et al., 2016), PvMYB85 (PvMYB42/85B; Pavir.Ba01239.1) (Bhatia et al., 2019). De outro lado ShMYB85 se posicionou proximamente a sequência de monocotiledôneas identificadas BdMYB85 (Bradi4g36210.1) e SbMYB85 (Sb02g030900) como descrito por Rao et al. (2019).

Outro potencial ativador da deposição da SCW e da biossíntese de lignina, identificado aqui para cana-de-açúcar foi o homólogo ShMYB61 (Lócus 5776). A árvore filogenética indica que ShMYB61 se agrupou com as sequências de homólogos caracterizadas de AtMYB61 (AtMYB61, AT1G09540) (Newman et al., 2004; Romano et al., 2012) e OsMYB55/61 (OsMYB55/61; Os01g18240.1) (Hirano et al., 2013; Huang et al., 2015), sendo até o momento os únicos organismos onde este TF foi identificado e caracterizado. Apesar de que as sequências de: PtMYB61(Potri.014G111200.2), ZmMYB61(GRMZM2G127490\_T01), SbMYB61(Sobic.003G136600.1), PvMYB61(Pavir.Eb01178.1) e BdMYB61 (Bradi2g11080.1) não foram identificadas, nós incluímos estas na análise filogenéticas e se agruparam proximamente no mesmo clado, quando usando como isca a ShMYB61, reforçando a provável função de ativador da SCW por parte deste TF em cana-de-açúcar.

Como mencionado acima, AtMYB52/54 (AtMYB52, AT1G17950) foi descrito como outro ativador de lignina e da SCW em *Arabidopsis* (Zhong et al., 2008), mas até o momento, em nenhum organismo modelo em eudicotiledôneas nem em monocotiledôneas este homólogo foi identificado ou caracterizado. Aqui identificamos em cana-de-açúcar, o homólogo ShMYB52/54 (Lócus 4916), usando como isca AtMYB52/54. Para reforçar o hipotético papel de que ShMYB52/54 é um potencial ativador de lignina em cana-de-açúcar incluímos na análise filogenética sequências ortólogas não caracterizadas de eudicotiledôneas e monocotiledôneas modelo que retornaram usando como isca ShMYB52/54: PtMYB52/54 (Potri.015G033600.1), SbMYB52/54 (Sobic.001G110900.1), ZmMYB52/54(GRMZM2G455869\_T01), PvMYB52/54 (Pavir.J14418.3) e BdMYB52/54 (Bradi1g10470.1), observando que estas sequências se agruparam proximamente a AtMYB52/54 e ShMYB52/54.

É importante assinalar que quando usando como iscas dos TFs: MYBs e NACs de *Arabidopsis*: AtMYB58, AtMYB63, AtMYB42, AtMYB85, AtMYB52, AtMYB54, AtMYB55, AtMYB61, AtMYB46, AtMYB83, AtNST1, AtNST2 e AtNST3/SND1 eles retornaram como um só clone para cana-de-açúcar, sendo representados neste estudo como: ShMYB58/63, ShMYB85 ShMYB52/54, ShMYB61, ShMYB46/83 e ShNST1 respectivamente, isto mesmo foi observado no trabalho de Brito et al. (2015) só para o caso de ShMYB58/63, ShMYB46/83 e ShNST1-2. Apesar da representativa cobertura no sequenciamento do genoma de cultivares híbridos cana-

de-açúcar em vários bancos de dados principalmente o do SUCEST, CTBE, Sugarcane Lignin RNA-seq Database e ultimamente o Sugarcane Genome Hub (SGH), possivelmente não tenha-se sequenciado outros genes expressando esses TFs mencionados acima. Outra possibilidade é que um gene só codifica para cada uns desse genes, apresentando redundância funcional, como o observado em *Arabidopsis* para AtNST1/NST2/SND1 (Mitsuda et al. 2005; Zhong et al.2007), AtMYB46/83 (McCarthy et al. 2009), AtMYB58/63 (Zhou et al. 2009) e AtMYB52/54 (Zhong et al.2008). O sequenciamento completo do genoma da cana-de-açúcar pode lançar luz sobre estas hipóteses (Brito et al.2015).

### **2.3 Perfil de expressão dos genes e fatores de transcrição relacionados à biossíntese de lignina em espécies de *Saccharum***

#### **Perfil de expressão de genes da rota biossintética de lignina em espécies de *Saccharum***

Os genes da rota biossintética de lignina, sequenciados e identificados nos parentais foram analisados por qPCR para melhor entendimento do controle genético. Na figura 10, de forma geral, os genes foram mais expressos em *S. spontaneum* para *C4H*, *4CL*, *C3H*, *CCoAOMT A* e *B*, *CCR* e *F5H*. *S. officinarum* teve maior expressão, no geral, para os genes *HCT*, *COMT* e *CAD B*. O gene *CAD A* teve um padrão misto, mas maior expressão em folha jovem e madura. Os internódios 3° e 5° apresentaram diferença quando na córtex e medula. *C4H* foi mais expresso na medula do que no córtex para *S. officinarum* e teve queda quando comparado em *S. spontaneum*. *4CL* não apresentou diferença nos internódios 3° para córtex e medula em ambas as espécies, mas teve queda no internódio 5° para *S. officinarum* e aumento para *S. spontaneum*. *HCT* teve aumento de expressão em ambas as espécies para córtex do internódio 3° e para córtex do internódio 5° e menor expressão geral quando comparamos córtex para medula, mas um aumento de medula 3° para 5° em *S. officinarum*. *C3H* quando analisada em todos os tecidos entre as espécies, fica claro uma maior expressão em *S. spontaneum* do que em *S. officinarum*. Além disso, uma maior expressão em córtex do internódio 3° do que na medula do internódio 3°. *CCoAOMT A* tem maior expressão em córtex e medula para internódio 5 do que internódio 3°. Entretanto, é mais expresso na medula (P3 e P5) do que córtex para *S. officinarum* e o oposto é observado em *S. spontaneum*. *CCoAOMT B* manteve a

expressão em córtex internódio 3° e um leve aumento entre medula 3° para 5° em *S. officinarum* e em *S. spontaneum* é mais expresso em tecidos do internódio 5° do que 3°. *CCR* tem destaque pela expressão relativamente maior em córtex e medula do internódio 5° em *S. spontaneum*. Em *S. officinarum* a expressão é menor em todos os tecidos e possui uma maior expressão em córtex internódio 5° do que córtex internódio 3°. Dentre os genes analisados, foi um dos mais expressos da via biossintética de lignina, acompanhado da *COMT*. *S. officinarum* não apresenta mudança de expressão entre córtex e medula para os internódios 3° e 5° para *F5H*, mas uma maior expressão em *S. spontaneum* em córtex e medula do internódio 5°. *COMT* foi o gene mais expresso dentre os analisados nos parentais de cana-de-açúcar, destacando-se a maior expressão na medula do internódio 5° para ambas as espécies. *CAD A* apresenta um padrão mais específico de folhas jovens e madura, baixa expressão em raiz e maior expressão na medula para os internódios 3° e 5° quando comparados com a córtex, respectivamente de cada internódio. *CAD B* em *S. officinarum* apresenta, entre todos os internódios nesta espécie, uma maior expressão em córtex internódio 5°. Em *S. spontaneum* temos uma maior expressão nos tecidos do internódio 5° quando comparados com o internódio 3°, tanto córtex quanto medula, respectivamente. Interessante ressaltar que os mesmos genes apresentam padrão de expressão distintos nos diferentes parentais. Isso revela um padrão complexo e distinto no controle da via biossintética de lignina. Os genes *CCR* e *COMT* merecem destaque pela maior expressão; *4CL* e *F5H* por terem um padrão de aumento em tecidos mais desenvolvidos; *C3H* e *CCR* pela maior expressão específica em *S. spontaneum*; *CAD B* por ser mais expressa especificadamente em *S. officinarum*.

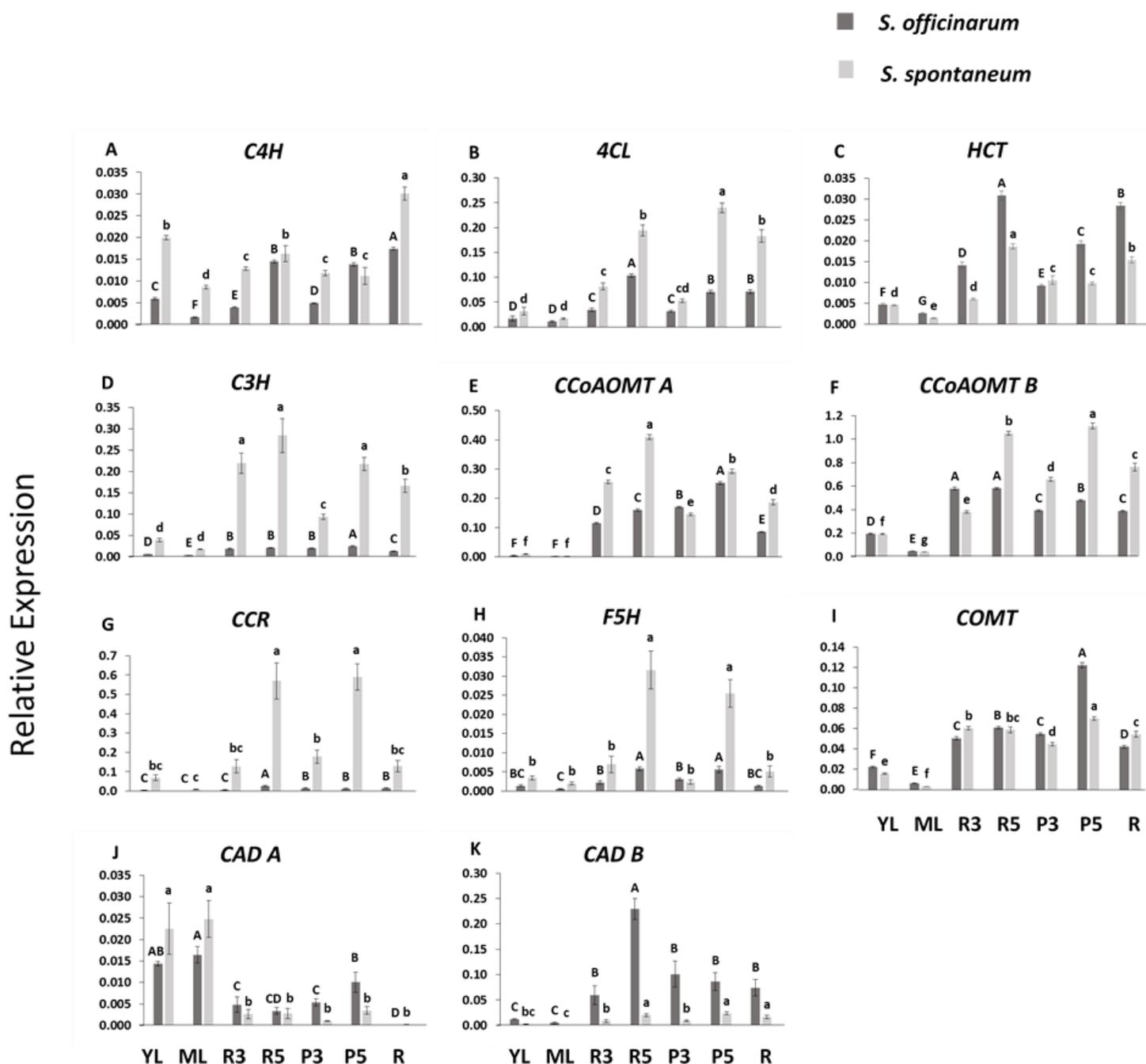


Figura 10 – Perfil de expressão dos genes da via biossintética dos monoclignóis analisados por qRT-PCR em espécies de *Saccharum* em relação ao gene de referência GADPH. YL= Folha Jovem, ML= Folha Madura, R3= córtex do internódio 3°, R5= córtex do internódio 5°, P3= medula do internódio 3°, P5= medula do internódio 5° e R= Raiz. Letras diferentes indicam diferenças significativas na expressão relativa do gene entre os tecidos de um mesmo genótipo. As medias foram comparados pelo teste de Willcoxon,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das medias de triplicatas biológicas independentes.

### **Perfil de expressão de fatores de transcrição (TFs) e a sua relação com a biossíntese e deposição de lignina em espécies de *Saccharum***

Com a informação obtida acima, usamos as sequências dos TFs ortólogos identificados para cana-de-açúcar, pertencentes à rede regulatória da biossíntese da SCW, para ser analisados por RT-qPCR em ambas as espécies (Figura 11), com o objetivo de ter um melhor entendimento do controle genético e uma ideia preliminar sobre a sua possível relação com a deposição e biossíntese de lignina em um tecido em particular e nas espécies. A expressão dos TFs identificados em cana-de-açúcar, foi dependente do tecido e do parental em estudo. De forma geral, indistintamente da região espacial (córtex ou medula) houve um aumento da expressão do internódio 3° para 5° para ambas as espécies. O córtex (R) e a medula (P) dos internódios 3° e 5° apresentaram maior expressão quando comparando com as folhas jovens, maduras e raízes, com algumas exceções para determinados TFs em ambas as espécies (Figura 11). Destacamos que entre as espécies houve um padrão específico de expressão para um TF em particular, com relação à região espacial do internódio 5°, assim *ShMYB85* e *ShMYB58/63* foram mais expressos no R5 de *S. officinarum* e no P5 em *S. spontaneum* dando indicação da função conservada destes TFs entre grupo de plantas como ativadores da deposição de lignina (Hirano et al., 2013; Scully et al., 2016; Noda et al., 2015; Rao et al., 2019; Bhatia et al., 2019; Zhou et al., 2009; Zhong et al., 2008). Um padrão comum se observou para *ShNST1*, *ShVNI2* e *ShKNAT*, *ShMYB42* desde que em ambas espécies a expressão destes TFs foi maior na R5 e de outro lado, *ShMYB61*, *ShMYB46/83*, *ShMYB31* e *ShWRKY12* tiveram maior expressão em P5 em ambas espécies, observando-se curiosamente que para cada uma das regiões do internódio 5°, houve expressão misturada de repressores e ativadores, possivelmente como um mecanismo de ajuste fino da deposição da SCW (Rao & Dixon, 2018). *ShMYB52/54* foi constitutivo desde que não houve diferenças entre R5 e P5 em *S. officinarum*, mas em *S. spontaneum* houve maior expressão em R5. *ShSHN1* em *S. officinarum* foi mais expresso em R5, enquanto *ShSHN1* em *S. spontaneum* não apresentou diferenças entre R5 e P5. Indistintamente da região espacial do internódio 5° e da espécie, o ativador específico de lignina: *ShMYB85* merece destaque pela maior expressão dentro de todos os TFs avaliados aqui. *ShMYB85* foi mais constitutivo em *S. spontaneum* desde que uma variação entre R5 e P5, não foi muito evidente, quando comparando com *S. officinarum* onde houve uma maior expressão no R5, o que tem coerência desde que o padrão espacial de

densidade de fibras e vasos xilemáticos é contrastante entre os parentais (Llerena et al., 2019). De forma contrária ao encontrado em *ShMYB85*, *ShSHN1* e *ShWRKY12* foram encontrados a ser os menos expressos entre todos os TFs avaliados aqui, o que indica o papel conservado destes TFs como másters switches, repressores da SCW entre monocotiledôneas e eudicotiledôneas (Wang et al., 2010; Yang et al., 2016; Gallego-Giraldo et al., 2016; Rao et al., 2019; Ambavaram et al., 2011; Martins et al., 2018; Liu et al., 2017). A expressão do *ShMYB31* homólogo filogeneticamente próximo do repressor específico de lignina em *Z. mays* (Agarwal et al., 2016; Fornalé et al., 2006, 2010; Sonbol et al., 2009) apresentou um perfil comum em ambas espécies sendo que a maior expressão destes TFs foi maior para o P5 o que pode ter sentido desde que a córtex tem mais fibras e vasos do que na medula (Llerena et al., 2019). Este perfil de expressão dos TFs encontrados no internódio 5° das espécies, possivelmente indica que os distintos TFs tem similares funções, mas são controlados de forma diferente e que dependem do parental e do tecido (córtex e medula).

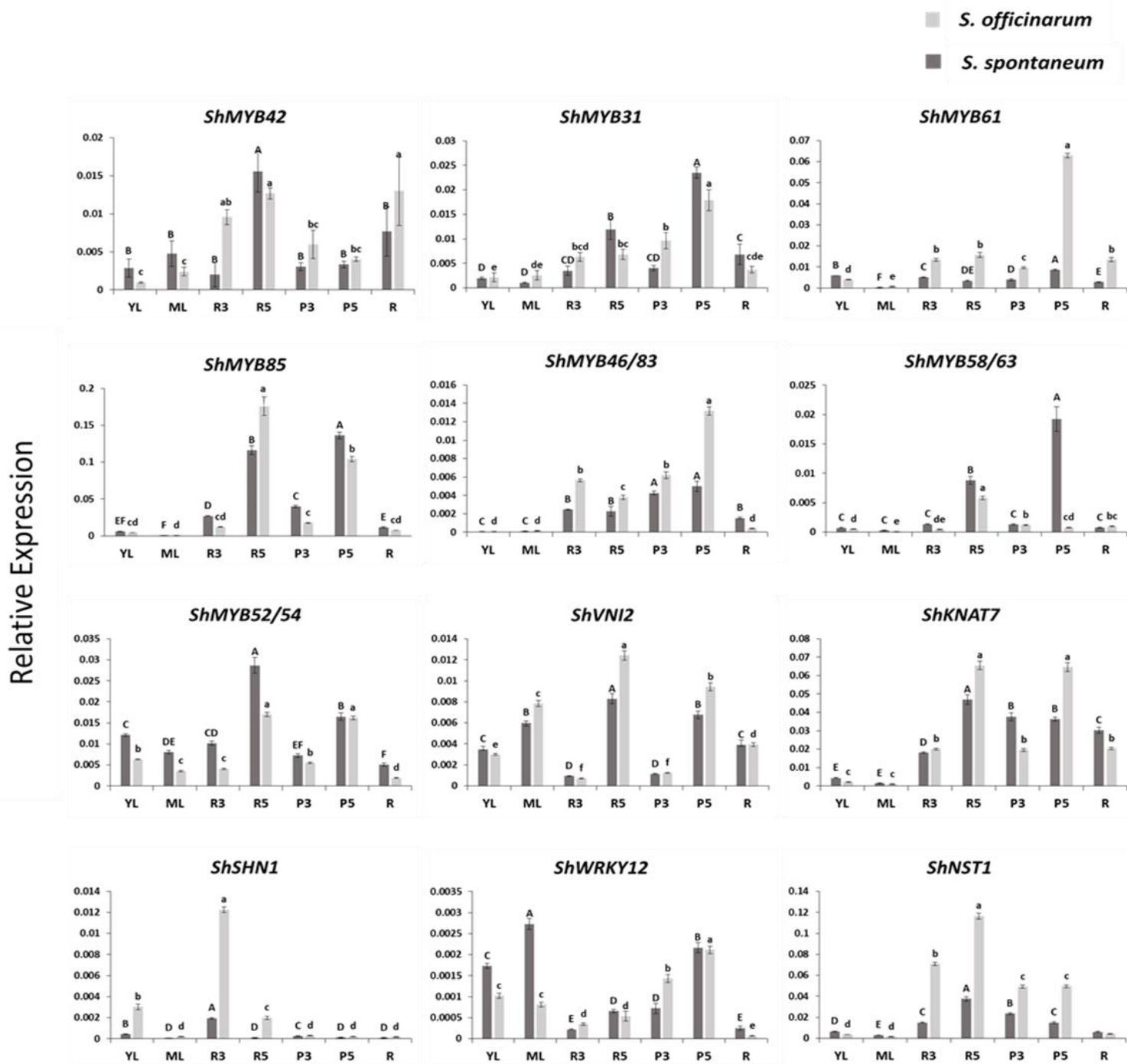


Figura 11 – Perfil de expressão de TFs relacionados com a regulação da deposição de SCW, analisado por qRT-PCR em espécies de *Saccharum* em relação ao gene de referência GADPH. YL= Folha Jovem, ML= Folha Madura, R3= córtex do internódio 3°, R5= córtex do internódio 5°, P3= medula do internódio 3°, P5= medula do internódio 5° e R= Raiz. Letras diferentes indicam diferenças significativas na expressão relativa do gene entre os tecidos de um mesmo genótipo. As medias foram comparados pelo test pos-hoc de Willcoxon,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das medias de triplicatas biológicas independentes

Uma análise de correlação de Pearson foi desenvolvida com o intuito de encontrar possíveis relações entre o perfil de expressão dos TFs identificados, a deposição de lignina (conteúdo e composição) e genes da biossíntese de lignina nas espécies de *Saccharum* usando para cada um destes parâmetros as mesmas amostras (7 tecidos) (Figuras 12 A e B). Os TFs homólogos identificados em cana-de-açúcar considerados como másters switches regulando positivamente a deposição da SCW: *ShNST1* e *ShMYB46/83* apresentaram uma correlação do conteúdo de lignina e S/G de: 0.45; 0.52 e -0.25; 0.19, respectivamente em *S. officinarum* e de outro lado a correlação foi de: -0.12; -0.28 e 0.86; 0.32, respectivamente em *S. spontaneum*, sugerindo um papel mais específico na deposição de lignina por parte de *ShNST1* em *S. officinarum* e *ShMYB46/83* em *S. spontaneum*. Dentro dos grupos dos ativadores específicos da deposição de lignina, é importante destacar que os TFs homólogos identificados em cana-de-açúcar: *ShMYB85*, *ShMYB58/63* foram os que tiveram a maior correlação positiva no conteúdo de lignina e razão S/G: 0.67; 0.97 e 0.78; 0.71, respectivamente em *S. officinarum* e 0.72; 0.80 e 0.8; 0.93, respectivamente em *S. spontaneum*. Em concordância com a razão S/G, *ShMYB85* e *ShMYB58/63* tiveram valores altos de correlação com a expressão de *F5H* em *S. officinarum* : 0.93; 0.61, respectivamente e *S. spontaneum* 0.92; 0.76, respectivamente, indicando uma possível modulação direta por parte desses TFs na biossíntese de *F5H* que dirige o braço que leva a biossíntese da lignina S nos parentais, operando de forma distinta a Arabidopsis onde a região regulatória de *F5H* não possui motivos AC para a ligação destes TFs, sendo descrito que a abundância de transcritos da *F5H* é governada por SND1 e MYB103 (Öhman et al., 2013; Zhao et al., 2010). As observações aqui apresentadas no que respeita aos parentais tem sido descritas consistentemente nos estudos em monocotiledôneas tais como : *Z. mays*, *B. distachyon*, *P. virgatum* e *S. bicolor* (Bhatia et al., 2019; Rao et al., 2019; Scully et al., 2016). Para os homólogos repressores específicos da biossíntese de lignina de *Z. mays* (Agarwal et al., 2016; Fornalé et al., 2006, 2010; Sonbol et al., 2009), identificados em cana-de-açúcar: *ShMYB31* e *ShMYB42*, foi encontrado uma correlação para o conteúdo e composição de lignina de: 0.08; 0.41 e 0.75; 0.59, respectivamente em *S. officinarum* e 0.82; 0.92 e -0.07 e 0.05, respectivamente em *S. spontaneum*, indicando uma possível divergência funcional entre estes sintélogos nas espécies, sendo *ShMYB31* um repressor em *S. officinarum* e em *S. spontaneum*, *ShMYB42*. Outros homólogos identificados em cana-de-açúcar, como repressores relacionados também com a deposição da SCW

que não pertencem à família MYB: *ShWRKY12*, *ShKNAT* e *ShSHN*, apresentaram pelo menos para uma das espécies uma correlação negativa ou baixa correlação no que respeita ao conteúdo e composição de lignina, assim em destaque *ShSHN1* homólogo de *AtSHN2* apresentou correlações negativas de -0.79; -0.62, respectivamente para *S. officinarum* e -0.79 e -0,2, respectivamente para *S. spontaneum*, além disso foi observado para todos os genes de lignina de ambas espécies uma muito baixa ou negativa correlação com o conteúdo e a razão S/G. Estas correlações para o homólogo repressor *ShSHN1* identificado aqui, podem indicar uma potencial funcionalidade conservada entre monocotiledôneas e eudicotiledôneas desde que foi observado que a super-expresssão de SHN de cana-de-açúcar em *O. sativa*, assim como também a de *PsnSH2* em *P. simonii* X *P. nigra* provocou redução do conteúdo de lignina e genes da biossíntese deste polímero (Martins et al., 2018; Liu et al., 2017). Controversamente a estas observações outro homólogo de *AtSHN2* em *P. virgatum*: *PvERF001*, foi focado para demonstrar a regulação da SCW, mas não se observou alteração do conteúdo de lignina nem da razão S/G (Wuddineh et al., 2015). *ShWRKY12* teve uma baixa correlação de 0.13; 0.24, respectivamente em *S. officinarum*, indicando uma conservação da função como repressor. Em *S. spontaneum* uma alta correlação 0.94; 0.78, respectivamente foi observada. Apesar de que *ShWRKY12* em *S. spontaneum* apresentou elevados coeficientes, para ser considerado como repressor, observamos correlações baixas e negativas para genes *SsC3H* (0.03), *SsC4H* (-0.49), *SsCCoAOMT2* (*CCoAOMT A*) (0.09) e *SsHCT* (0.00). Para *ShKNAT7* se observa que ambas espécies possuem um similar coeficiente de correlação, relacionado ao conteúdo de lignina (0.5-0.53) porém estas diferem na razão S/G, sendo que em *S. officinarum* e *S. spontaneum* houve uma alta e baixa correlação (0.97 e 0.07, respectivamente). Uma plausível explicação pode estar ligada à enzimas que governam a biossíntese da braço S da biossíntese de lignina desde que em *S. officinarum* houve um alto coeficiente no que respeita a *SoCOMT* e *SoF5H* (0.65, 0.97, respectivamente) em relação a *SsCOMT* e *SsF5H* (-0.13 e 0.62, respectivamente) de *S. spontaneum*. *AtVNI2* interage com alta afinidade com VNDs, atuando como repressor transcricional da diferenciação em vasos xilemáticos (Kubo et al., 2005; Yamaguchi et al., 2010). *ShVNI2* identificado em cana-de-açúcar, apresentou um similar coeficiente para conteúdo de lignina e maior para a composição de lignina em *S. spontaneum* do que *S. officinarum* (0.49; 0.62 e 0.67; 0.99, respectivamente), correlação que teria uma possível explicação desde que os

genes que dirigem o braço G da biossíntese de lignina: *ShCCoAOMT2* (*CCoAOMT A*) e *ShCCoAOMT1* (*CCoAOMT B*) tiveram uma correlação menor para *S. officinarum* do que *S. spontaneum* (0.5; 0.34 e 0.85; 0.94, respectivamente), influenciando na correlação observada a razão S/G nas espécies.

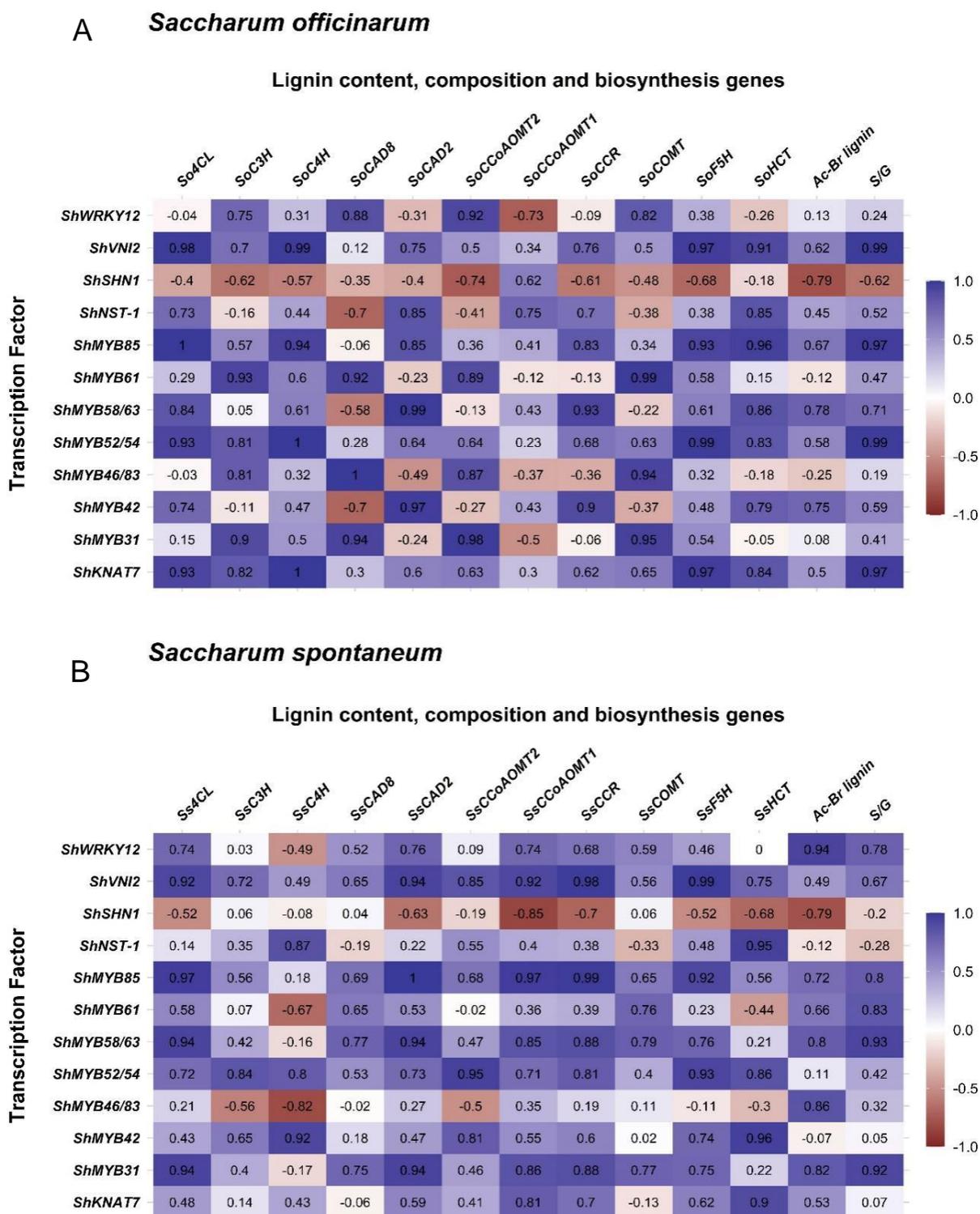


Figura 12 – Análise de correlação de Pearson entre dados do perfil de expressão dos TFs identificados, a deposição de lignina (conteúdo e composição) e genes da biossíntese de lignina nas espécies de *Saccharum* sendo todos avaliados nas mesmas amostras correspondentes a 7 tecidos. Os dados variam desde 1 (100% de positiva correlação- azul escuro) até -1(100% de negativa correlação- vermelho escuro). O valor de 0 indica nenhuma relação.

## 2.4 Análise de componente Principal (PCA)

Foi feita uma análise de componente principal (PCA) para simplificar as tendências e verificar possíveis agrupamentos entre a expressão de TFs, genes de biossíntese de lignina e parâmetros metabólicos (razão S/G e conteúdo de lignina) para cada uma das espécies de *Saccharum*, e revelar as variáveis mais importantes para cada agrupamento em apenas dois componentes (dimensões) que resumem todas as variáveis avaliadas. A análise de componente principal (PCA) para todos os parâmetros é apresentado na figura 13. Para *S. officinarum*, o primeiro (PC<sub>1</sub>) e o segundo (PC<sub>2</sub>) componente, representaram o 53.24 % e 37.12% da total variação, respectivamente. (Figura 13A). Para *S. spontaneum*, os dois componentes representam o 85.3% do total da variabilidade dos dados, com o PC<sub>1</sub> e o PC<sub>2</sub> representando 57.43% e 27.87% da variabilidade, respectivamente (Figura 13B). Tanto em *S. officinarum* como para *S. spontaneum* os parâmetros e a distribuição das variáveis no PC<sub>1</sub> revelam em destaque que o homólogo repressor da deposição da SCW *máster switch*: *ShSHN1* se agrupou de forma distante com os parâmetros de dois clusters relacionados de conteúdo de lignina, razão S/G, e a expressão dos restantes TFs e genes da biossíntese de lignina avaliados, reforçando as tendências encontradas nos baixos ou negativos coeficientes de correlação de Pearson, evidenciando a este TF homólogo como um potencial repressor da deposição de lignina nas espécies de *Saccharum*. No que respeita ao PC<sub>2</sub> em *S. officinarum* se observa um *cluster* que contem TFs repressores como *ShWRKY12*, *ShMYB31*, o ativador *máster switch* *ShMYB46/83* e o *ShMYB61* associando-se com os genes: *SoCAD A (ShCAD8)*, *SoCCoAOMT A (ShCCoAOMT2)*, *SoC3H* e *SoCOMT* indicando que os TFs e genes mencionados preferencialmente possuem maior expressão no P5, região espacial do internódio 5 de *S. officinarum* com menor proporção de fibras e vasos xilemáticos e portanto com menor conteúdo de lignina em relação a R5 (Llerena et al., 2019). O gene mais expresso da rota biosintética *SoCCoAOMT B (ShCCoAOMT1)* teve uma relação distante com *ShSHN1* coerente com a correlação de P5. Outro *cluster* se agrupou de forma separada onde podemos destacar a inclusão de três grupos proximamente relacionados entre si e definidos pelo conteúdo de lignina acetil bromide, razão S/G, e os TFs ativadores específicos de lignina *ShMYB85* e *ShMYB58/63*, novamente coincidindo com as tendências encontradas nas correlações de Pearson. A razão S/G se relacionou com o ativador *ShMYB52/54*, os

repressores *ShVNI2*, *ShKNAT*, *SoC4H* e interessantemente com o *SoF5H*. De outro lado o conteúdo de lignina se associou proximamente com o ativador *ShMYB85*, *SoHCT*, *SoCCR* e um dos genes mais expressos da rota biossintética de lignina: *So4CL*. *ShMYB58/63* se agrupou com o *SoCAD B (ShCAD2)*, *ShNST1* e *ShMYB42*. Curiosamente os TFs envolvidos neste *cluster* e os genes da rota biossintética da lignina foram mais expressos em R5, onde existe maior proporção de fibras e vasos e maior conteúdo de lignina (Llerena et al., 2019). O gene mais expresso da rota biossintética: *SoCCoAOMT B (ShCCoAOMT1)* teve uma relação distante com *ShSHN1* coerente com a correlação de Pearson. No que respeita ao PC<sub>2</sub> em *S. spontaneum* se observou um *cluster*, onde os TFs: *ShNST1*, *ShMYB42*, *ShMYB52/54*, *ShKNAT7* e *ShVNI2* se associaram com *SsC4H*, *SsHCT*, *SsCCoAOMT A (ShCCoAOMT2)*, *SsC3H* e *SsF5H*, respectivamente. Os mencionados genes foram expressos principalmente em R5 que de acordo com o conteúdo de lignina acetil bromide em *S. spontaneum* é menor do que o P5. Outro *cluster* se agrupou de forma separada incluindo três grupos, dos quais, dois foram próximos: *ShMYB85* que se relacionou proximamente com alguns genes mais expressos em *S. spontaneum*: *SsCCR*, *SsCCoAOMT B (ShCCoAOMT1)*, *Ss4CL* assim como com *SsCAD A (ShCAD8)* e *SsCAD B (ShCAD2)* e em destaque *ShMYB58/63* se associou com a razão S/G o conteúdo de lignina, com o gene *SsCOMT* e com *ShMYB31*. Os componentes dos grupos onde ficaram inclusos *ShMYB85* e *ShMYB58/63* apresentaram de moderadas a altas correlações na análise de Pearson. Um terceiro grupo mais disperso no que respeita aos seus componentes envolveu aos TFs: *ShMYB46/83*, *ShMYB61* e o homólogo repressor *ShWRKY12*. Neste *cluster* os TFs e genes da rota biossintética em *S. spontaneum* foram mais expressos em P5, que segundo o conteúdo de lignina acetilbromide avaliado este tecido foi mais lignificado do que R5.

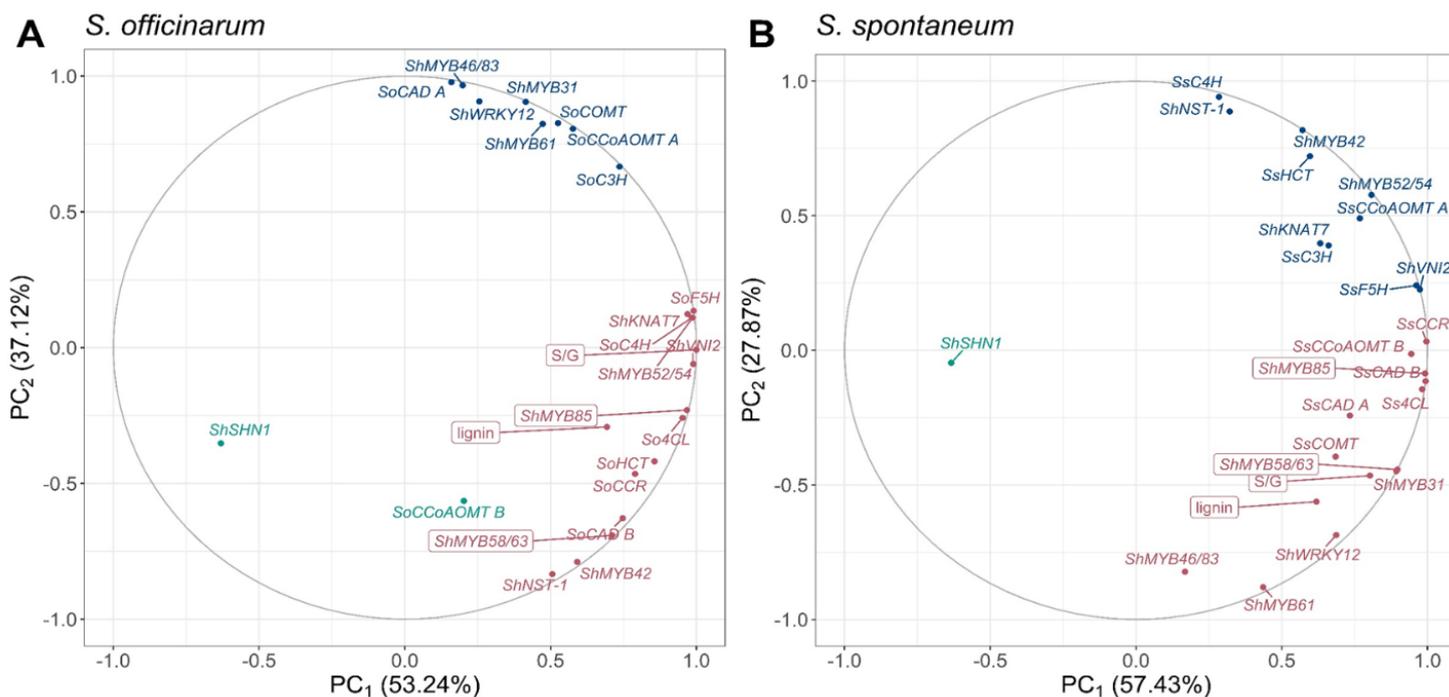


Figura 13 – PCA do metabolismo em espécies de *Saccharum*. Gráficos de variáveis de correlação de 11 genes da rota biossintética de lignina, 12 TFs reguladores da biossíntese da parede celular (R2R3-MYB, NAC, WRKY, KNOX e AP2), conteúdo de lignina e razão S/G de 7 tecidos de (A) *S. officinarum* e (B) *S. spontaneum*. Números em parêntese correspondem ao percentual de variação explicado pelo primeiro e segundo componente principal. As cores do texto indicam o algoritmo de agrupamento de médias usado para classificar as variáveis em 3 grupos.

## 2.5 Modulação da deposição e de genes da rota da biossíntese de lignina por Fatores de transcrição em espécies de *Saccharum*

Os dados correspondentes ao perfil transcricional dos genes da biossíntese de lignina, a razão S/G e o conteúdo de lignina acetilbromide de amostras de 7 tecidos em ambas espécies de *Saccharum* foram usados com os dados de expressão dos TFs relacionados com a regulação da SCW para desenvolver uma análise de rede Bayesiana, com o intuito de encontrar evidências de possíveis interações entre os TFs e os parâmetros anteriormente mencionados (Figura 14) Distintas interações resultaram da análise, observando em alguns casos interações entre genes, outros conectando genes e dados bioquímicos, ou mesmo com os tecidos analisados. Assim, interessante podemos indicar que o parâmetro relacionado ao conteúdo lignina acetilbromide apresentou uma positiva relação com o gene *4CL* (um dos mais expressos nas ambas as espécies), que é um gene chave na ativação dependente de

CoA do ácido cafeico, ferúlico e sinápico. *4CL* teve também interação positiva somente com os ativadores específicos da deposição de lignina: *ShMYB85* e *ShMYB58/63*. Em destaque, na rede Bayesiana se observa que *ShMYB85* ocupa um lugar central na rede Bayesiana desde que interage de forma positiva com a maioria dos genes da biossíntese de lignina: *C3H*, *CAD B (CAD2)*, *CCR*, *CCoAOMT A (ShCCoAOMT2)*, *F5H* e *HCT* e como mencionado acima foi constitutivo na sua expressão no que respeita a R5 e P5 em *S. spontaneum* (genótipo mais lignificado). A interação observada aqui entre *F5H* e *ShMYB85* é consistente com o descrito em *P. virgatum* (Rao et al., 2019). De outro lado *ShMYB58/63* se correlacionou positivamente com *CCoAOMT B (ShCCoAOMT1)*, *CCR* e interessante com *COMT*, como descrito para *S. bicolor* (Scully et al., 2016) e *P. virgatum* (Rao et al., 2019). Além disso *ShMYB85* interagiu positivamente com os ativadores: *ShMYB58/63* e *ShMYB52/54* e o repressor *ShKNAT7* revelando uma possível interação destes TFs para um ajuste fino na regulação da deposição de lignina nas espécies de *Saccharum* (Rao & Dixon, 2018). *ShMYB61* apresentou uma interação negativa com *ShMYB85* e positiva com *ShKNAT*. Em *Arabidopsis* *AtKNAT7* tem sido descrito que é alvo de *AtMYB61*, mas não de *AtMYB85* (Romano et al., 2012). *MYB61* em *Arabidopsis* tem sido considerado como um regulador transcricional que permite a alocação dos recursos da planta inteira em tecidos de armazenamento (Li et al., 2012; Romano et al., 2012) o que permitiria explicar plausivelmente sua maior expressão em P5, que constitui o tecido de reserva nos parentais. O repressor *ShMYB31* se correlacionou positivamente com *ShVNI2* (por ser repressores da SCW e com *ShMYB61* por a maior expressão em M5). P5 teve uma interação positiva com *ShMYB31*, *ShWRKY12* e *ShMYB46/83* desde que estes TFs em ambas as espécies tiveram a maior expressão neste tecido. De forma contrária também para ambas as espécies, houve correlação positiva entre R5 e *ShNST1*, *ShMYB42* e *ShVNI2* desde que foram mais expressos nesse tecido. Tecidos como R3 e P3 menos lignificados que R5 e P5 se correlacionaram positivamente com a expressão dos potenciais repressores em cana-de-açúcar, como *ShSHN1* e *ShKNAT*, respectivamente. A interação encontrada entre o padrão de expressão de *ShSHN1* e R3/P3 onde o processo de lignificação é ainda incipiente (Llerena et al., 2019) foi também encontrado em tomate (*SISHN*) onde a maior expressão foi detectada na córtex em estágios iniciais de desenvolvimento da fruta (Mintz-Oron et al., 2008), enquanto em *Arabidopsis* a expressão de *AtSHN1* foi detectado em tecidos jovens como pétalas de

botões muito jovens, fechados e no gineceu em desenvolvimento (Aharoni et al. 2004), portanto é possível inferir que a expressão de *SHN* está positivamente correlacionada com tecidos jovens em diferentes grupos de plantas. R3 interagiu positivamente com o repressor e o ativador *master switch* *ShSHN* e *ShNST1*, respectivamente, indicando possivelmente esta correlação que a deposição da SCW em R3 é regulada por estes TFs em um ajuste fino. Curiosamente R3 interage positivamente com *C3H* e *CCoAOMT A* (*ShCCoAOMT2*) devendo indicar que não todas as interações encontradas aqui são suportadas pela literatura. As raízes tiveram uma interação positiva com *C4H*, *HCT* e com o TF *ShMYB42* que retornou como homólogo do repressor *ZmMYB42*. O homólogo *ShMYB42* identificado em cana-de-açúcar teve maior proximidade filogenética com *AtMYB42* (*AtMYB7*) de *A. thaliana* (que é um repressor da biossíntese de flavonóides e induzido por estresse salino), portanto inferimos que a relação encontrada indica que *ShMYB42* regularia negativamente os genes que alimentam a rota das substâncias flavonóides dirigindo metabólitos à rota biossintética da lignina (Fornalé et al., 2014). ML (folha velha) se correlacionou positivamente *CAD A* (*ShCAD8*) e o repressor de *masters switch*: *VNI2*, enquanto YL (folha nova) só com este último TF. *CAD A* (*ShCAD8*) foi o gene da rota biossintética da lignina mais expresso no que respeita nas folhas de ambas espécies quando comparando com os outros tecidos em estudo e de outro lado parece coerente a associação com *VNI2* (*AtVNI2* interage com alta afinidade com VNDs, atuando como repressor transcricional da diferenciação em vasos xilemáticos (Kubo et al., 2005; Yamaguchi et al., 2010) desde que as folhas são tecidos menos lignificados do que os internódios. *S. officinarum* apresentou uma positiva interação com *CAD B* (*ShCAD2*) e negativa com *F5H*, isto é plausível desde que a expressão de *CAD B* (*ShCAD2*) foi exclusiva deste parental e em *F5H* foi residual quando comparando com *S. spontaneum*. Foi encontrado que *S. spontaneum* interagiu positivamente com *C3H* e negativamente com *ShNST1*, relação que tem lógica desde que a expressão de *C3H* foi exclusiva deste genótipo frente ao observado em *S. officinarum* e porque a expressão *ShNST1* neste parental foi maior em *S. spontaneum*.

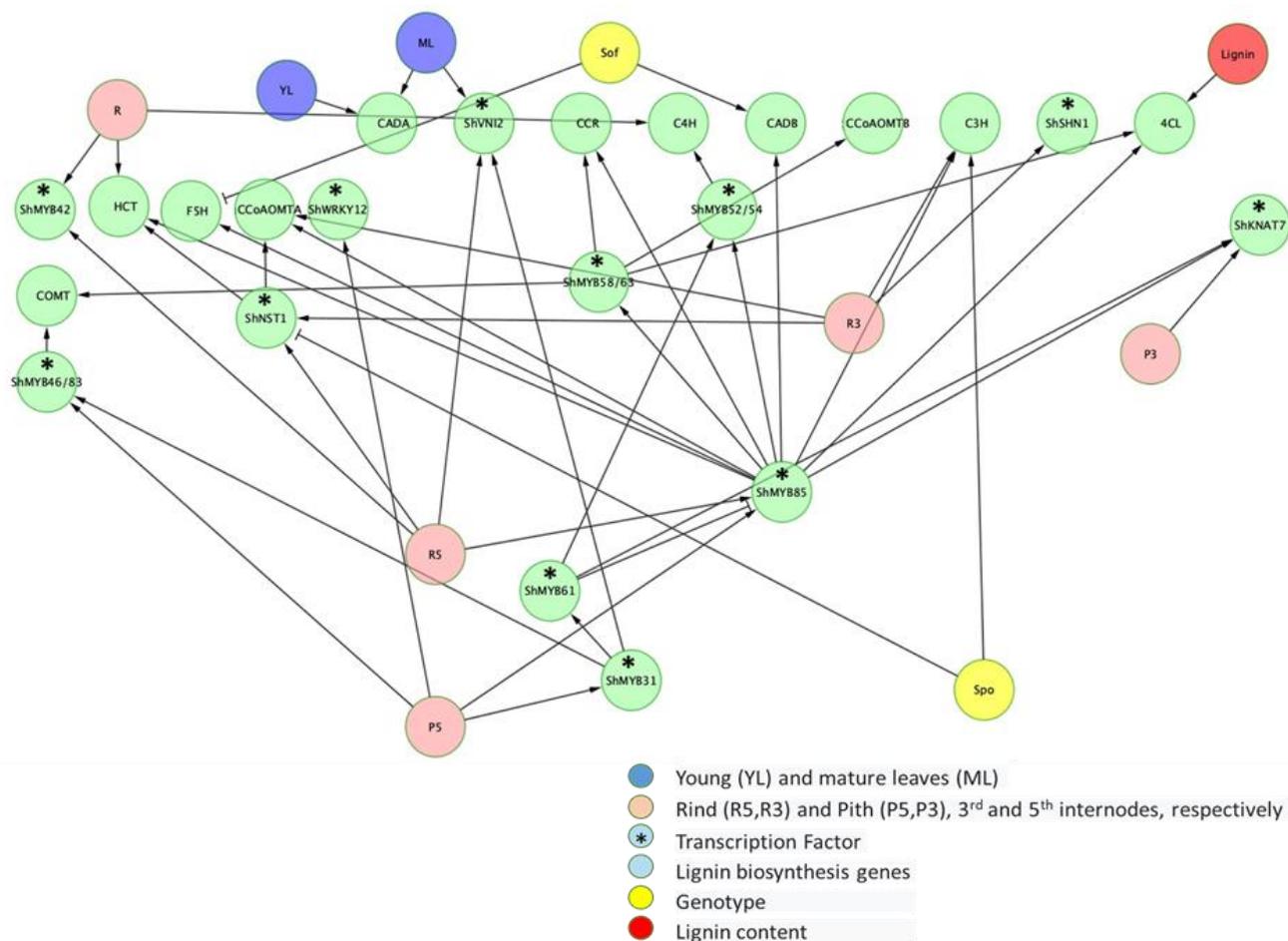


Figura 14 – Rede Bayesiana correspondente à expressão de fatores de transcrição (TFs) relacionados com a regulação da parede celular secundária, genes da biossíntese de lignina, e o conteúdo de lignina acetilbromide de amostras de 7 tecidos em ambas as espécies de *Saccharum*. YL e ML (folha nova e velha). R3; R5 e P3; P5, corresponde a córtex e medula do terceiro e quinto internódio, respectivamente. R(raiz). O círculo em vermelho indica o conteúdo de lignina. Sof e Spo, indicam: *S. officinarum* e *S. spontaneum*. Os genes relacionados à biossíntese da lignina são: cinamato 4-hidroxilase (C4H), 4-hidroxicinamoil CoA ligase (4CL), hidroxicinamoil CoA: shiquimato/quinato hidroxicinamoiltransferase (HCT), *p*-coumarato 3-hidroxilase (C3H), cafeoil-CoA O-metiltransferase (CCoAOMT A e B), cinamoil-CoA redutase (CCR), ferulato 5-hidroxilase (F5H), cafeato O-metiltransferase (COMT) e cinamil álcool desidrogenase (CAD A e B). Nas redes obtidas, foi assumido que a expressão de genes pode depender de condições (características genéticas, morfológicas e bioquímicas), mas não o contrário. Além disso, a orientação das interações regulatórias foi inferidas. Foi considerado os dados quantitativos de PCR como variáveis contínuas e outros dados (genótipo, tipo de tecido, razão S/G, conteúdo de lignina) como variáveis discretas. Para os TFs, os nomes dos genes identificados para os parentais foram baseados em sequencias caracterizadas de *Arabidopsis thaliana* e *Zea mays*.

## 2.6 Transativação de Protoplastos

Depois de analisar os resultados correspondentes à correlação de Pearson, análise de componente principal (PCA) e de redes Bayesianas que foram desenvolvidos com dados de expressão dos TFs identificados em cana-de-açúcar, genes da rota biossintética de lignina, conteúdo de lignina acetilbromide e razão S/G em tecidos das espécies de *Saccharum*, foi evidente que os TFs: ShMYB85 e ShMYB58/63 possuem em destaque um potencial papel como ativadores específicos da deposição de lignina. Para confirmar diretamente se essas correlações indicam um papel na regulação transcricional da deposição de lignina, se desenvolveu um ensaio de transativação de protoplastos, separadamente entre os TFs ShMYB85 e ShMYB58/63 (contidos nos vetores efetores) com as regiões regulatórias dos genes chaves dos parentais: *COMT*, *F5H* e *CAD A (CAD8)* (inclusos no vetores reporters) que governam o braço que determina a biossíntese de lignina S ( influenciando na razão S/G ,e portanto na sacarificação) e no conteúdo em lignina, respectivamente em ambos parentais (Figuras 15 e 16).

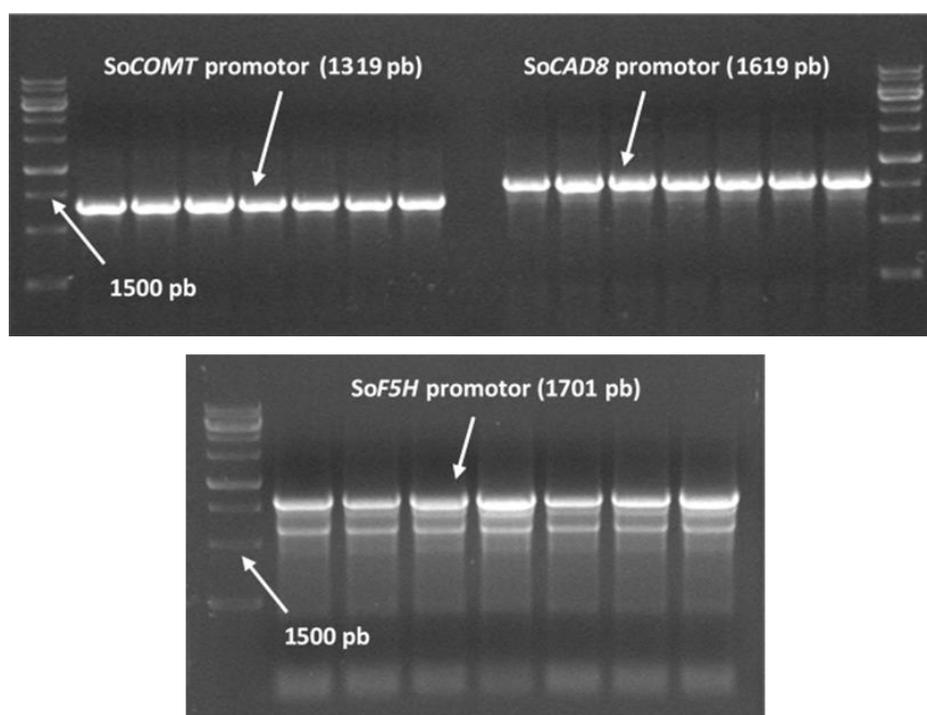


Figura 15 – Avaliação de plasmídeos pGWL7 (vetores de destino) depois da recombinação via LR clonase para verificar a presença do promotor dos genes *SoF5H*, *SoCOMT* e *SoCAD8* em *S. officinarum* por PCR usando primers attb1 e attb2 em gel de agarose ao 1% com brometo de etídeo.

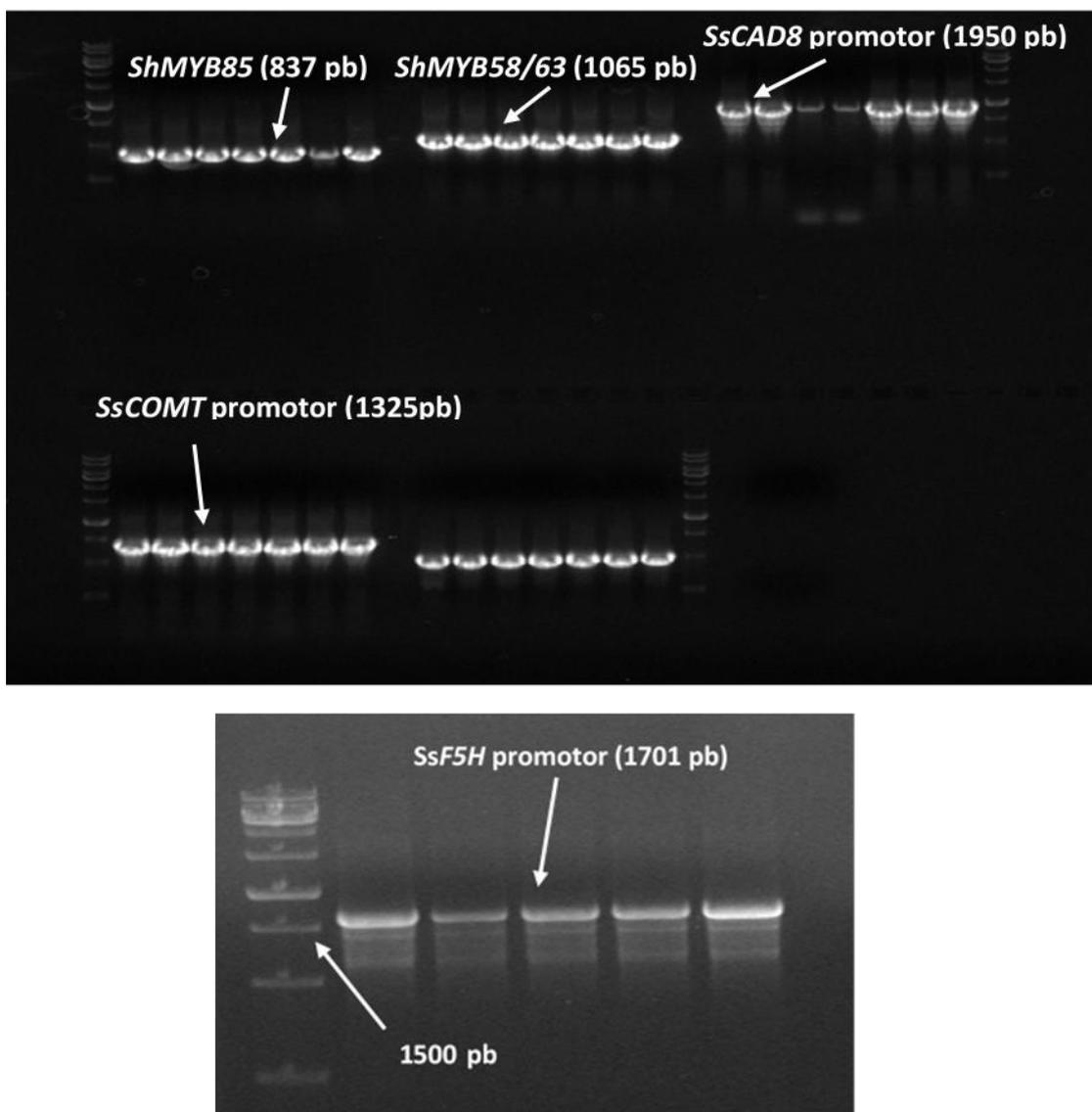


Figura 16 – Avaliação de plasmídeos p2GW7 e pGWL7 (vetores de destino) depois da recombinação via LR clonase para verificar a presença dos TFs: *ShMYB85* e *ShMYB58/63* e os promotores dos genes *ShCAD8*, *ShCOMT* e *ShF5H* em *S. spontaneum* por PCR usando primers attb1 e attb2 em gel de agarose ao 1% com brometo de etídeo.

Na figura 17 se observa que *pShCOMT* nas espécies, foi em destaque significativamente o promotor mais ativado por parte de *ShMYB58/63* e *ShMYB85*, sendo que a indução por este último TF foi maior do que *ShMYB58/63* em *S. officinarum* e *S. spontaneum*. O *pShF5H* em *S. officinarum* foi ativado por ambos TFs, sendo que houve uma maior indução por *ShMYB85*. Uma ligeira ativação não significativa de *pShF5H* de *S. spontaneum*, foi observada como também para *pShCAD8* de ambas espécies, mas não para nenhum dos casos houve diferenças

significativas de ativação entres os TFs. Uma interação que é importante assinalar é aquela observada entre os TFs: ShMYB85, ShMYB58/63 e os promotores da *pShF5H* indicando ativação significativa em *S. officinarum* e também em *S. spontaneum* apesar desta não foi significativa. Tem sido descrito que em *Arabidopsis* e *Medicago truncatula*, MYB58/63 aparece a ativar todos os genes da rota da biossíntese da lignina com exceção de *F5H* desde que o promotor desse gene nessas plantas carecem de motivos AC (Zhao & Dixon, 2011). De outro lado em *Arabidopsis*, a abundância de transcritos de *F5H* é influenciado positivamente por SND1 e MYB103 (Öhman et al., 2013). EgMYB2 um TF homólogo de AtMYB46/83 induz a expressão de *F5H* (Goicoechea et al., 2005) e o repressor em *Z.mays* (ZmMYB42) aparece a regular negativamente a *F5H*, indicando que o promotor deste gene possui motifs AC-like (Fornalé et al., 2010). A super-expresssão de *PvMYB4* de *P. virgatum* em tabaco e em *P. virgatum* provocou em ambas plantas a redução de expressão de *F5H* (Shen et al., 2012). Portanto a regulação da expressão de *F5H* é um caso particular no qual cada espécie de plantas pode ter diferentes mecanismos transcricionais, indicando que os parentais de cana-de-açúcar podem regular a biossíntese da lignina de uma forma espécie-específica. Dados que respaldam a interação de ShMYB85 e ShMYB58/63 com o promotor *F5H*, são aqueles obtidos por nós aqui no que respeita à correlação de Pearson entre os TFs, *F5H* e a razão S/G, onde *ShMYB58/63* e *ShMYB85* tiveram um coeficientes de 0.93, 0.61 e 0.97 e 0.71 respectivamente em *S. officinarum* em quanto os mesmos TFs tiveram coeficientes de 0.92, 0.76 e 0.8, 0.93 em *S. spontaneum*. Além disso ultimamente o único trabalho até o momento similar a nosso resultados obtidos na transativação foi o reportado em *P. virgatum* onde *PvMYB42/85A* e *PvMYB58/63* transativaram os promotores de *COMT* e *F5H* (Rao et al., 2019). Nossos resultados confirmam a capacidade dos TFs em regular a expressão destes genes representativos na biossíntese de lignina nas espécies de *Saccharum*, mas também é importante deixar claro que a modulação da expressão desses genes da biossíntese da lignina pode não ser só produto de um reconhecimento direto da sua região promotora por esses TFs, talvez seja também por co-ativação adicional por parte de outro TFs.

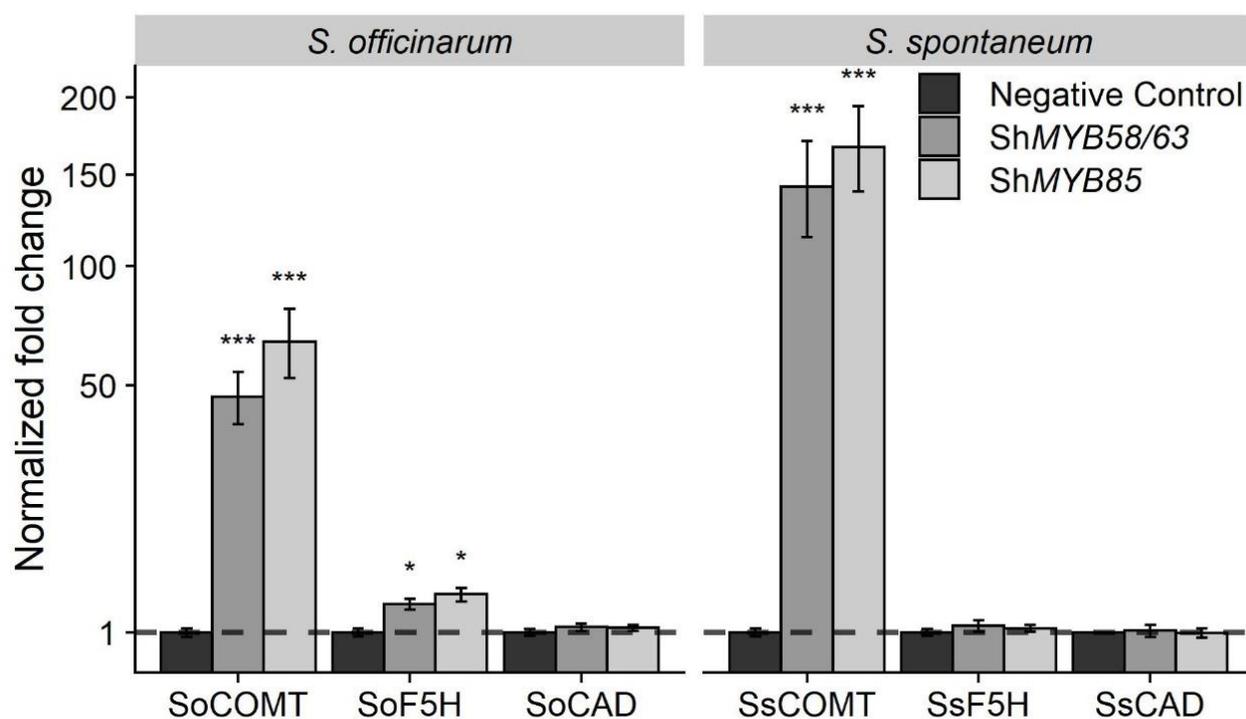


Figura 17 – Ensaio de transativação em protoplastos BY2 de *N. tabacum*. A figura amostra o *fold change* da ativação dos promotores *pShCOMT*, *pShF5H* e *pShCAD8* induzidos por *ShMYB58/63* e *ShMYB85* em *S. officinarum* (So) e *S. spontaneum* (Ss), em comparação da a sua atividade basal (controle negativo, gene GUS que não é fator de transcrição (TF)). Os dados são o produto da média de 6 replicatas  $\pm$  SE. Diferentes letras minúsculas indicam diferenças na ativação de um promotor baixo interação de um TF em particular. Diferenças significativas para a ativação de *pShCOMT*, *pShCAD* e *pSHF5H* foram determinadas pelo t-test ( $P < 0.05$ ).

### 3. Conclusões

O isolamento e análise *in silico* dos promotores de genes representativos da rota biossintética da lignina: *CAD8*, *F5H* e *COMT* nas espécies de *Saccharum*, permitiu de forma preliminar ter uma visão do repertório de TFs repressores e ativadores relacionados com a rede transcricional da regulação da SCW, existentes nas regiões regulatórias destes genes. A partir destes TFs identificados e outros reportados na literatura foram encontradas sequencias homólogas em cana-de-açúcar, sendo submetidas a uma análise filogenética, observando uma associação próxima com sequencias caracterizadas de TFs eudicotiledôneas e monocotiledôneas modelo, revelando uma potencial conservação funcional do controle da biossíntese da SCW. A análise de componente principal (PCA), correlação de Pearson e de redes bayesianas confrontando dados correspondentes à expressão dos genes da biossíntese de lignina, razão S/G, conteúdo de lignina acetilbromide frente ao perfil transcricional do TFs de ambos parentais, evidenciou em destaque que os TFs homólogos: *ShMYB85* e *ShMYB58/63* possuem a potencial função como ativadores específicos da deposição de lignina nas espécies de *Saccharum*. Para verificar isto um ensaio de transativação de protoplastos nos parentais, foi desenvolvido, entre estes TFs por separado frente às regiões regulatórias dos genes representativos da biossíntese de lignina (*F5H*, *COMT* e *CAD8*) que governam a composição e conteúdo de lignina, observando a ativação dos promotores de *COMT* e *F5H* em *S. officinarum* e *S. spontaneum*. Este conjunto de informações acrescentam o nosso conhecimento sobre a regulação da deposição de lignina nas espécies de *Saccharum* por parte de *ShMYB42/85* e *ShMYB58/63*, o que os converte como fortes candidatos para modificação genética e no desenvolvimento de novos genótipos em cana-de-açúcar para o aproveitamento de bioetanol lignocelulósico.

**CAPÍTULO II. Caracterização funcional de *Shmyb85*  
e *Shmyb58/63* em *Saccharum* spp. SP-3280**

---

## **CAPÍTULO II. Caracterização funcional de *Shmyb85* e *Shmyb58/63* em *Saccharum spp.* SP-3280**

### **1 Material e Métodos**

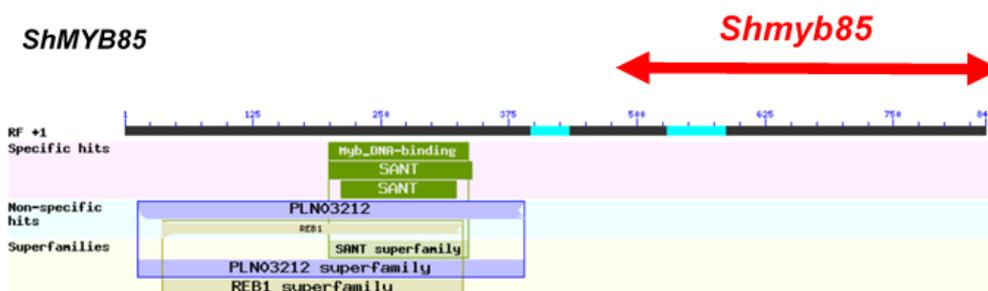
#### **1.1 Construção de pGVG\_hp\_*Shmyb85*(s+as+intron) e pGVG\_*Shmyb58/63* (s+as+intron)**

##### **Escolha de Sequência para formar RNAi no silenciamento de *ShMYB85* e *ShMYB58/63***

As mesmas sequências codificantes completas dos TFs: *ShMYB58/63* e *ShMYB85*, identificadas aqui em cana-de-açúcar e usadas para o ensaio de transativação de protoplastos no Primeiro Capítulo, foram submetidas na ferramenta online SiRNA do Galaxy Pasteur (<https://galaxy.pasteur.fr/#forms::sirna>) para procurar as regiões mais prováveis de formar *hairpin* (apresentando altos *scores*), com o intuito de silenciar a função destes genes. Se teve preferência em uma região que esteja fora dos domínios MYBR2R3, por tanto específica, indicando que para *ShMYB85* foi escolhido um fragmento de 357 pb (*Shmyb85*) e para *ShMYB58/63* foi 411 pb (*Shmyb58/63*), na direção 3', mais polimórfica (Figura 18). A sequência escolhida anteriormente mencionada para cada um dos TFs a silenciar foi submetida à ferramenta Blastn (Altschul & Madden, 1997) nos distintos bancos de cana-de-açúcar depositados no NCBI, Sugarcane Genome Hub (Garsmeur et al., 2018) assim como também no SUCEST (<http://sucest-fun.org/>; Vettore et al., 2001). Se visualizou os distintos alinhamentos, observando que a sequência escolhida para gerar o *hairpin* seja divergente o suficiente dos transcritos que não são alvos (*OFF target*), tendo o critério importante que as regiões alinhadas só podem apresentar no máximo 20 nucleotídeos contínuos idênticos com outro transcrito não alvo.

### >ShMYB85 CDS

ATGGGTCGGCAGCCGTGCTGCGACAAGCTGGGGGTGAAGCGGGGGCCGTGGACGGCGGAGGAGACCGAAGCTCATCAACTTCATCCTGACCA  
 ACGGCCATTGCTGCTGGCGCGCGGTGCCAAGCTCGCGGGCCTGCTGCGCTGCGGCAAGAGCTGCCGCTGCGCTGGACCAACTACCTCGGCC  
 GGACCTCAAGCGCGCCTCCTCAGGACGCCGAGGAGCAGGTGTCATCGACCTCCACGCAAGCTCGGCAACAGATGGTCGAAGATTGCTGCCA  
 AGCTACCGGGCAGGACTGACAACGAGATCAAGAACCACTGGAACACGCACATCAAGAAGAAGCTGATCAAGATGGGCATCGATCCAGTCACGCAG  
 AGCCCTCGATCGGAAGACAACCAGCAGCAGCCCGGCTACAACCTCGCAGTCAACCAAGTTCGACGAGGCAACCAAGCAGCAGAGCCCGCAGATC  
 GATGACGTCGTCGCGATGAGGGACATGCCGCCTGATGGTTGCAGCCCGACGGAATCGAGCACAAACACCGTGAGCACCGGCAGAAGCAGCAGCAG  
 CAGCAGCAGTGGCGGCGCCATGACCAAGACCCGCTGGTGAAGTGGCTCTTGGAAAGAGGAGCCTCCACCGGCAGCAGCCGTGGCTCAACTTCA  
 CTGGCAGTGTGATGTGGACGAGTTCAGCAGCATTGCCGCGGTCCAGAGTTGTTGCCGTGGGATGGCGCCACCGACTGGCTGCTCGACTACCAA  
 GATTTTGGATTGGGGGACTCGGCGACCTTGGTCGATGGCTACATGGTCAACAACAGCTCAAACGGAGCAAAGTTCTAG



### >ShMYB58/63 CDS

ATGGGGCGAGGGCGAGCGCGGTGCTGCGCAAGGTGGGGCTCAACAGGGGCTCCTGGACGCCGAGGAGGACATGCGCCTCATCGCTACAT  
 CCAGAAGCACGGCCACTCCAAGTGGCGCGCTGCCAAGCAAGCAGGTTTGTGCGGTGTTGGAAGAGCTGCAGGCTGCGGTGGATCAACTA  
 CCTGCGGCCGACCTCAAGCGCGCAGCTTACCGCCGAGGAGGAGACCATCATAAGGCTGCACGGTTTGTGCGGGAACAAATGGTCTAA  
 GATCGCTGCCTGCTGCCGGGAGGACGGACAACGAGATCAAGAACGTGTGGAACACGCACCTGAAGAAGAAGGTGGCGCCGCGGGAAGCA  
 GAAGGCTGGCGCGGCGACAAGAACGACGGCGGCGAGCGGCGACGCCGACGCCGCCACTGCCGCTCGTCTCGGTGTCCTCCTCGA  
 CGACGACCCACAACAGTTCCGGGGCAGCGACAACCTCCGGCGACCAGTGGCGCACGAGCAGGAGCCCGACGACGCGATCGACGTATCCCTCC  
 TCCGACTAGAGGACATCGACGTCTCGGACATGCTGGTGGACGCGCCCCGGCGGCGCAGCCATGCCAAGCGCAATGCTGTCGCGGTGCTCCT  
 CGTCTCCTGACGACGTGCGTCGGCGGGGTGGAGGAACTGATCGAGCTGCCGGTGATCGAGATCGAGCCGGAGATCTGGAGCATCATCGACG  
 GAGAGAGCGCTGACGCGCCGACGCCTCGGGAGGCGATGCAACAGCGCCATGTACGGGCACCGCCGTGCTCAGCACCAGCGAAGCAGAGGAA  
 CGCGCGAACGATTGGTGGTTGGAGAATTTGGAGAAGGAGCTGGGCTGTGGGGCCAGCGGAGGATACTCAGGCGGCCACCCTGACCTACTG  
 GACCACATAGCTGGCTTCCAGCCACTCGGCGCCTTGGAGTTGGAGAGGGACCCAGTCTCCACCTACTTCCAGACCAGCGCCCGCCTTGGGAAC  
 CTGAACTCCTGGTGGTGGACGAACCATCTGCCGTTCTGCTATGA

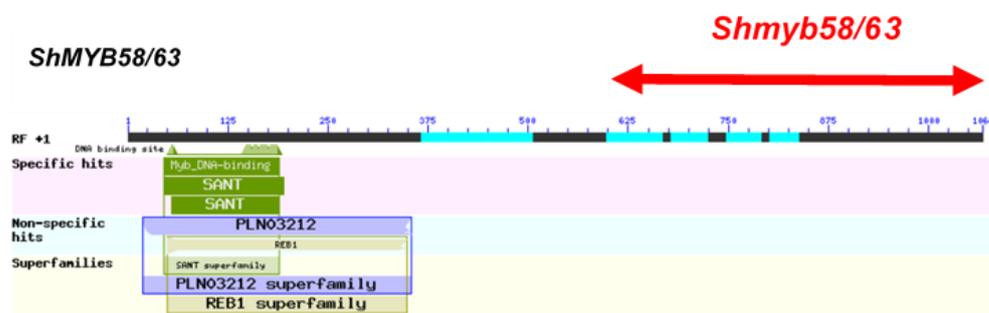


Figura 18 – Sequência codificante completa de *ShMYB85* e *ShMYB58/63* identificadas em *Saccharum* spp. cana-de-açúcar SP-3280. A região da CDS grifada em amarelo e as setas em vermelho representa as sequencias empregadas e a extensão, respectivamente para a construção do RNAi nos TFs para o seu silenciamento.



Tabela 7 - Lista de Primers desenhados para a construção do pUBI\_hp\_*Shmyb85* e pUBI\_hp\_*Shmyb58/63*

Primer	Sequência	Enzima de Restrição
<i>Shmyb85_rnai_sF</i>	<b>cgacgcgctc</b> GACGAGGCAACCAAGCAG	MluI
<i>Shmyb85_rnai_sR</i>	<b>ggggtacccc</b> AGTCCCCCAATCCAAAATCT	KpnI
<i>Shmyb85_rnai_asF</i>	<b>cgggatcc</b> AGTCCCCCAATCCAAAATCT	BamHI
<i>Shmyb85_rnai_asR</i>	<b>ggccagatatc</b> GACGAGGCAACCAAGCAG	EcoRV
<i>Shmyb58/63_rnai_sF</i>	<b>cgacgcgctc</b> CGGGGTGGAGGAACTGAT	MluI
<i>Shmyb58/63_rnai_sR</i>	<b>ggggtacccc</b> TCATAGCAGAACGGCAGATG	KpnI
<i>Shmyb58/63_rnai_asF</i>	<b>cgggatcc</b> TCATAGCAGAACGGCAGATG	BamHI
<i>Shmyb58/63_rnai_asR</i>	<b>ggccagatatc</b> CGGGGTGGAGGAACTGAT	EcoRV

O fragmento antisense (as) e sense (s) correspondente a *Shmyb85* e *Shmyb58/63* foram amplificados (Figura 21) a partir do cDNA sintetizado a partir do 5° internódio de *S. spontaneum*, purificados com o kit *PureLink Quick Gel Extraction and PCR Purification Combo kit* do gel de agarose ao 1.5%, digeridos com as enzimas de restrição correspondentes (ver tabela 7) e ligados (com T4 DNA ligase) ao vetor pUBI\_intron (também digerido preliminarmente com as enzimas de restrição correspondentes, para liberar os fragmentos asDella e sDella). O produto da ligação foi transformado em bactérias *E.coli* Dh10 $\beta$  termocompetentes (Novagen), e estas foram plaqueadas em meio LB semi-sólido contendo 100  $\mu$ g/mL de ampicilina. 25 colônias foram selecionadas para a mini-preparação plasmidial pelo método de lise alcalina. Para confirmar se as construções de *Shmyb85* (s+as+intron;837pb) e *Shmyb58/63* (s+as+intron;930pb), estarem dentro do vetor pUBI, se procedeu a fazer uma digestão com Bam HI e KpnI e o sequenciamento (para confirmar a identidade original do fragmento e descartar recombinações não desejadas). Na figura 22 pode se observar que houve sucesso na construção s+as+intron tanto para *Shmyb85* como *Shmyb58/63* e portanto na ligação desde que como se observa no gel as construções tiveram o tamanho esperado.

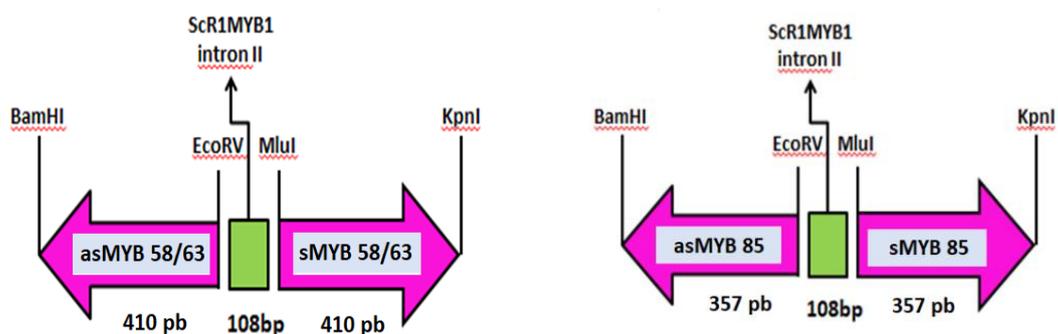


Figura 20 – Esquema do construção *Shmyb85* (s+as+intron) e *Shmyb58/63* (s+as+intron).

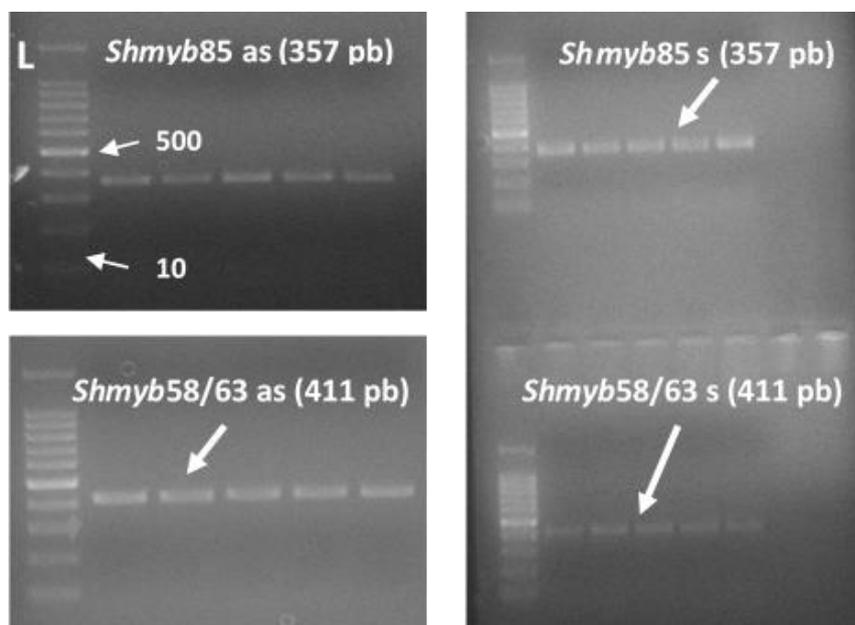


Figura 21 – Amplificação por PCR convencional de *Shmyb85* e *Shmyb58/63* tanto para o lado sense como antisense visualizados em gel de agarose ao 1.5% com brometo de etídeo. L=padrão de bandas 100pb, os valores da esquerda do padrão de bandas representam os tamanhos em pares de bases.

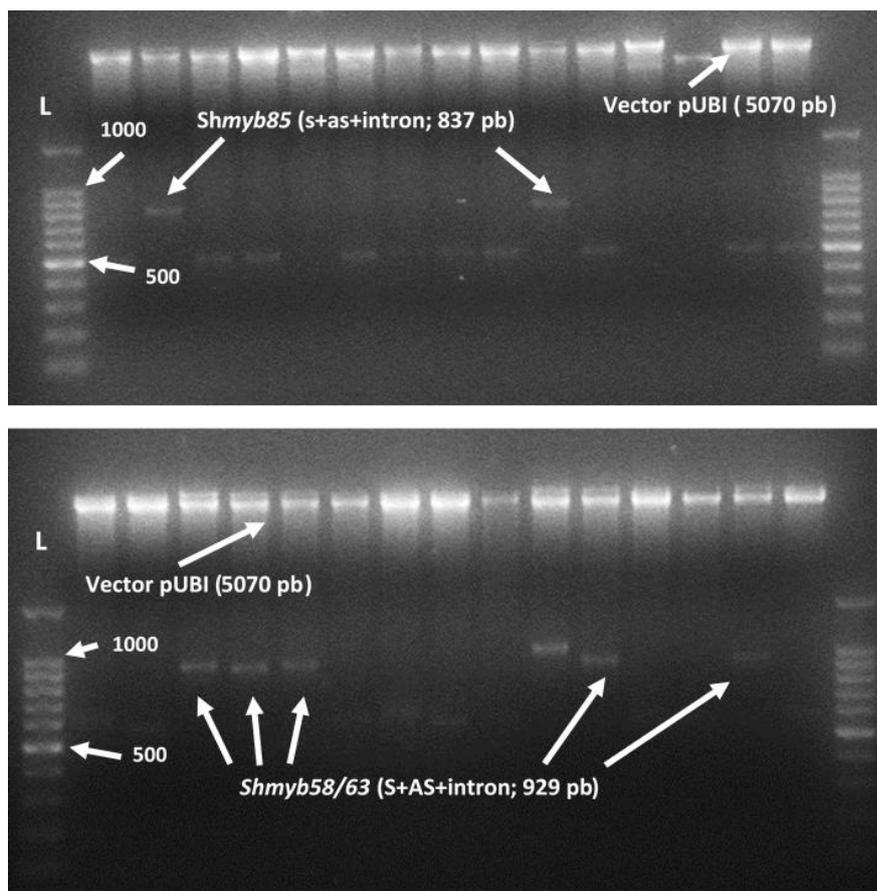


Figura 22 – Verificação da presença no plasmídeo pUBI das construções (s+as+intron) tanto para *Shmyb85* como para *Shmyb58/63* depois da reação de ligação com T4 DNA ligase. As verificações foram feitas pela digestão do pUBI com enzimas de restrição BamHI + KpnI as quais digerem dois sítios específicos que flanqueiam a construção s+as+intron e por sequenciamento Sanger. L=padrão de bandas 100pb, os valores da esquerda do padrão de bandas representam os tamanhos em pares de bases

Depois da confirmação da presença da construção *Shmyb85* (s+as+intron) e *Shmyb58/63* (s+as+intron) no pUBI, por digestão com BamHI+ KpnI, os plasmídeos positivos foram sequenciados. Uma vez sequenciados, as construções: pUBI\_hp\_*Shmyb85* (s+as+intron) e pUBI\_*Shmyb58/63* (s+as+intron), foram digeridas (Figura 23A) novamente com a mistura de enzimas BamHI+ KpnI para liberar as construções: *Shmyb85* (s+as+intron) e *Shmyb58/63* (s+as+intron) e ligar estas com T4 DNA ligase por separado no vetor modificado *pCR8/GW/TOPO* (2817 pb) do sistema gateway que também possui locais de restrição para BamHI e KpnI. O produto da ligação foi transformado em bactérias *E.coli* Dh10 $\beta$  termocompetentes (*Novagen*), e estas foram plaqueadas em meio LB semi-sólido contendo 50  $\mu$ g/mL de

Espectinomomicina. 10 colônias foram selecionadas para a mini-preparação plasmidial pelo método de lise alcalina. Para confirmar se as construções de *Shmyb85* (s+as+intron;837pb) e *Shmyb58/63* (s+as+intron;930pb), estarem dentro do vector *pCR8/GW/TOPO*, se procedeu a fazer uma digestão com BamHI e KpnI (Figuras, 23B e 23C) e o sequenciamento (para confirmar a identidade original do fragmento e descartar recombinações não desejadas). Nas Figuras 23B e 23C pode se observar que houve sucesso na ligação de s+as+intron tanto para *Shmyb85* como *Shmyb58/63* no *pCR8/GW/TOPO* (Figura 24) como se observa no gel as construções tiveram o tamanho esperado.

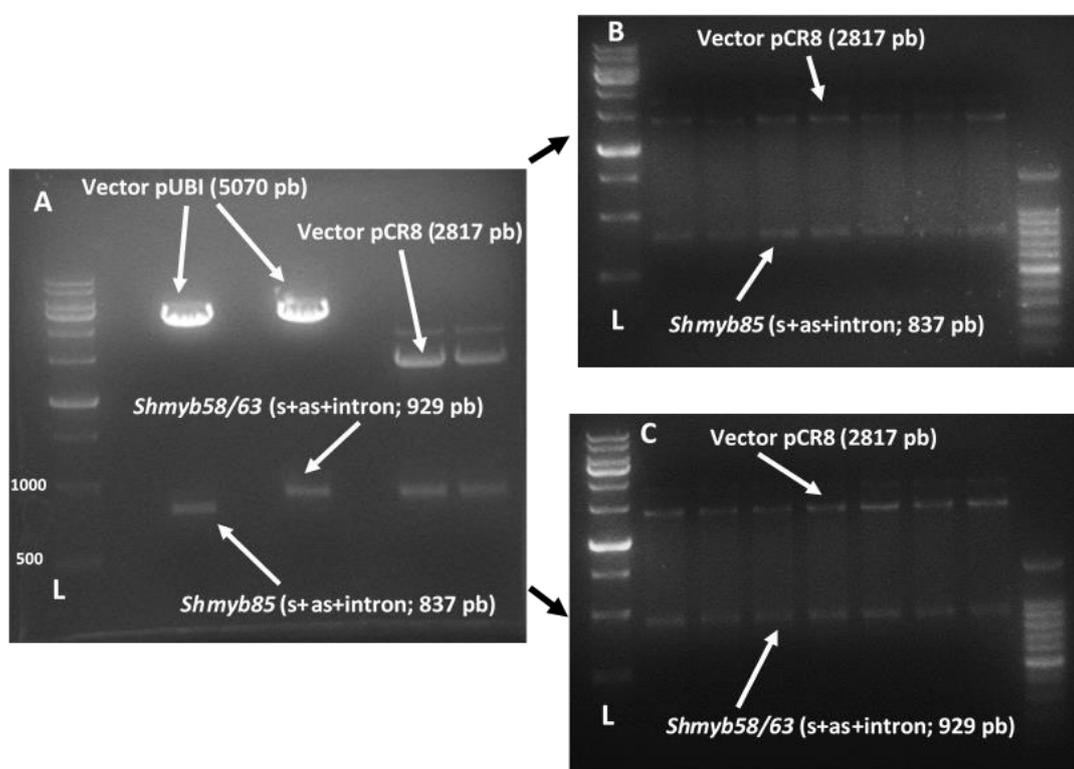


Figura 23 – A) (Lado esquerdo) Digestão com KpnI e BamHI no pUBI:Shmyb para liberar a construção *Shmyb85* (S+AS+intron) e *Shmyb58/63* (s+as+intron); (Lado direito) Digestão de pCR8/GW/TOPO com KpnI e BamHI para ligar *Shmyb85*(s+as+intron) e *Shmyb58/63*(s+as+intron) com T4 DNA ligase. B) Verificação das construções *Shmyb85*(s+as+intron) e *Shmyb58/63*(s+as+intron) no pCR8/GW/TOPO por Digestão com KpnI e BamHI. Gel de agarose 1.5 % com brometo de etídeo. L=padrão de bandas 1kb, os valores da esquerda do padrão de bandas representam os tamanhos em pares de bases

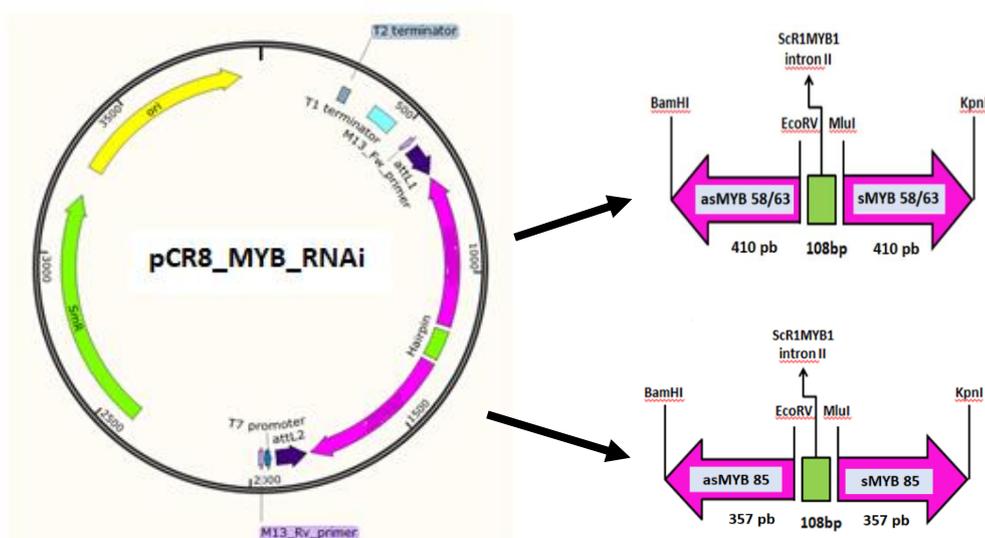


Figura 24 – Esquema do Vector pCR8\_hp\_*Shmyb85* (s+as+intron) e pCR8\_hp\_*Shmyb58/63* (s+as+intron).

Uma vez conseguido ligar as construções *Shmyb85* (s+as+intron) e *Shmyb58/63* (s+as+intron) no vetor modificado *pCR8/GW/TOPO*, foi feita uma reação de recombinação com o vetor de destino foi pGVG (12644 pb) (Guidelli et al., 2018) usando Invitrogen *LR Clonase II Plus Enzyme mix*. O produto da recombinação foi transformado em bactérias *E. coli* Dh10 $\beta$  termocompetentes (*Novagen*), e estas foram plaqueadas em meio LB semi-sólido contendo 50  $\mu$ g/mL de Kanamicina. 10 colônias foram selecionadas para a mini-preparação plasmidial pelo método de lise alcalina. Para confirmar se as construções de *Shmyb85* (s+as+intron;837pb) e *Shmyb58/63* (s+as+intron;930pb), estarem dentro do vetor pGVG, se procedeu a fazer uma digestão com BamHI e KpnI (Figura 25). A Figura 25, também amostra que teve sucesso a reação LR clonase de recombinação entre o vetor de entrada (pCR8\_hp\_*Shmyb85* (s+as+intron) e pCR8\_*Shmyb58/63* (s+as+intron)) e o vetor de destino pGVG. Da mesma forma como nos passos anteriores as construções: *Shmyb85* (s+as+intron) e *Shmyb58/63* (s+as+intron) inclusas no vetor pGVG (Figura 26) também foram sequenciados usando os primers Hairpin Forward pUBI: GCCTTTAGTTCCTTGGCAGT e *hairpin* Reverse pUBI: CACTGCCAAGGAACTAAAGG.



## **1.2 Transformação do híbrido SP-3280 de cana-de-açúcar (*Saccharum spp.*) para a obtenção de linhagens transgênicas silenciadas (*Shmyb85* e *Shmyb58/63*) em *ShMYB85* e *ShMYB58/63***

Para a produção das linhagens transgênicas, plantas de cana-de-açúcar (SP-3280) foram cultivadas em estufa (IAC, Ribeirão Preto, Brasil) por seis meses, e a região meristemática do ápice da parte aérea foi usada para gerar explantes. O material foi cultivado em meio MS (Murashige e Skoog, 1962), a 26°C no escuro, até a geração de calos embriogênicos. O pGVG\_hp\_*Shmyb85* e o pGVG\_hp\_*Shmyb58/63* foram transferidos para *Agrobacterium tumefaciens* (estirpe EHA105) por choque térmico. As culturas bacterianas foram incubadas com calos de cana-de-açúcar sob pressão de vácuo por 5 minutos e transferidas para meio de co-cultivo a 22°C, no escuro por 3 dias. Depois disso, os calos foram mantidos em meio de repouso a 26°C, no escuro por 6 dias. Após a fase de repouso, os calos transformados foram transferidos para um meio de regeneração seletiva a 26°C, durante 14 dias com 16 h de fotoperíodo. Os eventos transgênicos foram mantidos em meio sem induzir crescimento e enraizamento. A seguir, as linhagens transgênicas foram transplantadas em tubetes com substrato (0.5 L), sendo o conjunto destes colocado em uma bandeja contendo água e cobertos com filme de PVC, onde foram aclimatados durante 4 dias (dentro de uma câmara com alto teor de umidade). Depois, os tubetes foram transferidos para a estufa para adaptação à temperatura ambiente (2 dias ainda cobertos com filme de PVC e depois descobertos e mantidos assim durante 30 dias, até obter boa brotação e considerável densidade radicular (Figura 27). As melhores brotações de cada linhagem transgênica de *Shmyb85* e *Shmyb58/63* e controles (Vetor Vazio (EV) e Controle +(WT)) (saudáveis e com boa densidade radicular) foram separados em plântulas e transplantados para vasos de 10L, crescendo nas seguintes condições durante 3 meses: Irrigação por aspersão diariamente, fertilização com 50 ml de solução Hoagland de força  $\frac{1}{4}$  a  $\frac{1}{2}$  sucessivamente por semana. Só as linhagens transgênicas possuindo 3 replicatas biológicas (como mínimo) e saudáveis foram consideradas para o estudo.

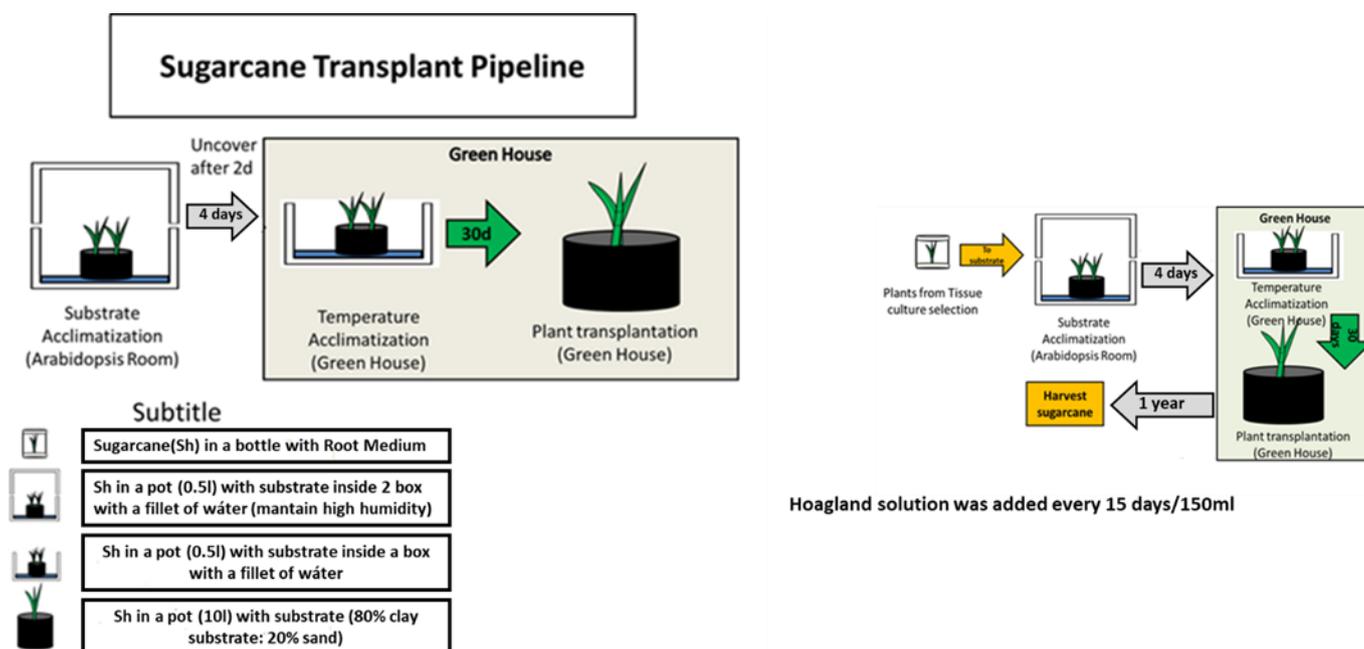


Figura 27 – Pipeline apresentando os passos depois da cultura de tecido e transplante de plântulas de *Saccharum* spp. cultivar híbrido SP-3280 para vasos com substrato.

### 1.3 Material vegetal e condições de crescimento

13 linhagens transgênicas primárias silenciadas para *Shmyb85* e 17 para *Shmyb58/63* de 1 ano de idade foram usadas neste estudo, obtidas a partir de calos embriogênicos do cultivar híbrido de *Saccharum* spp. 3280-SP transformados com as construções: pGVG\_hp\_ *Shmyb85* e pGVG\_hp\_ *Shmyb58/63* respectivamente. As plantas transformadas com vector vazio (EV) pGVG (Guidelli et al., 2018) e plantas de tipo selvagem (WT) foram consideradas como controles. As linhagens transgênicas assim como os controles cresceram em vasos de plástico com capacidade de 10 L contendo uma mistura de um Solo A, Solo B e areia (1:1:2 v/v). As plantas ficaram em uma estufa sem controle de humidade e temperatura, sob condições de luz natural (Figura 29). As plantas foram irrigadas manualmente com 500 ml de H<sub>2</sub>O a cada dois dias e fertilizadas a cada 15 dias com 150 ml de uma solução Hoagland (força total). A coleta do material foi feita em Agosto do 2020, sendo entre 3 e 5 replicatas biológicas de acordo com o tratamento. Colmos das repetições dos diferentes tratamentos, aparentemente saudáveis, do mesmo estágio fisiológico e de tamanho uniforme foram coletados usando um podão, separando os internódios 5°+6° e 7°, não incluindo o nó

e as folhas +4. Os testes histoquímicos foram executadas no internódio 7°. As análises bioquímicas e de expressão gênica foram feitos nos internódios 5°+6°, das 5 linhagens transgênicas de *Shmyb85* e *Shmyb58/63* que apresentaram a maior queda de lignina por efeito do silenciamento de *ShMYB85* e *ShMYB58/63* a partir da avaliação da lignina acetyl bromide na totalidade das linhagens transgênicas mencionadas no começo deste item. Após serem lavados com água da torneira, os internódios 5°+6° foram colocados em saco de polietileno e, em seguida, em recipientes térmicos contendo nitrogênio líquido, para posterior liofilização. O material liofilizado foi macerado em cadinhos esterilizados por autoclave e resfriado com nitrogênio líquido, transferidos para tubos Falcon, e armazenados em freezer a -80°C, até o seu posterior uso nas distintas análises moleculares e bioquímicas.

#### **1.4 Injúria (Indução da lignificação) e amostragem:**

Para o experimento da injúria, as melhores plantas (saudáveis) das linhagens transgênicas e controles mencionados no item anterior, com 10 meses de idade, do mesmo tamanho (2 m aproximadamente) e com boa produção de folhas (6 até 10 pares de folhas aproximadamente, sem clorose) foram usadas neste experimento. Aqui, nós escolhemos as folhas, usando o sistema de numeração de folhas proposto por (Moore, 1987), sendo que a primeira folha completamente expandida, com visível aurícula e fotossinteticamente ativa foi considerada como +1. Para estimular a lignificação, a folha +3 (de aproximadamente 50 cm de comprimento) foi mecanicamente injuriada usando uma pinça de dissecação anatômica com serrilha, na região média (30 cm de comprimento) a cada 2.5 cm, sem comprometer a nervura central. A folha +4 (de aproximadamente as mesmas dimensões que a folha +3) não foi machucada e serviu como controle (Figura 28). Depois de 5 dias, a região meia das folhas +3 e +4 foi separada (sem incluir as regiões basais e apicais) e coletada em tubos Falcon de 50 ml imediatamente imersa em nitrogênio líquido e armazenadas em freezer a -80°C até o momento das análises bioquímicas e moleculares.

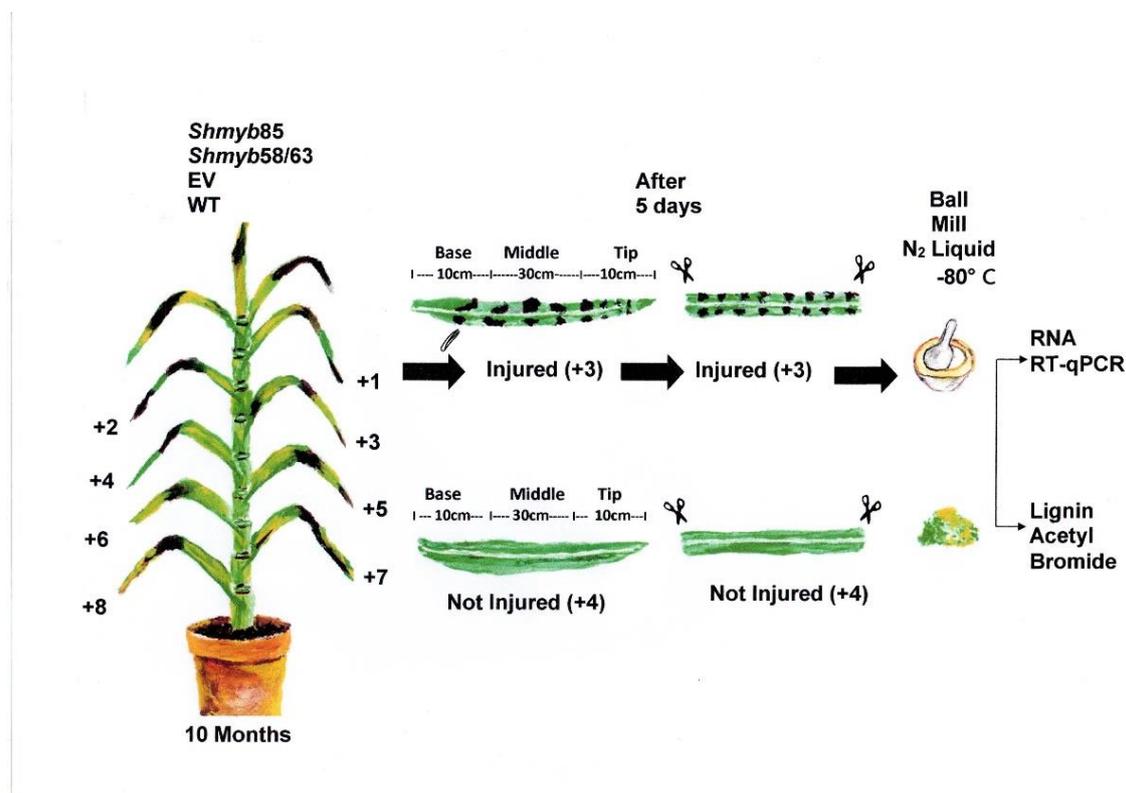


Figura 28 – Esquema geral do sistema de injúria estabelecido em folhas de linhagens transgênicas e controles de *Saccharum* spp. cultivar híbrido SP-3280: *Shmyb85*, *Shmyb58/63*. WT e EV são os controles. O sistema de numeração das folhas proposto, foi de acordo com Moore (1987). As Folhas (+3) injuriadas depois de 5 dias foram seccionadas na sua região média do limbo. As folhas (+4) serviram como controles. As folhas injuriadas e não injuriadas foram processadas para a avaliação de expressão génica e deposição de lignina Acetyl Bromide.

### 1.5 Caracterização funcional de linhagens transgênicas silenciadas (*Shmyb85* e *Shmyb58/63*) para *ShMYB85* e *ShMYB58/63*

#### Análises histoquímicas

Para os testes histoquímicos, os internódios 7° dos diferentes tratamentos foram utilizados nestas análises. As amostras, fazendo uso de uma lâmina de aço, foram seccionados longitudinalmente em blocos de aproximadamente de 2.5cm x 1cm de altura e largura respectivamente e imediatamente colocados em tubos Falcon de 15 ml de capacidade contendo uma solução fixadora FAA (formaldeído, ácido acético e etanol 50%, 5:5:90), levados a uma bomba de vácuo para a retirada do ar contido nos

tecidos e mantido assim por 5 dias. Posteriormente a solução fixadora foi retirada e trocada com etanol 50%, mantendo os blocos assim em vácuo por 24 horas. Terminado este período, a solução de 50% foi trocada por etanol 70% em vácuo até o começo dos testes histoquímicos. Os testes foram realizados em secções feitas a mão livre, usando uma lâmina de aço com o intuito de evidenciar a presença (Teste de Weisner/fluoroglucinol-HCL (Johansen, 1940)) e tipo de lignina (Teste de Maüle (Mäule, 1901)). Para a reação de Mäule, as amostras foram submetidas a  $\text{KMnO}_4$  0,5% por 10 minutos, seguido por lavagens com água destilada, mantido em HCl 10% por 2 min e a lâmina foi montada com  $\text{NH}_4\text{OH}$  concentrado. Para o floroglucinol as secções foram tratadas por 5 min com floroglucinol 1% em etanol 95% e a lâmina foi montada com HCl 25%. Para a documentação dos resultados utilizou-se câmera de vídeo Olympus DP71 acoplada ao microscópio Olympus BX 51.

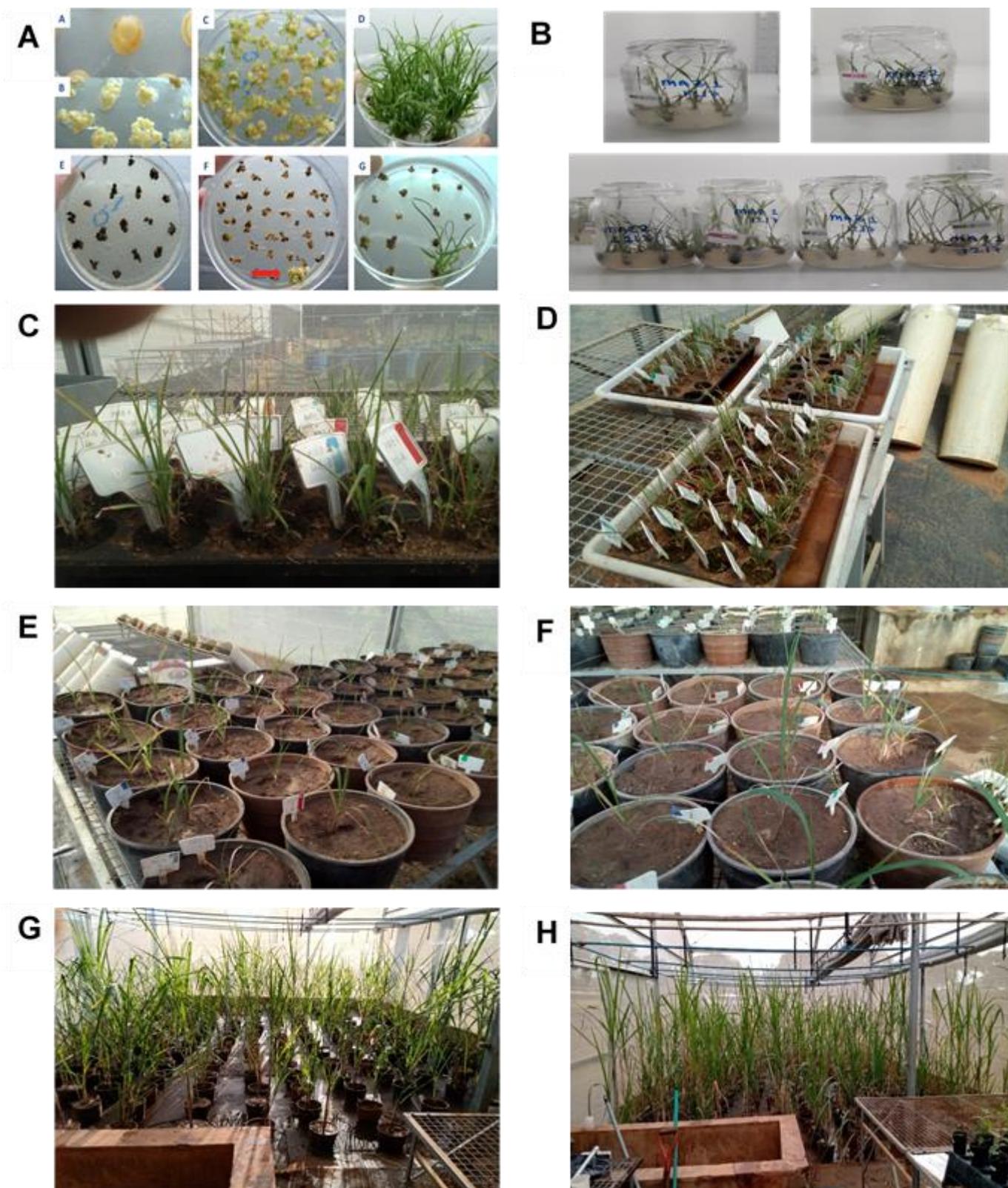


Figura 29 – A e B) Obtenção de calos embriogênicos e linhagens transgênicas de *Saccharum* spp. 3280-SP silenciadas para *ShMYB85* e *ShMYB58/63*, em meio MS. C e D) Linhagens transgênicas e controles transplantados em tubetes e transferidas a estufas para aclimação. E e F) Linhagens transgênicas e controles dos tubetes transplantados a vasos de 10L baixo condições de estufa. G e H) Linhagens transgênicas e controles crescendo em vasos de 10L com 6 e 12 meses de idade respectivamente.

## **Análises bioquímicas**

### **Determinação de Lignina Acetyl Bromide (AcBr)**

A determinação do conteúdo de Lignina Acetyl Bromide foi realizada de acordo com a metodologia descrita por (Hatfield & Fukushima, 2005) e conforme já descrito no primeiro capítulo.

### **Determinação da razão S/G**

A determinação da razão S/G foi feita de acordo com Kiyota et al. (2012) e conforme anteriormente descrito no primeiro capítulo.

### **Determinação dos polissacarídeos da parede celular**

As determinações de celulose, hemicelulose e pectina foram realizadas de acordo com a metodologia descrita por (Chen et al., 2002) com algumas modificações: 150 mg de amostra foi homogeneizada com 5 ml H<sub>2</sub>O destilada fria, usando um agitador e, em seguida, foi centrifugada durante 10 min, removendo-se o sobrenadante e conservando-se o pellet. O pellet foi lavado sequencialmente com 5 ml de H<sub>2</sub>O destilada fria, 5 ml de acetona e 5 ml de metanol/clorofórmio (1:1, v/v), sendo que em cada passo foi feita a centrifugação para recuperação do material sólido, que foi então seco em bancada por 24 h antes de ser extraído com o solvente seguinte. Para remover o amido, o resíduo obtido foi tratado a 37°C durante 3 h com 5 ml de 2 unidades.ml<sup>-1</sup> de amilase pancreática (tipo I-A, Sigma Chemical Co, St Louis, MO) em tampão 0,1 M acetato de sódio (pH: 6,5), e em seguida centrifugado para recuperar o material sólido. Este foi então extraído três vezes com 3 ml de 20 mM de oxalato de amônio (pH 4,0) a 70°C durante 1 h para se obter a fração correspondente às pectinas. Após centrifugação, o pellet foi então extraído com 3 ml de NaOH 0,1 M à temperatura ambiente sob vácuo durante 24 h no escuro, e o material foi recuperado por centrifugação. O sobrenadante foi denominado fração 1 de hemicelulose (HEM1). O pellet foi extraído três vezes com 2 ml de NaOH a 17,5% por 8 h, sob a vácuo e os sobrenadantes obtidos de cada extração foram reunidos e chamados fração 2 de hemicelulose (HEM2). A fração total de hemicelulose foi obtida pela soma das extrações (HEM1 + HEM 2). O resíduo alcalino insolúvel foi lavado sequencialmente com 5 ml de água destilada, 5 ml de 1 mM de ácido acético e 5 ml de etanol e em seguida seco a 37°C. Em seguida, esta fração foi dissolvida em 5 ml de uma solução

72% de H<sub>2</sub>SO<sub>4</sub> durante 1 h à temperatura ambiente e diluída 30 vezes com água destilada para obter a fração de celulose correspondente. O teor total de açúcares totais em cada fração foi obtido de acordo com o método do fenol sulfúrico, utilizando glucose como padrão (DuBois et al., 1956).

### **Determinação da porcentagem de sacarificação**

A determinação da porcentagem de sacarificação foi conduzida de acordo com a metodologia descrita por (Brown & Torget, 1996), com pequenas modificações. Em microtubos, foi pesada uma quantidade de biomassa liofilizada equivalente a 10 mg de celulose (previamente determinada, conforme protocolo acima e livre de açúcares solúveis e amido), e a esta foi adicionado 500 µl de tampão citrato de sódio (0,1 M, pH 4,8), 10 µl de azida sódica e uma quantidade necessária de água para completar 1 ml. Posteriormente, a mistura foi aquecida a 50°C e em seguida acrescentou-se 6,08 µl de uma mistura (1:4, v/v) de enzimas celulase (1,2 FPU/10 mg de celulose) e celobiose (1,26 U pNPGU/10mg celulose). Este coquetel foi preparado a partir de celobiose de *Aspergillus niger* e celulase de *Trichoderma reesei* (Sigma-Aldrich). Os microtubos contendo as amostras e a solução de reação, foram fechados hermeticamente e colocados inclinados (45°) em um agitador orbital, a 50°C, sendo mantidos a 160 rpm durante 5 dias. Passado esse período as amostras foram centrifugadas a 12000 rpm por 15 min. A determinação de glicose foi feita no sobrenadante pelo método do fenol sulfúrico (DuBois et al., 1956) mediante a leitura no espectrofotômetro a 495 nm, usando uma curva padrão de glicose no intervalo 20-100 µg/ml.

### **Determinação de compostos fenólicos solúveis totais**

O conteúdo total de compostos fenólicos solúveis foi determinado com o reagente de Folin-Ciocalteu (Swain & Hillis, 1959), utilizando-se ácido clorogênico como padrão. Amostras de 150 mg foram colocadas em microtubos de 1,5 mL e extraídas com 700 µL de etanol 80% sob sonicação por 30 min, sendo depois centrifugadas durante 10 min a 12000 rpm. Coletou-se o sobrenadante (por volta de 600 µL) e se repetiu a extração anterior. Os extratos de cada amostra foram reunidos e 50 µL foram adequadamente diluídos com 3,450 ml de água destilada. Em seguida, foram adicionados 0.25 ml do reagente de Folin-Ciocalteu, misturando a solução e incubando-a por 3 min em temperatura ambiente. Logo depois, 0.5 ml de Na<sub>2</sub>CO<sub>3</sub>

saturado (17,5g/50 ml) foram misturados à solução, agitados e mantidos em repouso por 1 h em temperatura ambiente, quando então se adicionou 0.75 ml de água destilada. A determinação do conteúdo de fenóis totais foi feita mediante a leitura da absorbância em 725 nm em um espectrofotômetro UV-VIS. Se utilizou uma curva padrão de ácido clorogênico calculada para o intervalo de concentração de 100-1000 ug/ml.

### **Análise por espectroscopia 2D-NMR**

As amostras correspondentes as linhagens transgênicas de *Shmyb85* (16,15 e 14), *Shmyb58/63* (17,16 e 12), assim como também os controles (EV e WT) foram sucessivamente tratadas com acetona (8h), metanol (8h) e água (8h) para remover materiais extrativos. Assim as matérias livres de extrativos foram finamente moídas em um Restch PM-100 planetary mill (Retsch, Haan, Germany) equipado com uma jarra de ágata e bolas de ágata (10X10 mm) a 600 rpm, durante 2 h (alternando 10 min de pausa a cada 20 min de moagem), previamente as análises por NMR. As paredes celulares inteiras das amostras foram analisadas por 2D-NMR em “estado de gel” de acordo com o método descrito por (Kim et al. 2008; Rencoret et al. 2009). Resumidamente, aproximadamente 55 mg de amostras livres de extrativos foram colocadas em 0.6 mL de DMSO- $d_6$ . Os espectros de 2D HSQC NMR foram registrados a 25°C em um equipo Bruker AVANCE III 500 MHz equipado com uma sonda de gradiente TCI de 5 mm, resfriada criogenicamente com geometria inversa nas instalações de RMN dos serviços de pesquisa geral da Universidade de Sevilla (SGI-CITIUS). Os experimentos de HSQC foram realizados usando programas de pulso padrão Bruker "hsqcetgpsisp.2" e os seguintes parâmetros: os espectros foram adquiridos de 10 a 0 ppm em F2 ( $^1\text{H}$ ) usando 1000 pontos de dados para um tempo de aquisição de 100 ms, um atraso interscan de 1s, e de 200 a 0 ppm em F1 ( $^{13}\text{C}$ ) usando 256 incrementos de 32 de scan, para um tempo total de experimento de 2h e 34 min. O  $^1J_{\text{CH}}$  usado foi de 145 Hz. O processamento usou apodização gaussiana típica em  $^1\text{H}$  e um sino cosseno quadrado em  $^{13}\text{C}$ . O pico central do solvente foi usado como uma referência interna ( $\delta_{\text{C}}/\delta_{\text{H}}$  39.5/2.49). Os sinais cruzados 2D-NMR foram atribuídos por comparação da literatura (del Río et al., 2012, 2015). Uma análise semiquantitativa dos integrais de volume dos picos de correlação HSQC foi realizada usando o software de processamento Topspin 3.5 da Bruker. Na região aromática/insaturada, os sinais de correlação de  $\text{H}_{2,6}$ ,  $\text{G}_2$  e  $\text{S}_{2,6}$  foram usados para

estimar o conteúdo das respectivos monômeros de lignina: H, G e S. Os sinais para  $pCA_{2,6}$  e  $FA_2$ , foram usados para estimar a abundância dos diferentes  $p$ -hidroxicinamatos ( $p$ -coumaratos;  $pCA$ , e ferulatos,  $FA$ ) e o sinal para  $T_8$  foi usado para estimar o conteúdo da flavona tricina (como os sinais  $H_{2,6}$ ,  $S_{2,6}$  e  $pCA_{2,6}$  envolvem dois pares próton-carbono, suas integrais de volume foram reduzidas à metade). Os dados foram referentes ao conteúdo total de carboidratos (estimado a partir dos sinais dos carbonos anoméricos que correspondem em sua maioria a xilose e glicose) e lignina.

### **Análise de Expressão por RT-qPCR**

A expressão de *ShMYB85* e *ShMYB58/63* e genes da rota biossintética da lignina foi analisada por RT-qPCR nas folhas injuriadas, não injuriadas e internódios 5°+6° das linhagens transgênicas (*Shmyb85* e *Shmyb58/63*) e controles respectivamente, usando primers específicos (Tabela 8). Desde que a folha é um tecido menos lignificado do que outros órgãos como os internódios foi feito um teste preliminar (folhas injuriadas e não injuriadas) de expressão só nas plantas controle WT para verificar se a folha possui capacidade de experimentar indução da expressão dos TFs em estudo, frente à injúria mecânica provocada neste estudo. De outro lado também nos controles foi feito um teste de expressão de 5 enzimas representativas da via biossintética da lignina identificadas para os parentais de *Saccharum* e cultivares contrastantes de cana-de-acucar em lignina (Llerena et al 2019; Bottcher et al.2013) (foi escolhida para o estudo final a enzima que experimentou a maior indução de expressão) com o intuito de correlacionar com a expressão dos TFs MYBs em estudo. A curva de eficiência dos primers foi determinada pelo software Step One Plus Software v2.3 (Life Technologies), Foram usados os pares de primers que amplificaram o cDNA com uma eficiência que oscilou entre 90-110%. Os cDNAs diluídos 7.5X, das folhas e internódios 5° +6° foram utilizados na análise. Utilizamos 5 replicatas biológicas e 3 técnicas (para o teste preliminar na folha do WT) e 1 replicata biológica e 3 técnicas, para avaliar a indução da expressão dos genes frente à injúria nas linhagens transgênicas de *Shmyb85* e *Shmyb58/63*. Para a expressão nos internódios 5°+6° usamos 3 replicatas biológicas e 3 técnicas. As reações foram preparadas com o *iTaq universal SYBR Green supermix BIORAD* e analisadas no StepOnePlus™ Real-Time PCR System, seguindo o programa de 95°C por 3 min e 40 ciclos de: 95°C por 10 s e 60°C por 30 s. A especificidade dos produtos amplificados foi avaliada pela análise da curva de dissociação gerada pelo equipamento. Para a

normalização interna adequada, foi utilizado os genes constitutivos Ubiquitin e GADPH (Bottcher et al.2013). A expressão relativa para folhas injuriadas, não injuriadas e internódios 5°+6° das linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles, foi calculada por  $2^{-\Delta Ct}$  e para o coeficiente da indução de expressão foi por  $2^{-\Delta\Delta Ct}$  (Livak and Schmittgen, 2001). A indução da expressão foi determinada pelo coeficiente entre a expressão em folhas injuriadas e não injuriadas.

### **Análises estatísticas**

Para as análises bioquímicas e perfil de expressão (RT-qPCR) avaliadas no 5to-6to internódio das linhagens transgênicas *Shmyb85*, *Shmyb58/63* e controles (WT e EV) foi realizada uma análise de variância (ANOVA) simples, e quando houve diferenças entre as médias a comparação foi feita pelo teste de Wilcoxon ( $p < 0,05$ ) usando o software R (Team, R 2013). Para as análises bioquímicas do 5°+6° internódio foram avaliadas entre 3 e 5 replicatas biológicas e três triplicatas técnicas. A composição de lignina, perfil de compostos fenólicos e espectrofotometria de 2D-NMR, a determinação foi feita entre 3-5 replicatas biológicas e 1 técnica. No que respeita a expressão (RT-qPCR) dos TFs: *ShMYB85* e *ShMYB58/63* e genes da biossíntese da biossíntese de lignina em espécies de *Saccharum* a avaliação foi feita com 3 replicatas biológicas e três replicatas técnicas. Os resultados das análises bioquímicas foram expressos pela média  $\pm$  erro padrão. No experimento de indução da lignificação nas linhagens transgênicas de *Shmyb85*, *Shmyb58/63*, WT e EV, a avaliação das análises bioquímicas e de expressão gênica feitas nas folhas +3 e +4 envolveu uma replicata biológica e uma técnica. O teste preliminar de expressão nas plantas controle (WT) no experimento da injúria envolveu 5 replicatas biológicas e 3 técnicas. Para expressão gênica os resultados foram expressos pela média  $\pm$  desvio padrão. Para o controle da transferência do erro no cálculo de expressão gênica foi utilizado um modelo linear de acúmulo de erro ( $\sigma_{\Delta Ct}^2 = \sigma_{Ct,ref}^2 + \sigma_{Ct}^2$ ) no cálculo do valor de  $\Delta Ct$  e não linear ( $\sigma_{2^{-\Delta Ct}}^2 = \left(\frac{d[2^{-\Delta Ct}]}{d[\Delta Ct]}\right)^2 \sigma_{\Delta Ct}^2$ ) no cálculo do valor de  $2^{-\Delta Ct}$  (Brown & Mac Berthouex, 2002).

## 2. Resultados

### 2.1 Caracterização funcional de linhagens transgênicas silenciadas (*Shmyb85* e *Shmyb58/63*) para *ShMYB85* e *ShMYB58/63*

#### Determinação do Conteúdo de Lignina Acetil bromide e Razão S/G

Desde que a frequência de inserção do transgene (vetor contendo o RNAi para silenciamento de *ShMYB85* e *ShMYB58/63*) é aleatória no genoma complexo de *Saccharum* spp. cultivar híbrido SP-3280 nós desenvolvemos preliminarmente um *screening* de todas as linhagens transgênicas geradas (*Shmyb85* e *Shmyb58/63*) avaliando o conteúdo de Lignina acetilbromide com o intuito de escolher aquelas que tem a maior queda deste parâmetro devido ao silenciamento para análises posteriores. Adicionalmente foi avaliado em todas as linhagens de *Shmyb85* e *Shmyb58/63* a razão S/G. No conteúdo de lignina podemos indicar que para *Shmyb85*, do total das linhagens transgênicas avaliadas, 7 delas (14, 16, 15, 17, 8, 19, 18) apresentaram uma queda significativa em relação ao WT, sendo de: 18.7%, 16.0%, 12.4%, 10.2%, 8.81%, 8.71% e 6.88% (Figuras 30 A e C) com  $p \leq 0.05$  (0.036; 0.036; 0.032; 0.016; 0.016; 0.016 e 0.032, respectivamente. Em *Shmyb58/63*, observamos nas linhagens (17,11, 12,1, 15, 18 e 20) uma redução porcentual significativa de: 15.3, 11.5, 10.6, 9.21, 8.97, 8.05, 7.40, respectivamente em relação ao WT (Figuras 30 B e D) e  $p \leq 0.05$  (0.036; 0.016; 0.032; 0.036; 0.036; 0.032, respectivamente. É importante destacar que a linha transgênica 16 de *Shmyb58/63* apresentou uma considerável redução (13.9%) da deposição de lignina mas esta não foi significativa ( $p=0.095$ ).

Na determinação da razão S/G, se observa redução deste parâmetro para algumas das linhagens transgênicas de *Shmyb85* mas para nenhuma destas houve alteração significativa (Figura 30E). Para *Shmyb58/63* (Figura 30F) se observou aumento significativo ( $p \leq 0.05$ ) desta razão na maioria das linhagens transgênicas, a exceção de: 18,20,16 e 1 ( $p \geq 0.05$ : 0.31; 0.10; 0.06 e 0.11), respectivamente. O aumento oscilou entre 29.18% e 69.68%. Daqui para frente nos selecionamos para as análises posteriores de: expressão gênica, composição de parede celular, compostos fenólicos e histoquímica as linhagens transgênicas 16,14,17,15 e 8 para *Shmyb85*, e 1, 17, 11, 16 e 12 para *Shmyb58/63*. Estas foram as plantas que experimentaram a maior queda de lignina acetyl bromide. Para as análises por espectroscopia 2D-NMR foram

consideradas para *Shmyb85* as linhagens 16, 15 e 14 e para *Shmyb58/63* foram as linhagens: 17, 16, 12.

### **Análises histoquímicas**

De maneira geral, os feixes vasculares da região periférica (córtex) adjacente à epiderme do internódio 7° apresentam uma menor densidade de camada de fibras nas linhagens transgênicas de *Shmyb85* do que *Shmyb58/63*. Os controles (WT e EV) apresentaram a maior concentração em relação a estes (Figuras 31, 32 e 33, colunas da esquerda). No que respeita à região central (medula) do internódio 7°, os distintos tratamentos possuem feixes vasculares dispersos entre as células do parênquima fundamental, sendo isto mais evidente para *Shmyb85*. O reagente floroglucinol-HCl evidenciou a presença de lignina total com a coloração vermelha (Figuras 31-32 A-J, respectivamente (*Shmyb85* e *Shmyb58/63*) e Figuras 33 A-D, WT e EV). Em destaque, as fibras ao redor dos feixes vasculares evidenciam incipiente deposição de parede celular secundária em linhagens transgênicas de *Shmyb85* (Figuras 31 A, E, G e I), quando comparando com linhagens transgênicas de *Shmyb58/53* (Figuras 32 A, E, G, I). O WT e o EV (Figuras 33 A e C) apresentam uma maior deposição nas fibras em relação a *Shmyb85* e *Shmyb58/63*. No sétimo internódio, para todos os tratamentos se encontram as células parenquimáticas completamente expandidas (Figuras 31, 32 e 33). No entanto as células parenquimáticas da região central de *Shmyb85* (Figuras 31 A, E, G e I) e de *Shmyb58/63* (Figuras 32 B, D, F e J) não estão lignificadas. Em sentido contrário para o WT e o EV (Figuras 33 B e D) o parênquima fundamental desta região está lignificado. O reagente de Maüle evidenciou a coloração avermelhada para a lignina siringil (S) e amarela para a lignina guaiacil (G) (Figuras 31 e 32 K-T, respectivamente, *Shmyb85* e *Shmyb58/63* e Figuras 33 E-H para WT e EV). No internódio 7° na região periférica, os feixes vasculares, as fibras ao redor destes e as células parenquimática nas linhagens transgênicas de *Shmyb85* (Figuras 31 K-S) apresentam um predomínio de coloração amarelada, indicando lignina G. De outro lado *Shmyb58/63* apresenta um padrão particular indicando que os feixes vasculares da região periférica apresentam fibras com deposição de lignina diferenciada, desde que as camadas de fibras mas internas dos feixes apresentam coloração amarela Lignina G, enquanto as mais externas coloração avermelhada lignina S, indicando também que as células parenquimáticas desta região tem coloração vermelha lignina S (Figuras 32 S K, M,O e Q). Os controles, em especial o

WT apresentou similar configuração de deposição de lignina diferenciada observada nas linhagens de *Shmyb58/63* com a diferença que houve mais fibras que envolvem os feixes periféricos. As células parenquimáticas foram de cor vermelho igual a *Shmyb58/63* (Figuras 33E e 33G). Os feixes vasculares da região central dos distintos tratamentos apresentaram coloração amarela, lignina G, não se observou coloração nem vermelha nem amarela para as células parenquimáticas (Figuras 31, 32 e 33, colunas da direita). As diferenças marcantes encontradas entre as linhagens (*Shmyb85* e *Shmyb58/63*) e os controles (EV e WT) foram a espessura da parede celular das fibras dos feixes vasculares na região periférica e a lignificação das células parenquimáticas na região central. Em *Shmyb85* e *Shmyb58/63* os feixes vasculares próximos a epiderme apresentam fibras com parede celular mais delgada quando comparadas com os presentes nos controles (EV e WT). Na região periférica, as células parenquimáticas de todas os tratamentos estão lignificadas no internódio 7. Porém, na região central de *Shmyb85* e *Shmyb58/63* as células parenquimáticas permanecem não lignificadas comparadas aos controles (EV e WT).

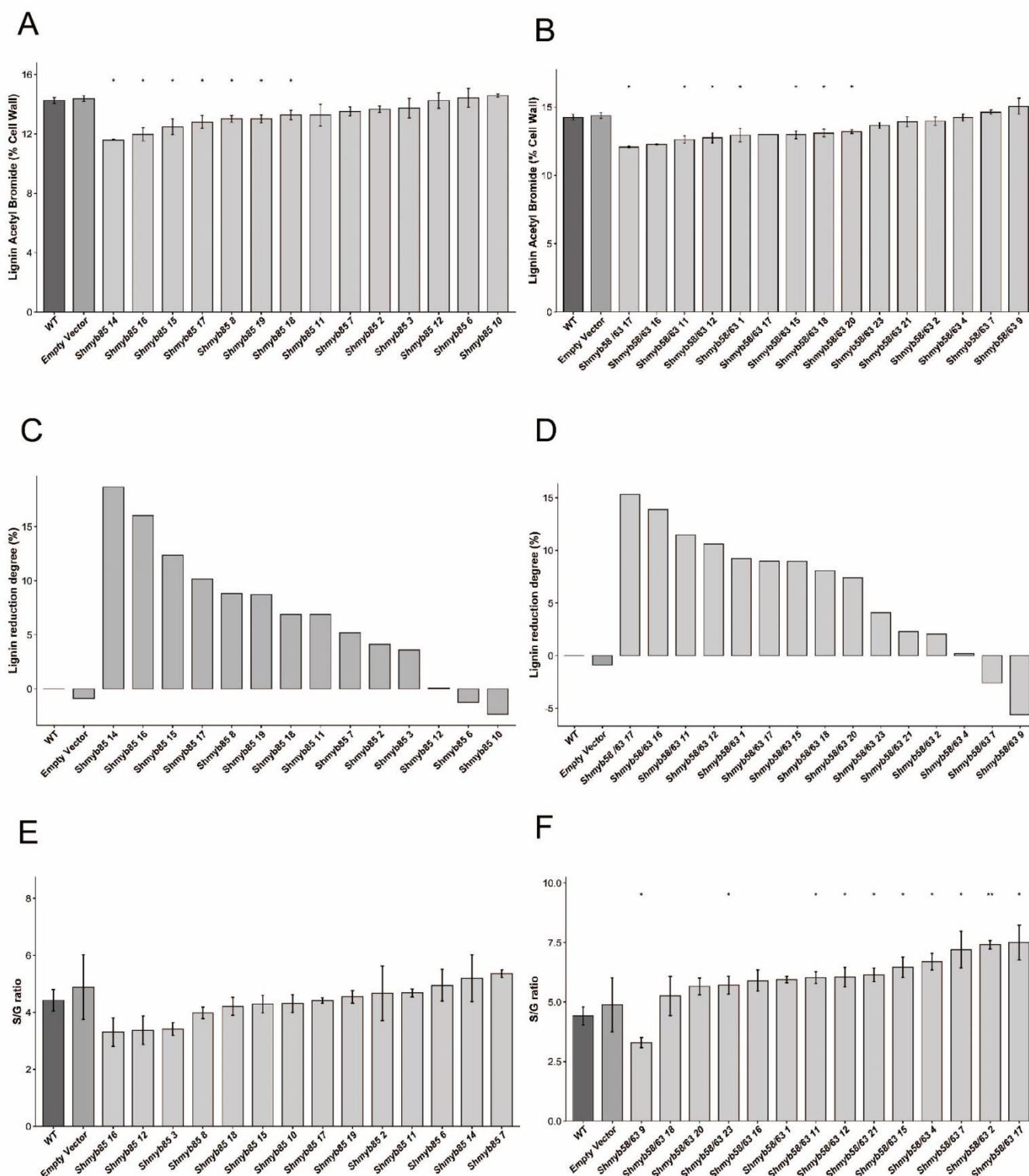


Figura 30 – A e B) Conteúdo, C e D) grau de redução de lignina Acetyl bromide e E e F) razão S/G no 5º internódio em linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo de lignina e razão S/G entre os tratamentos. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das médias produto de 3-5 replicatas biológicas.

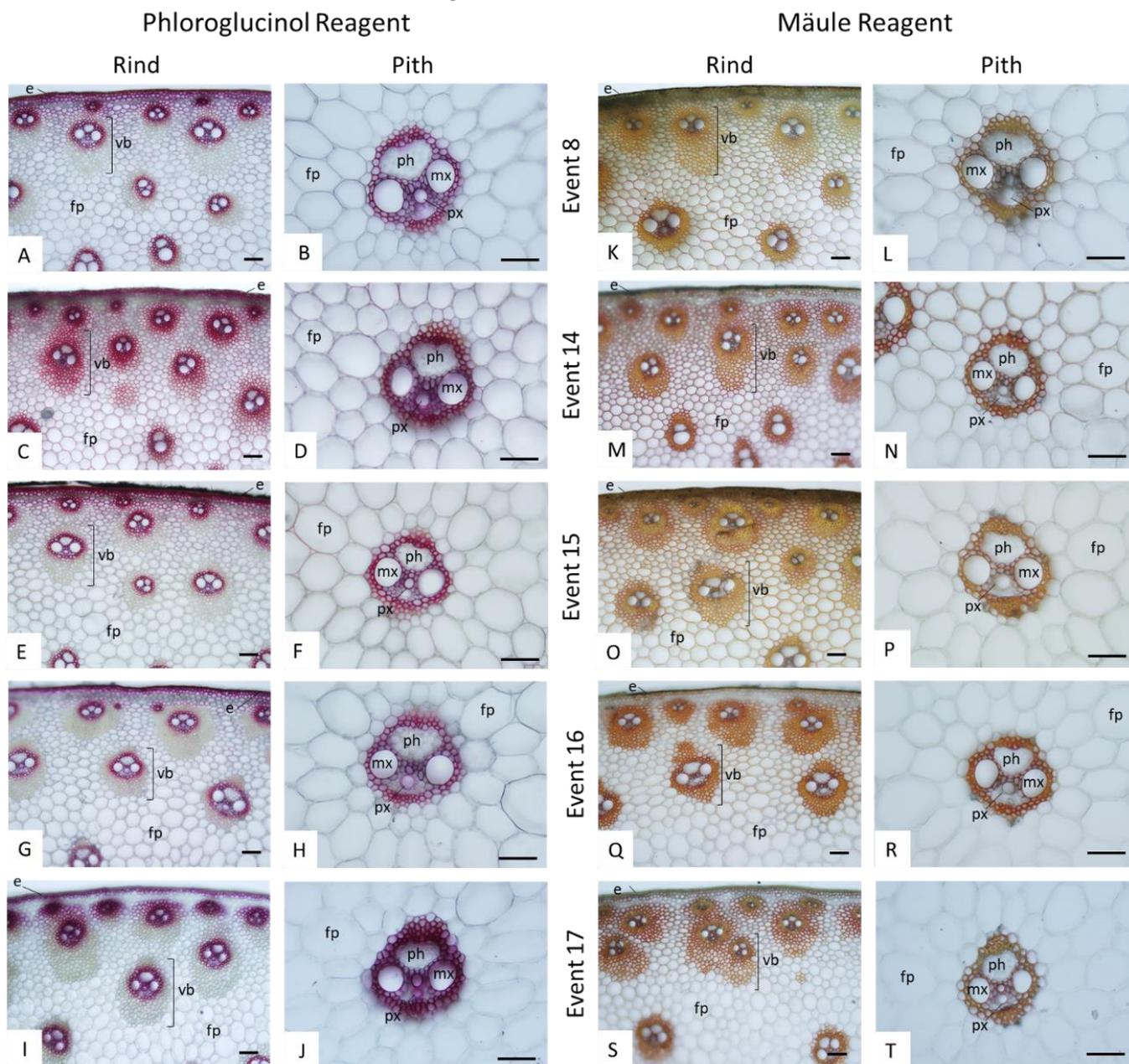
***Shmyb85* – 7th internode**

Figura 31 – Seções transversais de diferentes regiões do internódio 7º em linhagens transgênicas de *Shmyb85 Saccharum* spp. cultivar híbrido SP-3280 submetidas ao reagente Fluoroglucinol e de Mäule para detecção de lignina total, S e G, respectivamente. Colunas da esquerda: Região periférica (Rind); Colunas da direita: Região central do colmo (Pith). e: epiderme; f: fibras; fp: parênquima fundamental; mx: metaxilema; ph: Floema; px: protoxilema; vb: feixe vascular. Barras = 50 µm.

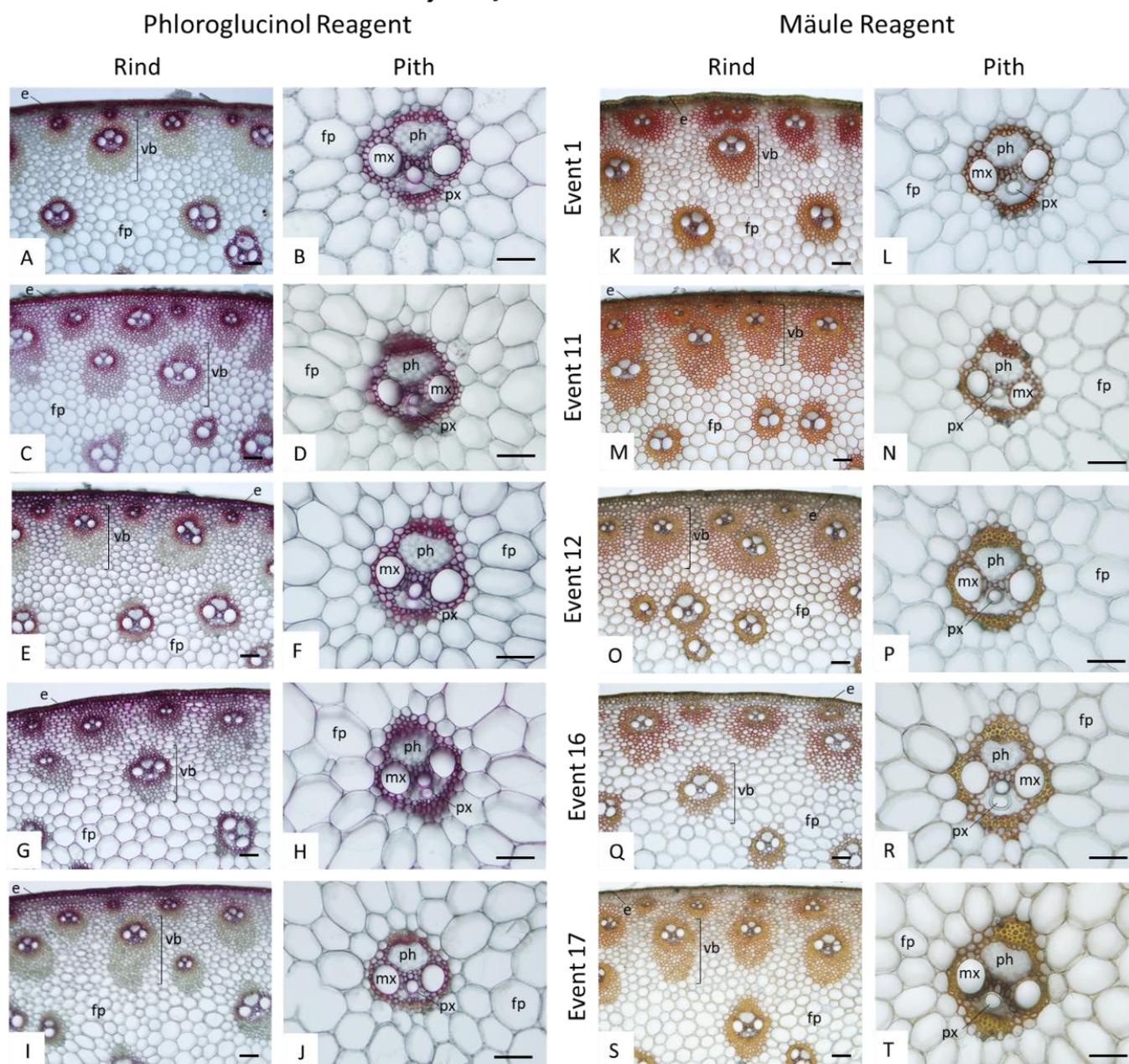
***Shmyb58/63* – 7th internode**

Figura 32 – Seções transversais de diferentes regiões do internódio 7º em linhagens transgênicas de *Shmyb58/63 Saccharum* spp. cultivar híbrido SP-3280 submetidas ao reagente Fluoroglucinol e de Maüle para detecção de lignina total, S e G, respectivamente. Colunas da esquerda: Região periférica (Rind); Colunas da direita: Região central do colmo (Pith). e: epiderme; f: fibras; fp: parênquima fundamental; mx: metaxilema; ph: Floema; px: protoxilema; vb: feixe vascular. Barras = 50 µm.

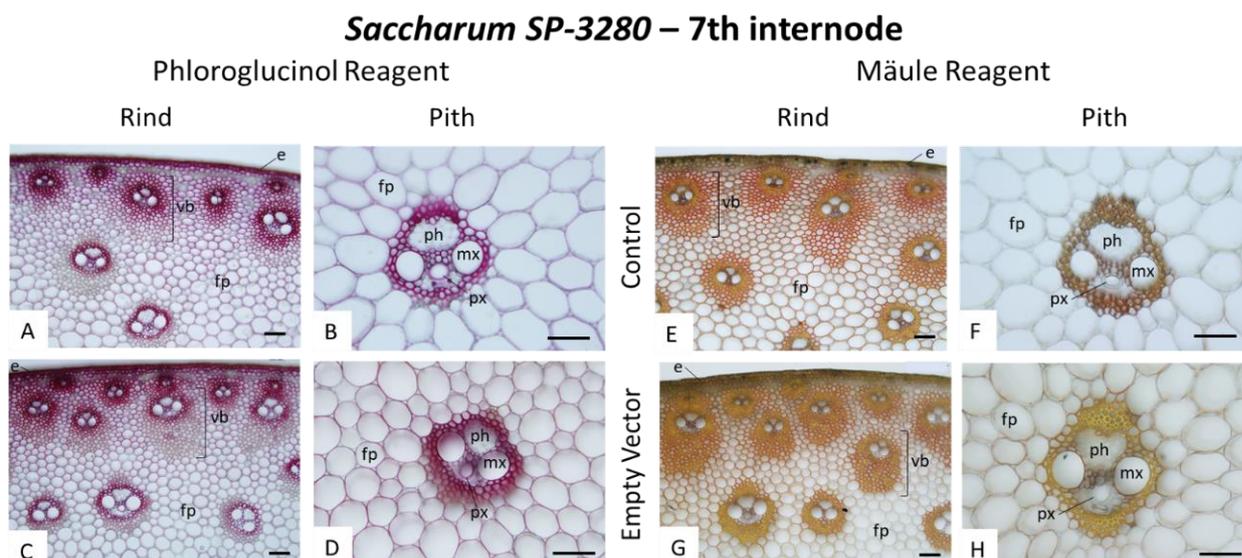


Figura 33 – Secções transversais de diferentes regiões do internódio 7º em controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280 submetidas ao reagente Fluoroglucinol e de Mäule para detecção de lignina total. Colunas da esquerda: Região periférica (Rind); Colunas da direita: Região central do colmo (Pith). e: epiderme; f: fibras; fp: parênquima fundamental; mx: metaxilema; ph: Floema; px: protoxilema; vb: feixe vascular. Barras = 50 µm.

Tabela 8 - Sequências dos pares de primers utilizados na amplificação por RT-qPCR e tamanho do amplicon de genes da rota biossintética da lignina e TFs em linhagens transgênicas: *Shmyb85*, *Shmyb58/63*, EV e WT de *Saccharum* spp. cultivar híbrido SP-3280.

Gene	Forward 5'-->3'	Reverse 5'-->3'	Amplicon (bp)
<i>Sh4CL</i> (RT-qPCR)	AGCCGTTCCAGGTCAAGTC	ACTCGGGGTCGTTCAAGTA	161
<i>ShCAD2</i> (RT-qPCR)	ATCAGCTCGTCGTCCAAGAA	CGATGATGTAGTCCAGCGAGT	120
<i>ShCAD8</i> (RT-qPCR)	TCAAGAACGACTGGGGAAC	GCAGGAGCCGACGAAGTA	140
<i>ShCOMT</i> (RT-qPCR)	GAGGACAAGGACGGCAAGTA	ACCGCGTCCTTGAGGTAGTA	154
<i>ShHCT</i> (RT-qPCR)	TCAGACGACACCGCCTTC	GTCCGCCACGAGTTGAT	137
<i>ShF5H</i> (RT-qPCR)	AGACGCAGGACGGAGTGTT	AAGAGCTTCATCACGCACAG	138
<i>ShMYB85</i> (RT-qPCR)	TTGATGTGGACGAGTTCAGC	TCATGTAGCCATCGACCAAG	132
<i>ShMYB58/63</i> (RT-qPCR)	ACCTGACGTACTGGACCAC	CACCACCAGGAGTTCAGGTT	125
<i>ShCCoAOMT2</i> (RT-qPCR)	CTCGTGACCGACAAGCAC	AGGGAGTAGCCCGTGAACA	130
<i>ShCCoAOMT1</i> (RT-qPCR)	ACGCCGACAAGGACAACACTAC	GCGGTAGAAGCGGATGTACT	151

### **Determinação dos polissacarídeos da parede celular**

De forma geral, *Shmyb85* teve mais impacto sobre a composição da parede celular do que *Shmyb58/63* no internódio 5. Assim, se observa na figura 34A que as linhagens transgênicas 16, 14, 17, e 15 de *Shmyb85* tiveram uma redução porcentual significativa (com respeito ao WT) do conteúdo de celulose (-41.66, -32.41, -26.76 e -14.79, com  $p \leq 0.05$  de 0.0167; 0.0167; 0.0051 e 0.0424, respectivamente). Para o caso de *Shmyb58/63* (Figura 34B) as linhagens 1 e 16 apresentaram redução porcentual do conteúdo deste polímero de -28.27 e -25.86, com  $p \leq 0.05$  de 0.017 e 0.033, respectivamente. No que se refere ao conteúdo de hemicelulose, só a linhagem 15 de *Shmyb85* apresentou redução porcentual significativa (-26.1,  $p \leq 0.05$ , 0.017, Figura 34C). Em *Shmyb58/63* também só uma linhagem transgênica (12) experimentou alteração no conteúdo deste polímero, com um aumento significativo porcentual (29.58,  $p \leq 0.05$ , 0.012, Figura 34D). É importante destacar que apesar de que em ambos mutantes só uma linhagem transgênica apresentou alteração significativa, é evidente a resposta oposta entre estes desde que em *Shmyb85* se evidencia uma tendência de aumento deste polímero e o contrário acontece com *Shmyb58/63*. De outro lado, para pectina, as linhagens transgênicas 16, 14, 17, e 15 de *Shmyb85* apresentaram uma redução porcentual significativa de -53.87, -51.89, -43.18, -42.79 ( $p \leq 0.05$  com 0.0227; 0.0074; 0.0244 e 0.0188, respectivamente, Figura 34E). Não houve alteração significativa do conteúdo de pectina em nenhuma das linhagens de *Shmyb58/63*,  $p \geq 0.05$  (Figura 34F).

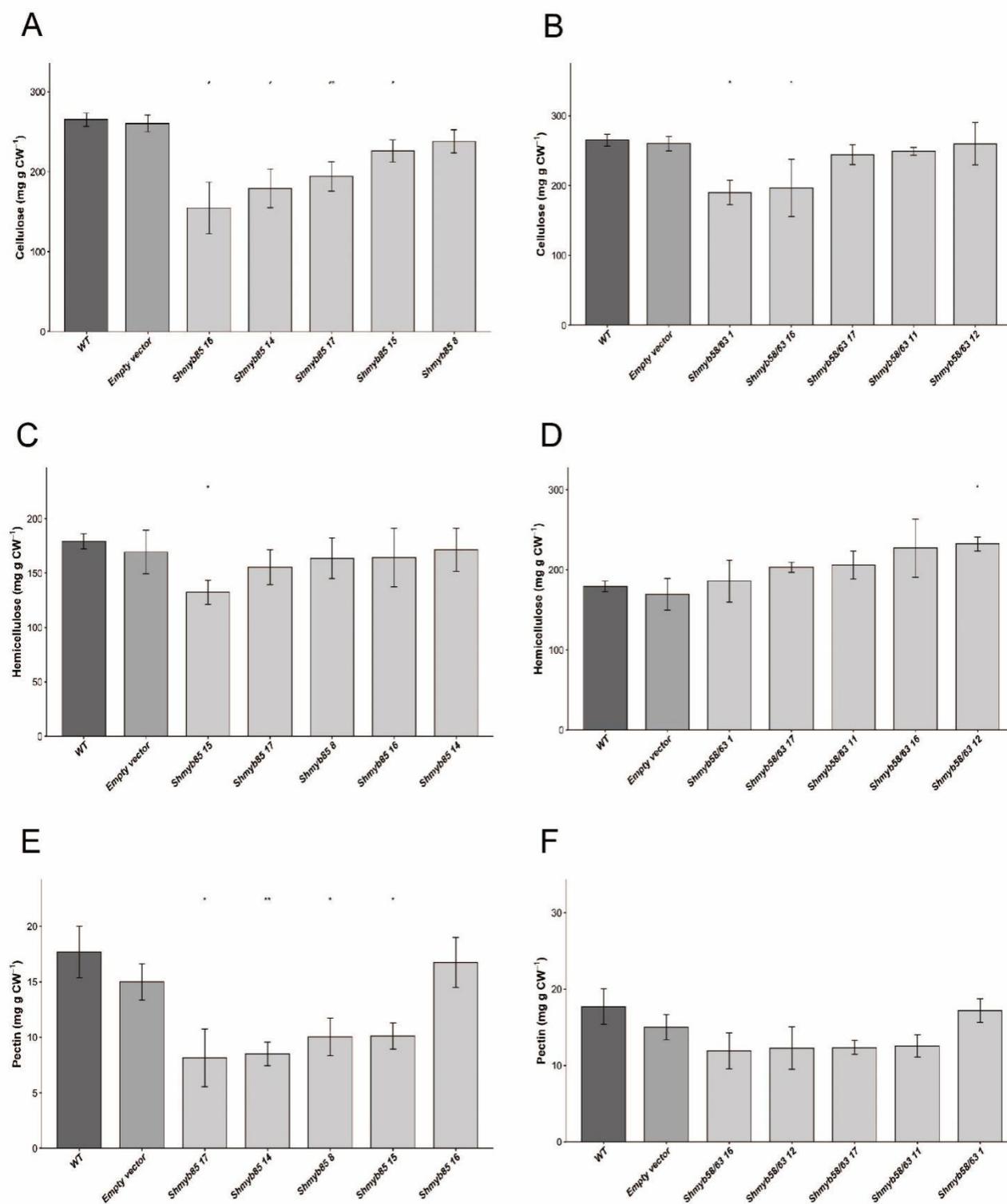


Figura 34 – A e B) Conteúdo de celulose C e D) hemicelulose e E e F) pectina no 5° internódio em linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo de lignina e razão S/G entre os tratamentos. As médias

foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das médias produto de 3-5 replicatas biológicas.

### **Determinação da porcentagem de sacarificação**

Na figura 35A, em *Shmyb85* se observa um aumento porcentual significativo nas linhagens 8, 15, 17, 14, 16 de 50.11%, 63.28%, 90,11%, 109,89% e 257.71%, respectivamente ( $p \leq 0.05$  de 0.0051; 0.0061; 0.0025; 0.0167 e 0.0167). Só as linhagens 1 e 16 de *Shmyb58/63* experimentaram aumento porcentual significativo de 89.13 e 101.79, respectivamente ( $p \leq 0.05$  de 0.017 e 0.017, respectivamente, Figura 35B).

### **Determinação e perfil de compostos fenólicos solúveis totais**

Na figura 36A, se observa que só as linhagens 16 e 17 de *Shmyb85* apresentaram um aumento porcentual significativo nos fenóis totais de 41.54 e 45.55, respectivamente ( $p \leq 0.05$  de 0.0167; 0.0025, respectivamente). Não houve alteração significativa do conteúdo de fenóis totais em nenhuma das linhagens de *Shmyb58/63* (Figura 36B). No que diz respeito ao perfil de compostos fenólicos (Tabela 9) de maneira geral se observa que *Shmyb85* teve mais impacto do que *Shmyb58/63*. Na Tabela 9A, em *Shmyb85*, os conteúdos dos ácidos clorogênico e coumárico só apresentaram alterações percentuais significativas na linhagem 16, 547.32 e -70.66, respectivamente ( $p \leq 0.05$  de 0.000171; 0.01805, respectivamente). Para o ácido quinico só a linhagem 17 experimentou um aumento porcentual significativo de 125.82 ( $p \leq 0.05$ , 0.016963). Não houve alterações no ácido shiquimico. *Shmyb58/63* (Tabela 9B) só apresentou alteração no ácido clorogênico, para o qual se observou redução porcentual nas linhagens 12, 1, 17 e 11 (-85.49, -71.75, -35.87, -59.54, respectivamente)  $p \leq 0.05$ , 0.024847, 0.010530, 0.049137, 0.012081).

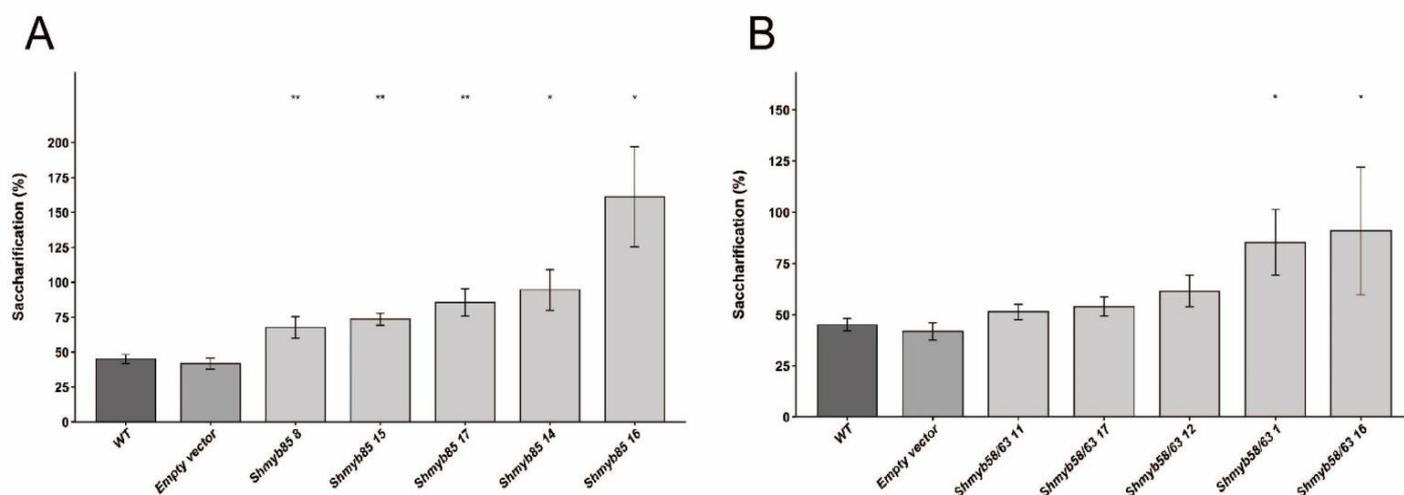


Figura 35 – Porcentagem de sacarificação no 5° internódio em linhagens transgênicas de *Shmyb85* (A), *Shmyb58/63* (B) e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo de lignina e razão S/G entre os tratamentos. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das médias produto de 3-5 replicatas biológicas.

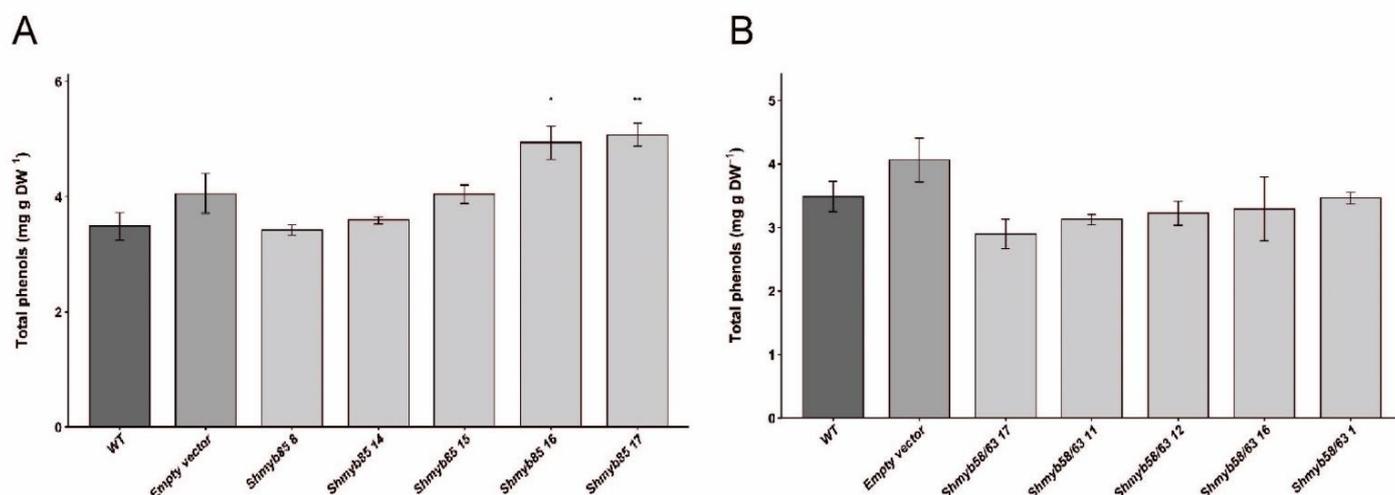


Figura 36 – Conteúdo de fenóis totais no 5° internódio em linhagens transgênicas de *Shmyb85* (A), *Shmyb58/63* (B) e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo de lignina e razão S/G entre os tratamentos. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das médias produto de 3-5 replicatas biológicas.

Tabela 9A - Perfil de compostos fenólicos em linhagens transgênicas de *Shmyb85* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Letras diferentes indicam diferenças para um composto fenólico entre distintos genótipos. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . O erro padrão e produto da média de 3-5 replicatas biológicas.

Genotype	Chlorogenic acid (mg g <sup>-1</sup> DW)	Coumaric acid (mg g <sup>-1</sup> DW)	Quinic acid (mg g <sup>-1</sup> DW)	Shikimic acid (mg g <sup>-1</sup> DW)
WT	0.131 ± 0.026 b	0.075 ± 0.011 a	1.615 ± 0.189 b	0.026 ± 0.005 a
EV	0.166 ± 0.037 b	0.075 ± 0.010 a	1.597 ± 0.318 b	0.016 ± 0.003 a
<i>Shmyb85 8</i>	0.039 ± 0.012 b	0.039 ± 0.003 ab	3.070 ± 0.241 ab	0.037 ± 0.005 a
<i>Shmyb85 16</i>	0.848 ± 0.273 a	0.022 ± 0.001 b	2.124 ± 0.520 ab	0.024 ± 0.002 a
<i>Shmyb85 17</i>	0.154 ± 0.053 b	0.055 ± 0.004 ab	3.647 ± 0.339 a	0.041 ± 0.007 a
<i>Shmyb85 15</i>	0.302 ± 0.104 b	0.037 ± 0.005 ab	1.746 ± 0.536 b	0.027 ± 0.011 a
<i>Shmyb85 14</i>	0.057 ± 0.043 b	0.044 ± 0.013 ab	2.565 ± 0.547 ab	0.033 ± 0.009 a

Tabela 9B - Perfil de compostos fenólicos em linhagens transgênicas de *Shmyb58/63* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Letras diferentes indicam diferenças para um composto fenólico entre distintos genótipos. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . O erro padrão e produto da média de 3-5 replicatas biológicas.

Genotype	Chlorogenic acid (mg g <sup>-1</sup> DW)	Coumaric acid (mg g <sup>-1</sup> DW)	Quinic acid (mg g <sup>-1</sup> DW)	Shikimic acid (mg g <sup>-1</sup> DW)
WT	0.131 ± 0.026 ab	0.075 ± 0.011 a	1.615 ± 0.189 a	0.026 ± 0.005 a
EV	0.166 ± 0.012 a	0.075 ± 0.010 a	1.597 ± 0.318 a	0.016 ± 0.003 a
<i>Shmyb58/63 1</i>	0.037 ± 0.019 b	0.055 ± 0.003 a	1.615 ± 0.719 a	0.018 ± 0.006 a
<i>Shmyb58/63 17</i>	0.084 ± 0.062 b	0.056 ± 0.005 a	2.696 ± 0.461 a	0.030 ± 0.005 a
<i>Shmyb58/63 16</i>	0.133 ± 0.051 ab	0.038 ± 0.005 a	2.234 ± 0.567 a	0.013 ± 0.006 a
<i>Shmyb58/63 12</i>	0.019 ± 0.008 c	0.065 ± 0.009 a	2.506 ± 0.350 a	0.031 ± 0.004 a
<i>Shmyb58/63 11</i>	0.053 ± 0.020 b	0.066 ± 0.013 a	1.253 ± 0.277 a	0.023 ± 0.006 a

### Determinação de Flavonóides e Antocianinas

Na Figura 37A se observa que para as linhagens 16, 11, 14, 15, 17 e 19 de *Shmyb85* houve redução percentual de flavonóides em -54.11, -61.33, -50.63, -55.28, -72.30 e -50.16, respectivamente ( $p \leq 0.05$ , 0.0173, 0.0012, 0.0037, 0.0059, 0.0002 e 0.0212). Para *Shmyb58/63* (Figura 37B) as linhagens 20, 23, 16 e 4 apresentaram redução percentual de -58.52, -53.10, -57.39, -56.50, respectivamente ( $p \leq 0.05$ , 0.0188, 0.0042, 0.0097, 0.0337). *Shmyb85* apresentou redução percentual de antocianinas (Figura 37C) nas linhagens 17, 18 e 2 em -39.55, -60.86 e -37.60, respectivamente ( $p \leq 0.05$ , 0.025, 0.00018 e 0.0438). Não houve mudanças em antocianinas em nenhuma linhagem para *Shmyb58/63* (Figura 37D).

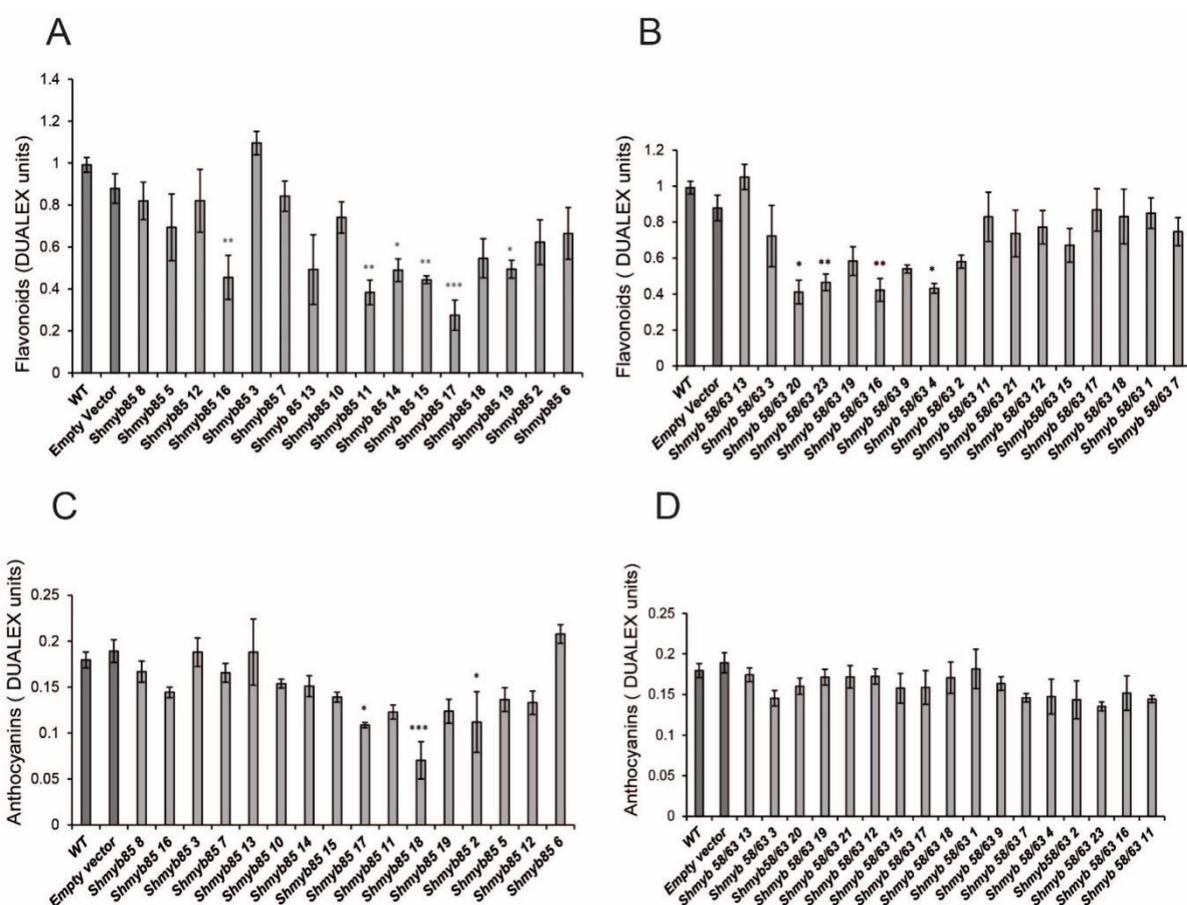


Figura 37 – A e B) Unidades DUALEX de Flavonóides e C e D) de Antocianinas na folha +3 em linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas no conteúdo das substâncias flavonóides e antocianinas. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o erro padrão das médias produto de 3-10 replicatas biológicas.

### Indução da Lignificação (injúria)

O teste preliminar de expressão nas plantas controle (WT) para verificar se a folha possui capacidade de experimentar indução da expressão dos TFs, frente à injúria mecânica provocada neste estudo, evidenciou um significativo aumento da expressão de *ShMYB85* e *ShMYB58/63* (3.19 e 5.42 vezes, respectivamente,  $p=0.01$ ) (Figura 38). Por outro lado, houve uma estimulação da expressão para *Sh4CL*, *ShCOMT*, *ShF5H*, *ShHCT* (5.55, 8.94, 4.64 e 3.84 vezes, respectivamente,  $p=0.01$ ). Não houve estimulação para *ShCAD8* e *ShCAD2*. Desde que a maior indução foi observada para *ShCOMT* dentro deste grupo de genes representativos na rota de biossíntese de lignina, *ShCOMT*, foi avaliada nas folhas injuriadas e não injuriadas das linhagens transgênicas de *Shmyb85* e *Shmyb58/63* com o intuito de corroborar o papel destes TFs na lignificação. Assim na Figura 39 se observa uma queda paralela no coeficiente da indução de expressão relativa de *ShMYB85* e *ShCOMT* nas linhagens transgênicas 7, 3, 2, 18, 16 e 14 de *Shmyb85*. No que se refere a *Shmyb 58/63* houve queda paralela frente a *ShCOMT* nas linhagens transgênicas 9, 7, 4, 2, 21, 20, 1, 18, 15, 13, 12 e 11. Entre os coeficientes da indução de deposição de lignina acetyl bromide e indução da expressão relativa das linhagens transgênicas de *Shmyb85* e *Shmyb58/63* houve queda paralela para as linhagens transgênicas 8, 7, 3, 19, 18, 16, 14 e 9, 7, 4, 2, 23, 1, 18, 17, 16, 15, 12, 11, respectivamente (Figura 40). Dos resultados anteriormente indicados podemos inferir que existe uma relação direta entre os coeficientes de indução de expressão de *ShCOMT*, *ShMYB85* e *ShMYB58/63* e o coeficiente de indução de deposição de lignina acetyl bromide das linhagens transgênicas mencionadas acima para *Shmyb85* e *Shmyb58/63*.

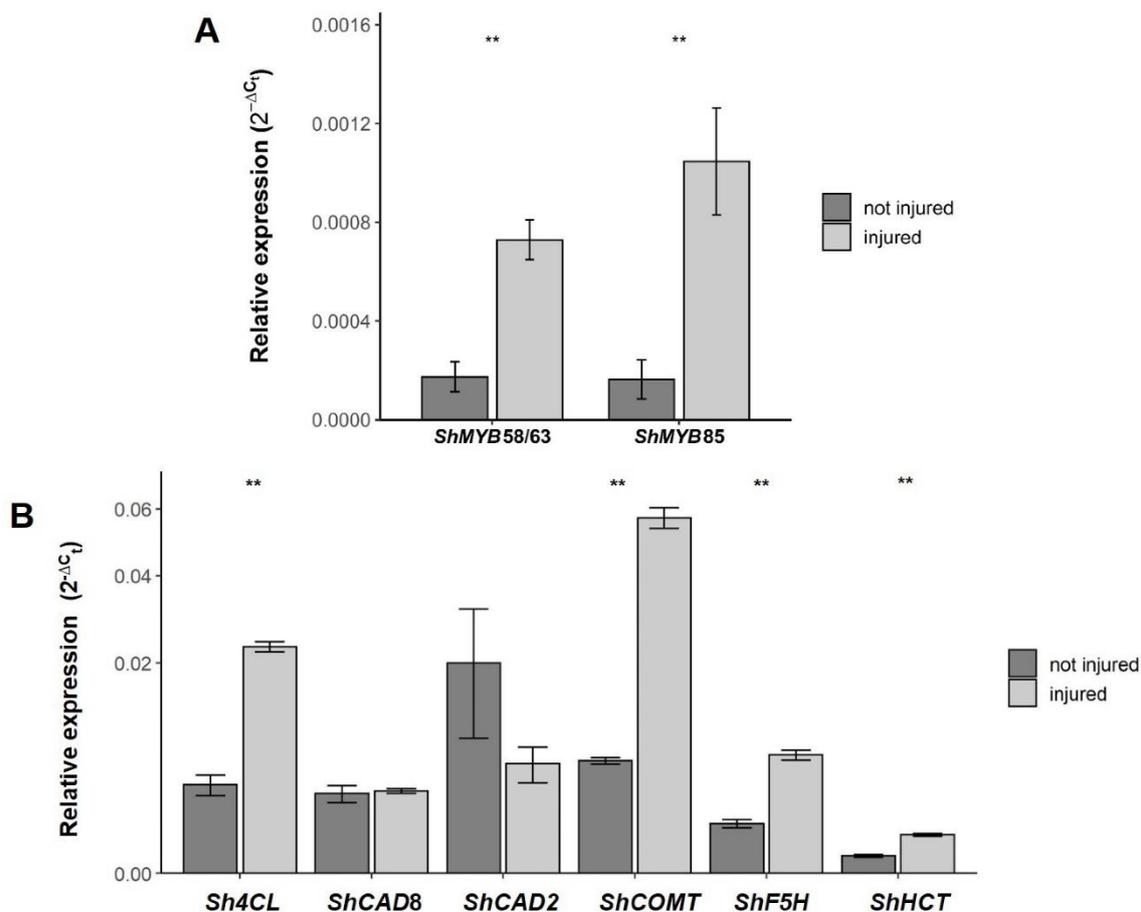


Figura 38 – Expressão relativa (RT-qPCR) de genes TFs (A): *ShMYB85*; *ShMYB58/63* e da rota biossintética da lignina (B) em folhas injuriadas e não injuriadas do controle (WT) de *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas na expressão relativa do gene entre os tratamentos. As médias foram comparados pelo t-test,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das médias de quintuplicatas biológicas independentes.

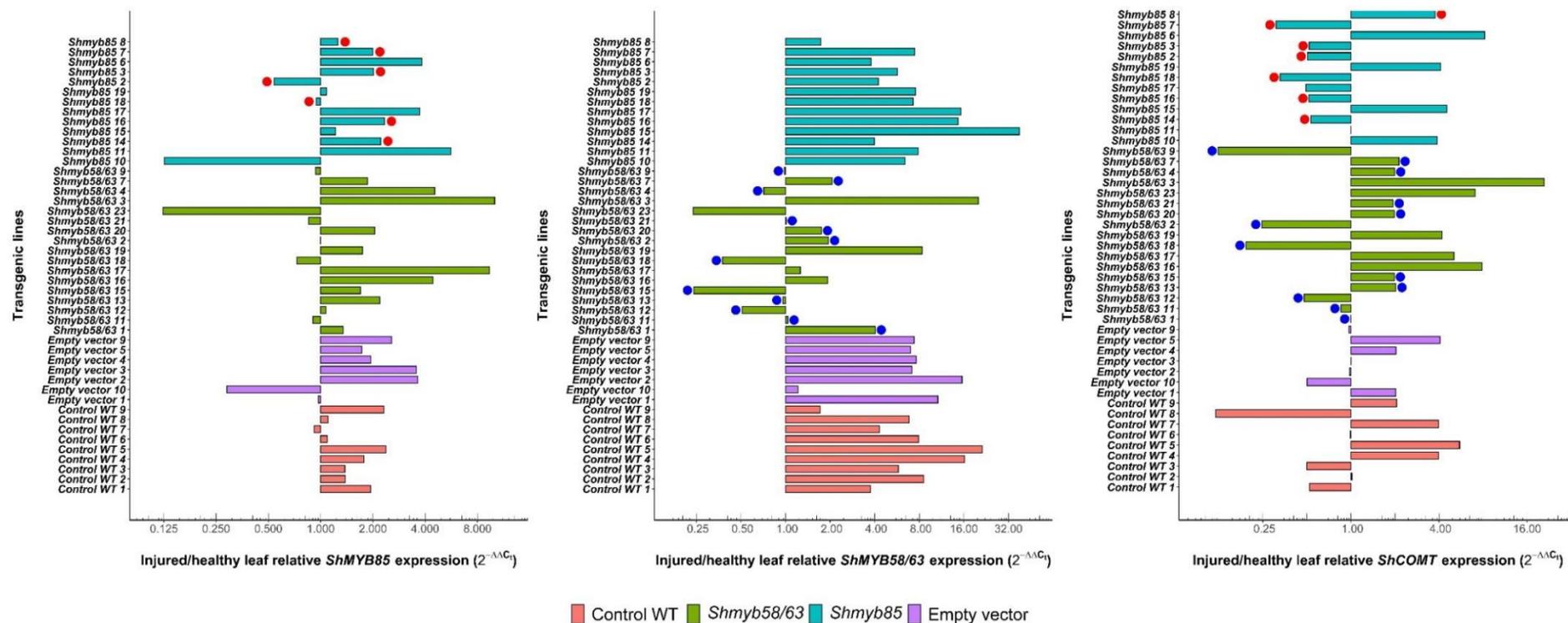


Figura 39 – Coeficiente da indução de expressão relativa (folhas injuriadas/folhas não injuriadas) de *ShMYB85*, *ShMYB58/63* e *ShCOMT* em linhagens transgênicas de *Shmyb85* e *Shmyb58/63*. Os círculos vermelhos representam a queda paralela do coeficiente em *ShMYB85* e *ShCOMT* nas linhagens transgênicas de *Shmyb85*. Os círculos azuis representam a queda paralela do coeficiente em *ShMYB58/63* e *ShCOMT* nas linhagens transgênicas *Shmyb58/63*. WT e EV são considerados como controles + e – respectivamente.

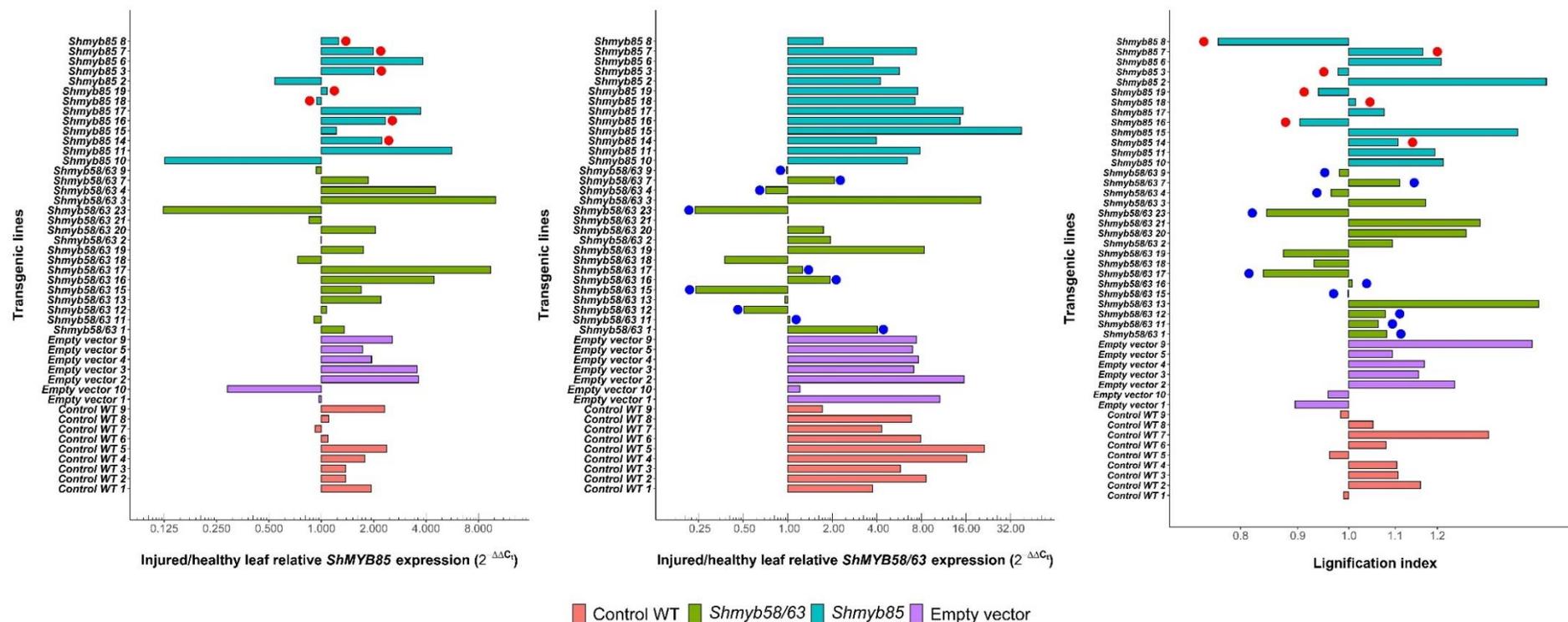


Figura 40 – Coeficiente da indução de expressão relativa (folhas injuriadas/folhas não injuriadas) de *ShMYB85*, *ShMYB58/63* e coeficiente de indução da lignificação em distintas linhagens transgênicas de *Shmyb85* e *Shmyb58/63*. Os círculos vermelhos representam a queda paralela do coeficiente de expressão de *ShMYB85* e lignificação em linhagens transgênicas de *Shmyb85*. Os círculos azuis representam a queda paralela do coeficiente de expressão de *ShMYB58/63* e lignificação em linhagens transgênicas de *Shmyb58/63*. WT e EV são considerados como controles + e – respectivamente.

### Análise por espectroscopia 2D-NMR

As modificações estruturais produzidas no polímero de lignina das paredes celulares dos transgênicos de *Shmyb85* e *Shmyb58/63* foram avaliadas “*in situ*” usando experimentos de 2D-HSQC NMR no estado de gel, sem a necessidade de isolar a lignina, comparando os espectros dos transgênicos com aqueles correspondentes aos controles WT e EV. As regiões aromáticas/insaturadas ( $\delta_C/\delta_H$  90–148/5.7–7.9) dos espectros 2D-HSQC NMR das amostras dos transgênicos e controles, em conjunto com as principais subestruturas identificadas, são apresentados na Figura 41. Os sinais cruzados atribuídos nos espectros HSQC estão listados na Tabela 10. Nas amostras, os principais sinais cruzados observados na região aromática/insaturada dos espectros HSQC corresponderam aos anéis aromáticos e às cadeias laterais insaturadas das diferentes unidades de lignina: *p*-hidroxifenil (**H**), guaiacil (**G**) e siringil (**S**), e dos *p*-hidroxicinamatos (*p*-cumaratos, **pCA** e ferulatos, **FA**) que estão associados à lignina. Sinais para tricina (**T**), uma flavona que está presente na lignina de gramíneas foram também observados em todas as amostras. As abundâncias relativas das diferentes unidades de lignina, *p*-hidroxicinamatos e tricina, estimadas a partir das integrais do volume de seus sinais nos espectros das linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e os controles, são apresentados nas tabelas 11A e 11B.

A análise de 2D-NMR, evidencia uma diminuição significativa dos diferentes monômeros de lignina nas linhagens transgênicas de *Shmyb85* e *Shmyb58/63*. Assim em *Shmyb85*, os monômeros: H, G e S, na linhagem transgênica 16 tiveram a maior queda (0.70%, 7.10% e 5.50%, respectivamente) ( $p \leq 0.05$ ; 0.00015, 0.00016, 0.00015) e no caso de *Shmyb58/63* foi para a linhagem transgênica 17 (0.87%, 5.80% e 5.85%, respectivamente) ( $p \leq 0.05$ ; 0.005, 0.00015 e 0.00106) quando comparando ambas linhagens com os controles WT e EV (1.06-1.20, 8.70-10.57 e 7.42-10.50, respectivamente). No que respeita à razão S/G, em *Shmyb85*, as linhagens 16 e 15 apresentaram redução neste parâmetro, sendo só significativa para esta última (0.69) ( $p \leq 0.05$ ; 0.046457) e em *Shmyb58/63*, houve aumento na razão, sendo significativo esta alteração para as linhagens 17 e 16 (1.00 e 0.98, respectivamente) ( $p \leq 0.05$ ; 0.01204 e 0.03074) em relação aos controles WT e EV (0.83% e 0.88%, respectivamente). Em destaque, a abundancia relativa das unidades totais de lignina (soma das unidades H, G e S, em relação às unidades totais de lignina + carboidratos)

apresentou redução significativa em todas as linhagens transgênicas tanto para *Shmyb85*: 16; 13.30%,15; 15.63% e 14; 15.37% ( $p \leq 0.05$ ; 0.00015, respectivamente) como para *Shmyb58/63*: 17; 12.45%,16; 13.63%,12; 13.43% ( $p \leq 0.05$ , 0.000152, respectivamente) quando comparando aos controles WT e EV (17.18% e 21.07%, respectivamente). A flavona tricina apresentou redução significativa somente na linhagem de *Shmyb85* 15: 0.53% ( $p \leq 0.05$ ; 0.02698), enquanto nenhuma das linhagens de *Shmyb58/63* apresentou alteração desta flavona ( $p = 0.45703$ ) em relação aos controles WT e EV (0.73% e 0.80%, respectivamente). De outro lado todas as linhagens avaliadas para *Shmyb85* apresentaram queda para o caso do *p*-coumarato, sendo que a menor abundância relativa foi para a linhagem 16: 7.60% ( $p \leq 0.05$ ; 0.00254) em sentido contrário com as linhagens de *Shmyb58/63* onde não houve alterações ( $p = 0.52035$ ) neste *p*-hidroxicinamato com respeito aos controles WT e EV (9.35% e 8.93%). O outro *p*-hidroxicinamato avaliado aqui: o Ferulato, apresentou redução significativa somente na linhagem transgênica de *Shmyb85* 15: 4.40% ( $p \leq 0.05$ ; 0.01152) e *Shmyb58/63* 17: 4.87% % ( $p \leq 0.05$ ). O coeficiente *pCA/FA* não apresentou alterações tanto para *Shmyb85* ( $p = 0.12463$ ) nem para *Shmyb58/63* ( $p = 0.28509$ ).

Por outro lado, adicionalmente na análise de 2D-NMR, foi verificado os carboidratos presentes na parede celular. A região dos espectros 2D-NMR mostrando as correlações dos carbonos anoméricos dos diferentes carboidratos presentes nas amostras das linhagens de *Shmyb85*, *Shmyb58/63* e controles (WT e EV) é apresentada na Figura 42, e as atribuições dos sinais observados estão listadas na Tabela 12. Os principais sinais corresponderam aos arabinoxilanos, e a análise 2D-NMR não detectou alterações estruturais importantes na composição dos carboidratos.

Tabela 10 - Atribuições dos sinais de correlação de lignina nos espectros 2D HSQC em linhagens transgênicas de *Shmyb85* e controles (EV e WT) de *Saccharum* spp. cultivar híbrido SP-3280. Os sinais foram atribuídos por comparação com a literatura (del Río et al., 2012, 2015).

Label	$\delta_C/\delta_H$	Assignment
T <sub>8</sub>	93.9/6.54	C <sub>8</sub> /H <sub>8</sub> in tricin units ( <b>T</b> )
T <sub>6</sub>	98.6/6.20	C <sub>6</sub> /H <sub>6</sub> in tricin units ( <b>T</b> )
S <sub>2,6</sub>	103.8/6.69	C <sub>2</sub> /H <sub>2</sub> and C <sub>6</sub> /H <sub>6</sub> in etherified syringyl units ( <b>S</b> )
T <sub>2',6'</sub>	104.0/7.30	C <sub>2</sub> /H <sub>2</sub> and C <sub>6</sub> /H <sub>6</sub> in tricin units ( <b>T</b> )
G <sub>2</sub>	110.9/6.99	C <sub>2</sub> /H <sub>2</sub> in guaiacyl units ( <b>G</b> )
FA <sub>2</sub>	110.9/7.31	C <sub>2</sub> /H <sub>2</sub> in ferulates ( <b>FA</b> )
pCA <sub>8</sub> and FA <sub>8</sub>	113.9/6.48	C <sub>8</sub> /H <sub>8</sub> in <i>p</i> -coumarates ( <b>pCA</b> ) and ferulates ( <b>FA</b> )
H <sub>3,5</sub>	114.5/6.68	C <sub>3</sub> /H <sub>3</sub> and C <sub>5</sub> /H <sub>5</sub> in <i>p</i> -hydroxyphenyl units ( <b>H</b> )
G <sub>5/6</sub>	114.9/6.72	C <sub>5</sub> /H <sub>5</sub> and C <sub>6</sub> /H <sub>6</sub> in guaiacyl units ( <b>G</b> )
FA <sub>5</sub>	115.3/6.77	C <sub>5</sub> /H <sub>5</sub> in ferulates ( <b>FA</b> )
pCA <sub>3,5</sub>	115.5/6.77	C <sub>3</sub> /H <sub>3</sub> and C <sub>5</sub> /H <sub>5</sub> in <i>p</i> -coumarates ( <b>pCA</b> )
G <sub>6</sub>	119.0/6.76	C <sub>6</sub> /H <sub>6</sub> in guaiacyl units ( <b>G</b> )
FA <sub>6</sub>	123.1/7.11	C <sub>6</sub> /H <sub>6</sub> in ferulates ( <b>FA</b> )
H <sub>2,6</sub>	127.6/7.17	C <sub>2</sub> /H <sub>2</sub> and C <sub>6</sub> /H <sub>6</sub> in <i>p</i> -hydroxyphenyl units ( <b>H</b> )
pCA <sub>2,6</sub>	130.1/7.45	C <sub>2</sub> /H <sub>2</sub> and C <sub>6</sub> /H <sub>6</sub> in <i>p</i> -coumarates ( <b>pCA</b> )
pCA <sub>7</sub> and FA <sub>7</sub>	145.2/7.56	C <sub>7</sub> /H <sub>7</sub> in <i>p</i> -coumarates ( <b>pCA</b> ) and ferulates ( <b>FA</b> )

Tabela 11A - Análise 2D-RMN semiquantitativa em linhagens transgênicas de *Shmyb85* e controles (EV e WT) de *Saccharum* spp. cultivar híbrido SP-3280. A composição da amostra representa a abundância relativa de unidades de lignina (**H**, **G**, **S**), tricina (**T**) e *p*-hidroxicinamatos (**pCA**, **FA**), de acordo com a soma das unidades totais de lignina e carboidratos\*. Letras diferentes representam diferenças para um composto fenólico entre distintos genótipos. As medias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . O erro padrão e produto da média de de 3 replicatas biológicas.

Genotype	H(%)	G(%)	S(%)	(H+G+S)	S/G	Tricine(%)	pCA(%)	FA(%)	pCA/FA (%)
WT	1.06±0.03 b	8.70±0.24 b	7.42±0.16 b	17.18±0.17 b	0.83±0.02 b	0.73±0.04 a	9.35±0.27 a	5.27±0.11 a	1.79±0.08 a
EV	1.20±0.00 a	10.57±0.08 a	9.30±0.05 a	21.07±0.12 a	0.88±0.01 ab	0.80±0.00 a	8.93±0.20 a	5.57±0.17 a	1.62±0.03 a
<i>Shmyb85 16</i>	0.70±0.00 c	7.10±0.30 c	5.50±0.36 c	13.30±0.66 d	0.77±0.01 bc	0.73±0.03 a	7.60±0.28 b	5.10±0.20 ab	1.49±0.01 a
<i>Shmyb85 15</i>	0.46±0.03 d	8.97±0.29 b	6.20±0.17 c	15.63±0.44 c	0.69±0.01 c	0.53±0.06 b	7.70±0.20 b	4.40±0.36 b	1.77±0.10 a
<i>Shmyb85 14</i>	0.70±0.00 c	7.37±0.29 c	7.30±0.41 bc	15.37±0.33 c	1.00±0.08 a	0.63±0.06 a	7.97±0.14 b	4.90±0.10 ab	1.63±0.06 a

Tabela 11B - Análise 2D-RMN semiquantitativa em linhagens transgênicas de *Shmyb58/63* e controles (EV e WT) *Saccharum* spp. cultivar híbrido SP-3280. A composição da amostra representa a abundância relativa de unidades de lignina (**H**, **G**, **S**), tricina (**T**) e *p*-hidroxicinamatos (**pCA**, **FA**), de acordo com a soma das unidades totais de lignina e carboidratos\*. Letras diferentes indicam diferenças para um composto fenólico entre distintos genótipos. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . O erro padrão e produto da média de 3 replicatas biológicas

Genotype	H(%)	G(%)	S(%)	(H+G+S)	S/G	Tricine(%)	pCA(%)	FA(%)	pCA/FA (%)
WT	1.06±0.03 a	8.70±0.24 b	7.42±0.16 b	17.18±0.17 b	0.83±0.02 b	0.73±0.04 a	9.35±0.27 a	5.27±0.11 a	1.79±0.08 a
EV	1.20±0.00 a	10.57±0.08 a	9.30±0.05 a	21.07±0.12 a	0.88±0.01 ab	0.80±0.00 a	8.93±0.20 a	5.57±0.17 a	1.62±0.03 a
<i>Shmyb58/63 17</i>	0.87±0.06 b	5.80±0.30 c	5.85±0.25 c	12.45±0.44 c	1.00±0.01 a	0.67±0.03 a	8.97±0.46 a	4.87±0.13 b	1.84±0.08 a
<i>Shmyb58/63 16</i>	1.03±0.03 a	6.40±0.51 c	6.20±0.20 c	13.63±0.74 c	0.98±0.06 a	0.80±0.10 a	8.97±0.31 a	4.90±0.11 a	1.83±0.05 a
<i>Shmyb58/63 12</i>	1.08±0.07 a	6.45±0.15 c	5.90±0.13 c	13.43±0.26 c	0.91±0.00 ab	0.70±0.04 a	8.65±0.24 a	5.28±0.16 a	1.65±0.08 a

\* As abundâncias das unidades de lignina (**H**, **G**, **S**), tricina (**T**) e *p*-hidroxicinamatos (**pCA**, **FA**) foram calculadas a partir da integração de seus respectivos sinais e as unidades de carboidratos (xilose, glicose e 4-O-metil- $\alpha$ -D-ácido glucurônico) da integração dos sinais de carbono anomérico, sendo referidos à lignina total (H+G+S) e unidades de carboidrato (unidades de lignina + carboidrato = 100).

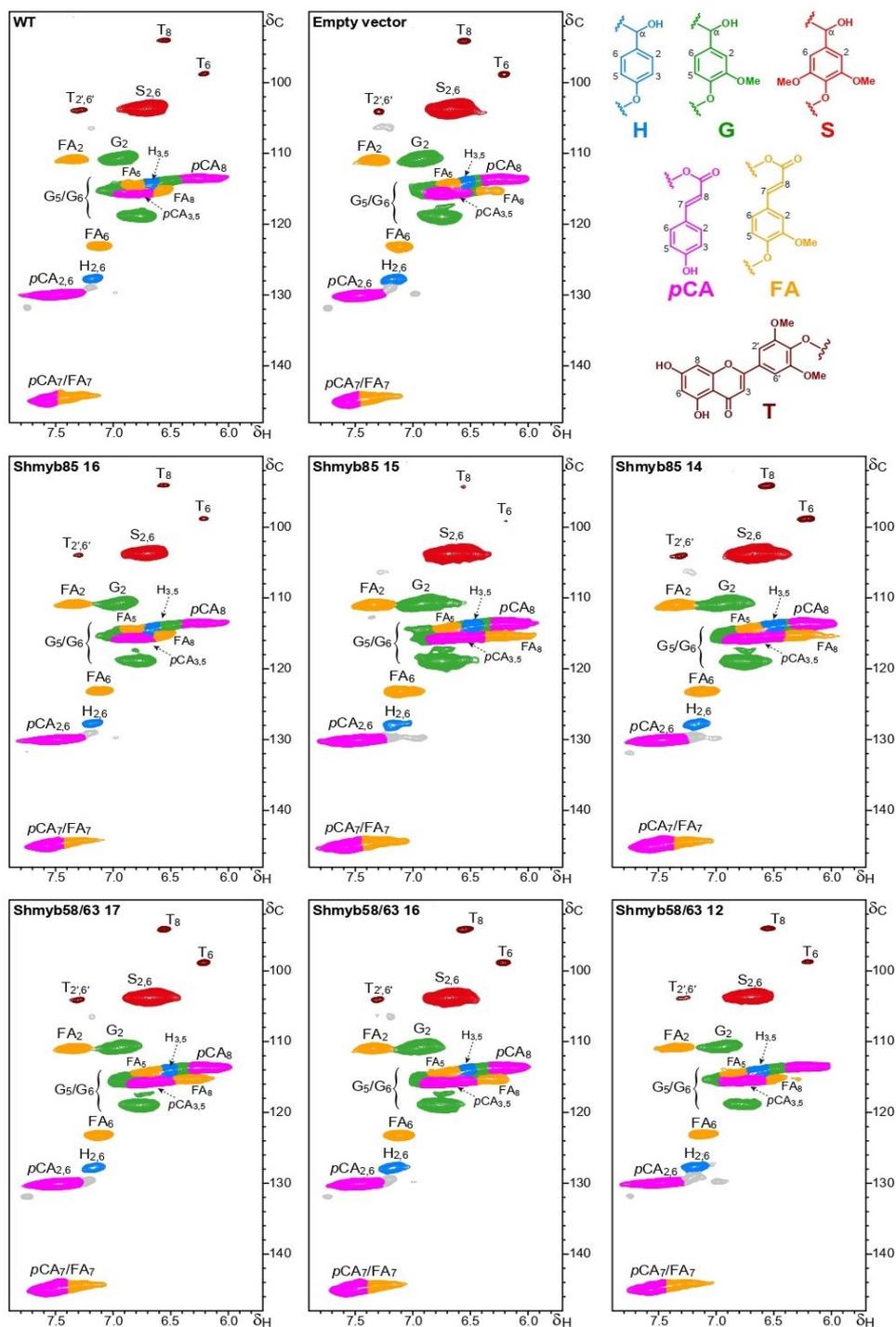


Figura 41 – Caracterização estrutural da lignina por 2D-RMN. Os espectros 2D-HSQC NMR (em DMSO-  $d_6$ ) são apresentados para as regiões aromáticas/insaturadas ( $\delta_C/\delta_H$  90–148/5.7–7.9) de paredes celulares das linhagens de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) em *Saccharum* spp. cultivar híbrido SP-3280. As principais estruturas encontradas são as unidades: p-hidroxifenil (H), guaiacil (G) e siringil (S), e os p-hidroxicinamatos (*p*-cumarato, pCA e ferulato, FA). As estruturas são coloridas para coincidir com os contornos atribuídos nos espectros de NMR. Valores integrados para cada monômero de lignina (H, G, S), pCA e FA (de acordo com a soma das unidades totais de lignina e carboidratos). As atribuições de sinal estão listadas na Tabela 10.

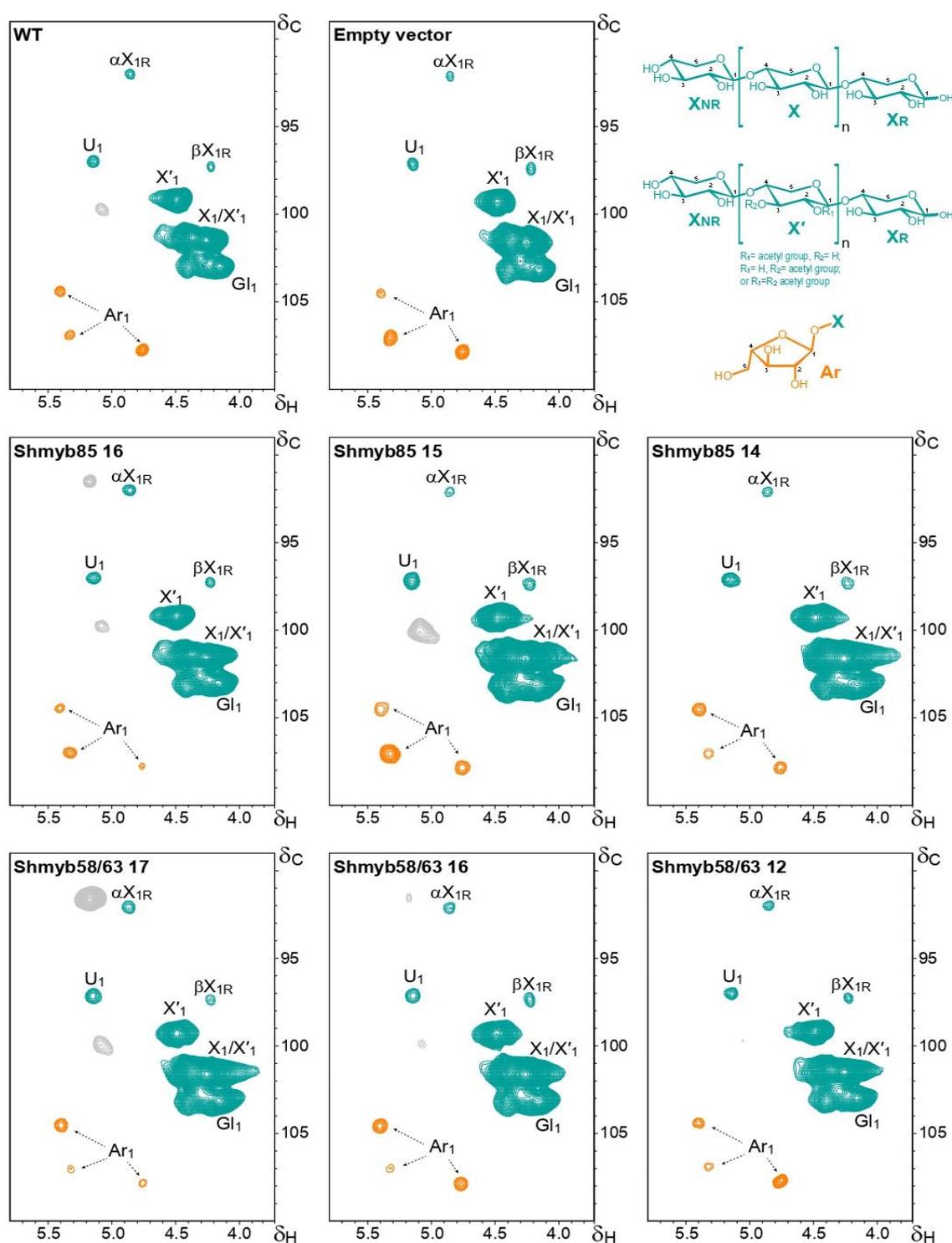


Figura 42 – Caracterização estrutural de carboidratos por 2D-RMN. Regiões anoméricas de carboidratos ( $\delta_C/\delta_H$  89-110 / 3,7-5,8) dos espectros de 2D-RMN HSQC 2D (em DMSO- $d_6$ ) de paredes celulares das linhagens de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) em *Saccharum* spp. cultivar híbrido SP-3280. As estruturas são coloridas para coincidir com os contornos atribuídos nos espectros de NMR. As atribuições dos sinais de carboidratos estão listadas na Tabela 12.

Tabela 12 - Atribuições dos sinais de correlação para os carbonos anoméricos nos espectros 2D HSQC dos diferentes carboidratos identificados presentes em linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) de *Saccharum* spp. cultivar híbrido SP-3280.

Label	$\delta_C/\delta_H$	Assignment
$\alpha X_{1(R)}$	92.1/4.88	$\alpha$ -D-xylopyranoside (R) [ $\alpha$ -D-glucopyranoside (R)]
$\beta X_{1(R)}$	97.2/4.23	$\beta$ -D-xylopyranoside (R) [ $\beta$ -D-glucopyranoside (R)]
U <sub>1</sub>	97.2/5.15	4-O-methyl- $\alpha$ -D-glucuronic acid
X' <sub>1</sub>	99.3/4.49	2-O-acetyl- $\beta$ -D-xylopyranoside
X <sub>1</sub> /X' <sub>1</sub>	101.5/4.26	$\beta$ -D-xylopyranoside + 3-O-acetyl- $\beta$ -D-xylopyranoside
Gl <sub>1</sub>	102.8/4.23	(1→4)- $\beta$ -D-glucopyranoside
Gl <sub>1</sub>	103.1/4.17	(1→3)- $\beta$ -D-glucopyranoside + (1→6)- $\beta$ -D-glucopyranoside
Ar <sub>1</sub>	104.4/5.41	$\alpha$ -L-arabinofuranoside
Ar <sub>1</sub>	106.9/5.33	$\alpha$ -L-arabinofuranoside
Ar <sub>1</sub>	107.7/4.76	$\alpha$ -L-arabinofuranoside

**Expressão genica de TFs: *ShMYB85* e *ShMYB58/63* e genes representativos da rota biossintética da lignina em linhagens transgênicas de *Shmyb85*, *Shmyb58/63* e controles (EV e WT) *Saccharum* spp.**

**Expressão de Fatores de Transcrição: *ShMYB85* e *ShMYB58/63***

Nas figuras 43A e 43C se observa que as linhagens 15, 16, 17 e 8 de *Shmyb85* apresentaram uma redução significativa na expressão de *ShMYB85* de -0.67, -0.82, -0.84 e -0.78 vezes, respectivamente ( $p \leq 0.05$ , 0.0051, 0.0167, 0.0025, 0.0025). Para o caso de *Shmyb58/63* (Figuras 43B e 43D), nas linhagens 1, 11, 12 e 16 houve redução de -0.27, -0.22, -0.05, -0.08 vezes, respectivamente ( $p \geq 0.05$ ). Particularmente na linhagem 17 houve um aumento significativo em destaque de 1.35 vezes ( $p \leq 0.05$ , 0.038)

**Expressão de genes representativos da rota biossintética da lignina**

As figuras 44A e 44B apresentam o perfil de expressão relativa para os genes *ShCAD8*, *ShCAD2*, *ShCCoAOMT2*, *ShCOMT*, *ShF5H* e *ShCCoAOMT1* em linhagens de *Shmyb85* e *Shmyb58/63* e controles (EV e WT), onde se observa uma tendência

marcante na redução de expressão destes genes. Assim, primeiramente para *Shmyb85* (Figura 44A) em *ShCAD8* (Figura 44A1) houve redução significativa da expressão nas linhagens 14, 15 e 16 (-0.68, -0.55 e -0.75 vezes, respectivamente,  $p \leq 0.05$ , 0.0167, 0.0025, 0.0025). Para *ShCAD2* (Figura 44A2), *ShCCoAOMT1* (Figura 44A6) e *ShF5H* (Figura 44A5) houve reduções significativas para todas as linhagens, oscilando entre -0.48 e -0.80; -0.29 e -0.67; -0.76 e -0.92, respectivamente ( $p \leq 0.05$ ). *ShCCoAOMT2* (Figura 44A3) não experimentou queda significativa da expressão. *ShCOMT* (Figura 44A4, apresentou redução significativa na maioria das linhagens (queda que oscilou entre -0.01 e -0.64, vezes,  $p \leq 0.05$ ). Particularmente é importante informar que na linhagem 8 os genes *ShCAD8*, *ShCCoAOMT2*, *ShCCoAOMT1* e *ShCOMT* apresentaram aumento de 0.8, 4.8, 6.6 e 1.3 vezes, respectivamente ( $p \leq 0.05$ , 0.0025 para todos os genes). No que diz respeito a *Shmyb58/63*, o gene *ShCAD8* (Figura 44B1) apresentou queda significativa da expressão na maioria das linhagens (a exceção da linhagem 1), oscilando entre -0.66 e -0.74 vezes ( $p \leq 0.05$ ). Em *ShCAD2* (Figura 44B2) houve queda significativa apenas na linhagem 1 ( $p \leq 0.05$ , 0.017). *ShCCoAOMT2* (Figura 44B3) experimentou alteração no teor de expressão, evidenciando um aumento só na linhagem 1, em 10.24 vezes ( $p \leq 0.05$ , 0.017). Para *ShCCoAOMT1* (Figura 44B6) houve redução nas linhagens 11 e 16, oscilando entre -0.48 e -0.61 vezes, respectivamente ( $p \leq 0.05$ , 0.012 e 0.017). O teor de expressão de *ShCOMT* (Figura 44B4) foi reduzido nas linhagens 1, 11, 12, e 16, oscilando entre -0.70 e -0.92 vezes ( $p \leq 0.05$ ). Finalmente, em três linhagens transgênicas (11, 12 e 16) de *Shmyb58/63* houve redução significativa de *ShF5H* (Figura 44B5), que oscilou entre -0,60 e -0.77 vezes ( $p \leq 0.05$ ).

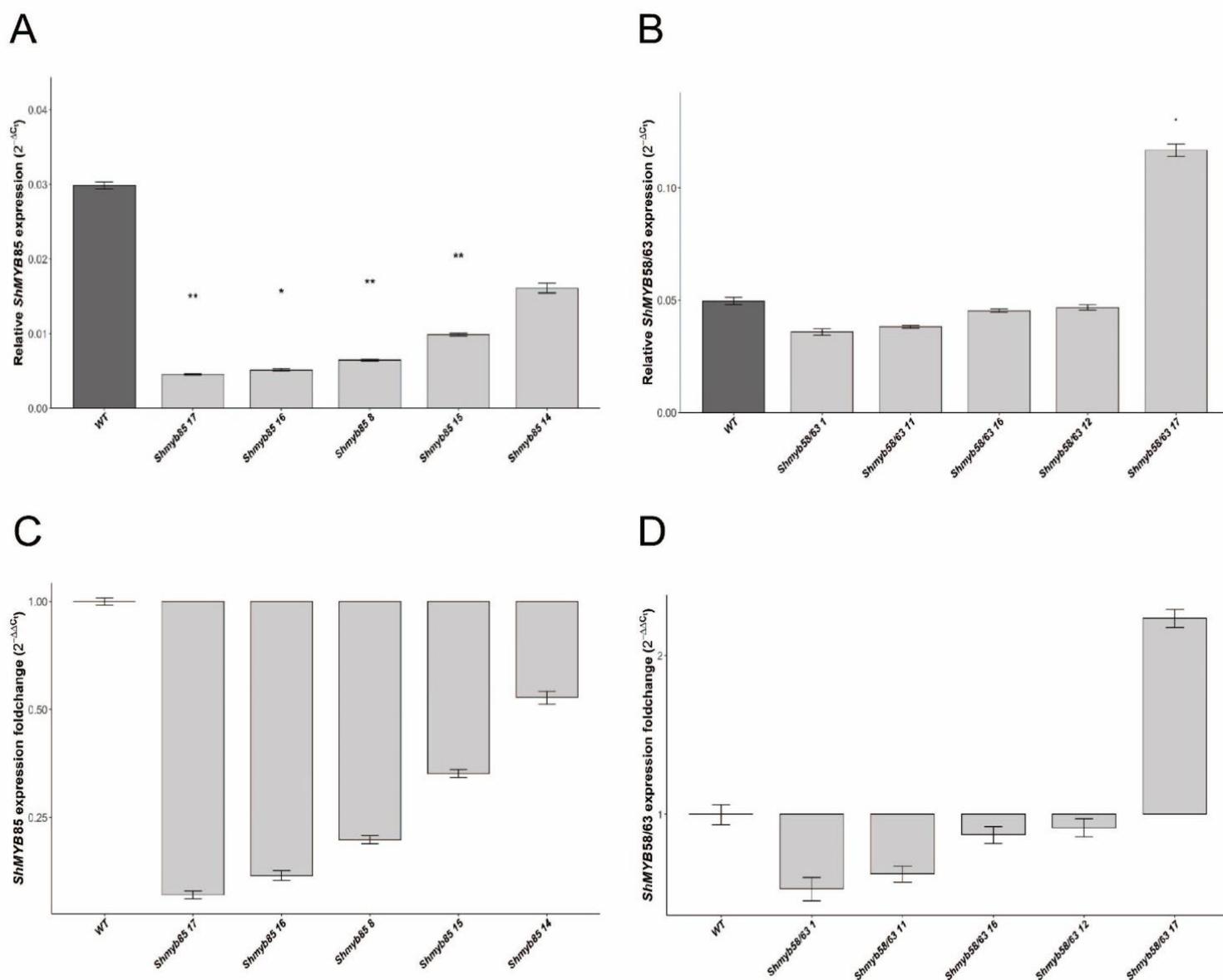


Figura 43 – Expressão relativa (RT-qPCR) e fold-change de genes TFs (A e C): *ShMYB85* e (B e D) *ShMYB58/63*, respectivamente no 5° internódio de linhagens de transgênicas de *Shmyb85* e *Shmyb58/63* *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas na expressão relativa dos genes entre os tratamentos. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das médias de quintuplicatas biológicas independentes.

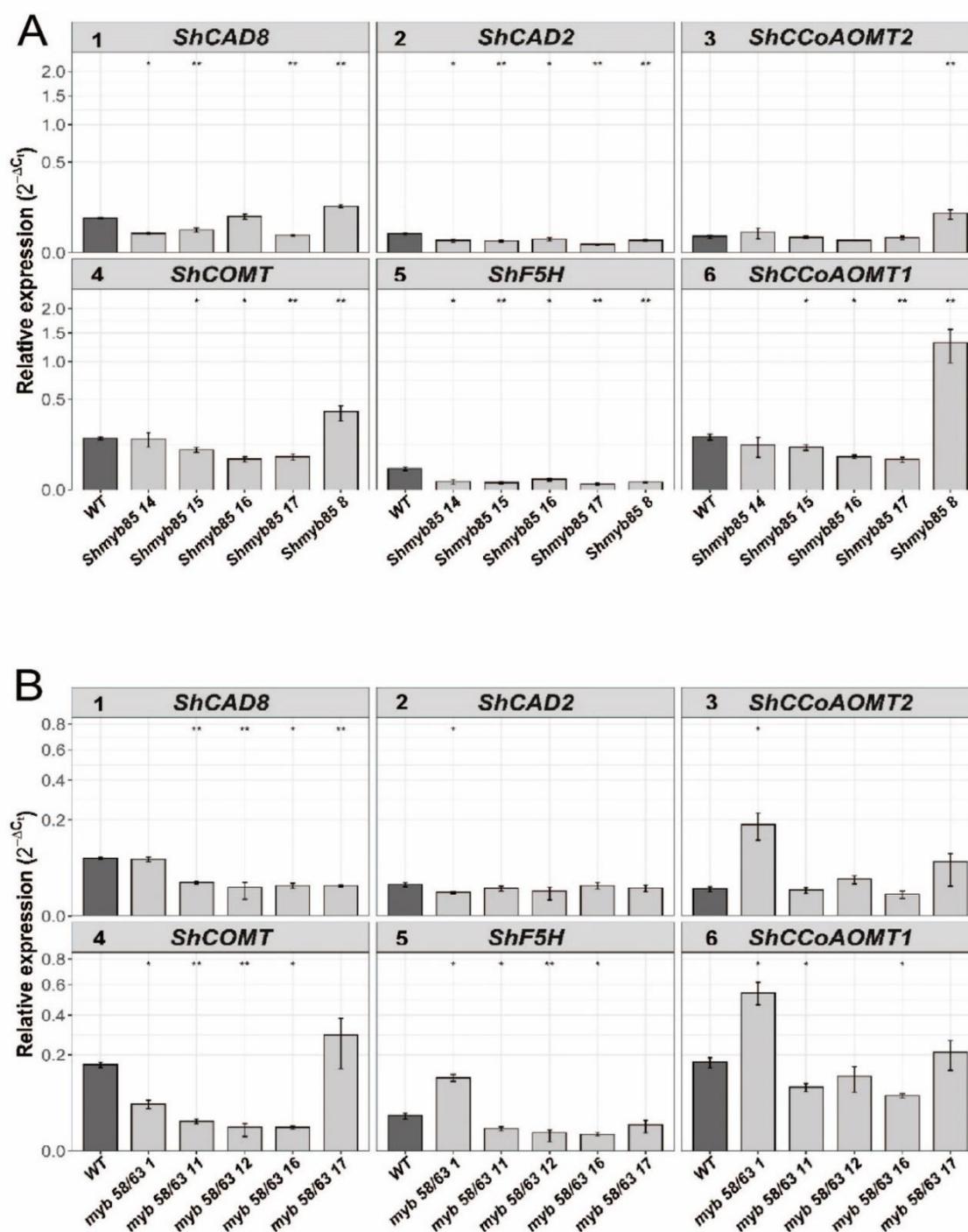


Figura 44 – Expressão relativa (RT-qPCR) de genes representativos da rota biossintética de lignina, respectivamente, no 5<sup>o</sup> internódio de linhagens de transgênicas de *Shmyb85* (A) e *Shmyb58/63* (B) *Saccharum* spp. cultivar híbrido SP-3280. Asteriscos representam diferenças significativas na expressão relativa de um gene particular entre os tratamentos. As médias foram comparados pelo teste de Wilcoxon,  $p \leq 0.05$ . As barras verticais indicam o desvio padrão das médias de quintuplicatas biológicas independentes.

### 3. Discussão

Plantas com metabolismo C4, especialmente aquelas pertencentes à subfamília Panicoideae como: cana-de-açúcar, *Sorghum*, *Miscanthus* e *Panicum* emergem como potenciais culturas bioenergéticas devido ao seu alto potencial de produção de biomassa e fixação de carbono (van der Weijde et al., 2013). A parede celular vegetal é um recurso abundante (40% à 80% da biomassa é vegetal) e promissor para a produção de etanol 2G (Pauly & Keegstra, 2008; Welker et al., 2015). Uma das principais dificuldades para a conversão dos polissacáridos em açúcares fermentáveis é a presença da lignina, devido principalmente a sua propriedade de resistir à degradação e à diversidade de ligações de baixa reatividade, fazendo que este polímero fenólico seja o principal responsável pela recalcitrância da parede celular secundária (SCW) (Chen & Dixon, 2007; Van Acker et al., 2013; Wilkerson et al., 2014). Portanto uma compreensão dos mecanismos reguladores subjacentes à biossíntese e deposição de lignina é essencial para o desenvolvimento de novos genótipos de gramíneas energia para o aproveitamento de bioetanol lignocelulósico.

Muitos fatores de transcrição envolvidos diretamente ou indiretamente na regulação da biossíntese da lignina tem sido identificados e caracterizados, principalmente em eudicotiledôneas modelo como *Arabidopsis*, *Populus* e *Nicotiana* (Gray et al., 2012; McCann & Carpita, 2008; Vogel, 2008; Zhao & Dixon, 2011), porém a informação no que respeita a gramíneas é ainda limitada (Bhatia et al., 2019) e muito pouco ou quase não compreendida em cana-de-açúcar. (Brito et al., 2015; Martins et al., 2018; Poovaiah et al., 2016). Apesar da distinta composição de SCW que permite inferir um distinto controle transcricional da lignificação entre eudicotiledôneas e monocotiledôneas (Handakumbura & Hazen, 2012), TFs ortólogos de eudicotiledôneas regulando a SCW foram identificados no genoma de gramíneas e o perfil de expressão tem ajudado a identificar possíveis reguladores candidatos implicados na deposição do SCW neste tipo de plantas. Uma das primeiras tentativas ainda usadas e o desenvolvimento de sistemas heterólogos para avaliar a função de um grupo de TFs em arroz, milho e outras gramíneas na regulação da deposição da SCW (Cesarino et al., 2016). Todos estes estudos em sistemas heterólogos tem mostrado informações importantes, mas é necessário mais identificações e caracterizações de TFs *in planta* em gramíneas para entender a regulação de forma específica na deposição da lignina, visando a produção e o desenvolvimento do E2G.

O presente trabalho teve como objetivo principal, identificar e caracterizar fatores de transcrição que regulem enzimas da biossíntese da lignina em *Saccharum* spp. visando a alteração e modulação da deposição de lignina na SCW.

Preliminarmente a análise filogenética revelou que ShMYB85 (Sh\_250M06\_t000170) e ShMYB58/63 (Sh\_227K15\_t000030) são potenciais ortólogos das sequências identificadas e ou caracterizadas dos TFs ativadores específicos de lignina: AtMYB85 (Zhong et al., 2008), PtrMYB92 (Zhong et al., 2011), OsMYB42/85 (Hirano et al., 2013), PvMYB42/85B (Rao et al. 2019), ZmMYB152 (Zhang et al., 2016; Bhatia et al., 2019) e AtMYB58/63 (Zhou et al., 2009), OsMYB58/63 (Hirano et al., 2013; Noda et al., 2015), PvMYB58/63C (Rao et al., 2019), SbMYB60 (Scully et al., 2016), respectivamente (Figura 9, Cap 1). As sequências aminoácidas de ShMYB85 e ShMYB58/63 indicam que são TFs R2R3-MYB com motivos R2 e R3 conservados (Figura S2). Além disso o conjunto de informações relacionadas à correlação de Pearson (Figura 12, Cap1), a análise de componentes principal (PCA) (Figura 13, Cap1) de redes bayesianas (Figura 14, Cap 1), o ensaio de transativação de protoplastos (Figura 17, Cap1) e o experimento da indução da injúria em folhas (Figuras 39 e 40), acrescentou o nosso conhecimento sobre a regulação da deposição de lignina nas espécies de *Saccharum* por parte de ShMYB85 e ShMYB58/63, o que os converte como potenciais candidatos para modificação genética e no desenvolvimento de novos genótipos em cana-de-açúcar para o aproveitamento de bioetanol lignocelulósico. Para corroborar esta hipótese se ShMYB85 e ShMYB58/63 regulam a deposição de lignina e/ou da SCW *in planta*, nós desenvolvemos linhagens transgênicas silenciadas para estes TFs (*Shmyb85* e *Shmyb58/63*) em *Saccharum* spp. cultivar híbrido SP-3280 enfatizando as oportunidades e desafios de usar estes TFs MYBs como ferramenta de engenharia genética.

Destacamos que *Shmyb85* e *Shmyb58/63* tiveram similitudes e diferenças no que respeita a mudanças no metabolismo de lignina relacionados ao conteúdo e composição de lignina, associados a testes histoquímicos e perfil de expressão de genes representativos da biossíntese deste polímero. Um dos aspectos a assinalar como se observa nas Figuras 30C-D e 43C-D, é que as linhagens de *Shmyb85* apresentaram redução significativa no conteúdo de lignina acetylbromide e da expressão de *ShMYB85* em 18.7% e 0.67-0.84 vezes, respectivamente, no entanto as linhagens de *Shmyb58/63* apresentaram redução não significativa da expressão

de *ShMYB58/63* e significativa do conteúdo de lignina acetylbromide em 0.05-0.27 vezes e 15.3%, respectivamente. A diferença da redução do conteúdo de lignina acetylbromide também se observou nas análises histoquímicas com o teste de Weisner (floroglucinol-HCl). Em comparação com *Shmyb58/63*, nas distintas linhagens de *Shmyb85* as córtex dos entrenós seccionados transversalmente têm menor coloração nos feixes vasculares e as paredes dos elementos celulares como a hipoderme, epiderme, esclerênquima e as fibras vasculares aparecem mais finos e menos lignificados, contribuindo significativamente com o menor conteúdo deste polímero (Figuras 31 e 32). Reforçando a presença de fenótipo nestes transgênicos, a análise 2D-RMN das linhagens elites dos transgênicos (Tabela 11A-11B) revelou queda diferencial do conteúdo de lignina em ambos transgênicos. Devemos indicar que apesar da redução da expressão de *ShMYB58/63* em *Shmyb58/63*, não foi significativa, houve em várias linhagens deste transgênico apresentando fenótipo (queda de lignina acetylbromide e lignina por dosagem 2D-RMN). Os únicos estudos apresentando silenciamento para *MYB85* e *MYB58/63* são aqueles em *O. sativa* (Hirano et al., 2013) e *P. virgatum* (Rao et al. 2019) até o momento. Similar ao descrito em *O. sativa* onde o silenciamento de *OsMYB58/63* não provocou fenótipo (Hirano et al., 2013), em *P. virgatum* silenciando *PvMYB42/85A* e *PvMYB58/63B* (com redução de transcritos destes TFs), não foi observado fenótipo traduzido em conteúdo de lignina acetylbromide, mas com redução dos genes da biossíntese deste polímero (Rao et al., 2019). O acontecido em *Shmyb58/63* em termos de redução não significativa da expressão de *ShMYB58/63*, mas com evidente fenótipo hipoteticamente pode ser devido principalmente a: 1) Uma leve supressão, consequência de uma baixa frequência de inserção do transgene aleatória no genoma complexo de *Saccharum* spp. cultivar híbrido SP-3280 (vetor contendo o RNAi para silenciamento de *ShMYB58/63*) e/ou, 2) redundância de genes parálogos de *ShMYB58/63* não identificados neste estudo, desde que o genoma de cana-de-açúcar até o momento devido a sua complexidade não está sequenciado completamente. O sequenciamento completo do genoma e banco de dados mais robustos em cana-de-açúcar pode lançar luz sobre estas hipóteses.

Os genes representativos da rota biossintética da lignina relacionados ao conteúdo e composição (aqueles que governam o braço da biossíntese da Lignina S e G que influem na razão S/G) avaliados aqui (Figuras 44A-B) apresentaram de forma geral,

redução no teor do seus transcritos, tanto *Shmyb85* como para *Shmyb58/63*, demonstrando com isto que existe uma associação de *ShMYB85* e *ShMYB58/63* com a regulação da rota da biossíntese de lignina. Um dado que é de destacar é a redução da expressão de *ShF5H* observada para *Shmyb85* e *Shmyb58/63*, resultado que foi coerente com o observado no ensaio de transativação de protoplastos (Figura 17, Cap1) executado com os parentais de cana-de-açúcar, onde *ShMYB85* e *ShMYB58/63* impulsionou a ativação dos promotores de *ShF5H*, mais especificadamente em *S. officinarum*. Tem sido descrito que em *Arabidopsis* e *Medicago truncatula*, *MYB58/63* aparece a ativar todos os genes da rota da biossíntese da lignina com exceção de *F5H* desde que o promotor desse gene nessas plantas carecem de motivos AC (Zhao & Dixon, 2011). De outro lado em *Arabidopsis*, a abundância de transcritos de *F5H* é influenciado positivamente por *SND1* e *MYB103* (Öhman et al., 2013). *EgMYB2* um TF homólogo de *AtMYB46/83* induz a expressão de *F5H* (Goicoechea et al., 2005) e o repressor em *Z. mays* (*ZmMYB42*) aparece a regular negativamente a *F5H*, indicando que o promotor deste gene possui motifs AC-like (Fornalé et al., 2010). A super-expresssão de *PvMYB4* de *P. virgatum* em tabaco e em *P. virgatum* provocou em ambas plantas a redução de expressão de *F5H* (Shen et al., 2012). Portanto a regulação da expressão de *F5H* é um caso particular no qual cada espécie de plantas pode ter diferentes mecanismos transcricionais, indicando que os parentais de cana-de-açúcar e os híbridos podem regular a biossíntese da lignina de uma forma espécie-específica. Dados adicionais que respaldam a regulação de *ShMYB85* e *ShMYB58/63* na expressão *F5H*, neste item, são aqueles obtidos por nós aqui no que respeita à correlação de Pearson entre a expressão destes TFs e *F5H*, onde *ShMYB58/63* e *ShMYB85* tiveram um coeficientes de 0.93, e 0.97 e respectivamente em *S. officinarum* em quanto os mesmos TFs tiveram coeficientes de 0.92, e 0.8 em *S. spontaneum* (Figura 12, Cap1) Além disso ultimamente o único trabalho até o momento similar a nossos resultados obtidos na transativação e o silenciamento foi o reportado em *P. virgatum* onde *PvMYB42/85A* e *PvMYB58/63* transativaram os promotores de *COMT* e *F5H* (Rao et al., 2019).

De outro lado, *F5H* e *COMT*, são considerados como pontos críticos e limitantes na biossíntese da lignina S (Chen & Dixon, 2007; Li et al., 2003; Ralph et al., 2006) influenciando na razão S/G que é uma das características importantes para prever o grau e natureza da condensação do polímero e, conseqüentemente, sobre a recalcitrância

da biomassa vegetal (Kiyota et al., 2012), no processo de digestibilidade e/ou o rendimento da sacarificação (Takeda et al., 2017), desde que ligninas S são mais susceptíveis à hidrólise do que lignina G (Ferrer et al., 2008) como demonstrado em mutantes *comt* e *f5h* por silenciamento e super-expresssão em espécies tais como *Arabidopsis* (Meyer et al., 1998), *P. virgatum* (Wu et al., 2019), *S. bicolor* (Tetreault et al., 2020), *O. sativa* (Takeda et al., 2017), e *Populus* (Huntley et al., 2003). A digestão da biomassa (internódio 5°) dos transgênicos com NaOH para avaliar o coeficiente S/G, permitiu evidenciar que as linhagens de *Shmyb58/63* apresentaram aumento significativo da razão S/G (29.18%-69.68%) do que aquelas de *Shmyb85* onde não se observa redução significativa (Figuras 30 E-F). Quando analisando as linhagens elites de ambos transgênicos nos resultados de espectroscopia 2D-NMR estes confirmam o aumento da razão S/G para *Shmyb58/63* (linhagens: 17 e 16), enquanto para *Shmyb85* só a linhagem 15 apresentou queda significativa deste parâmetro. De outro lado *Shmyb85* apresentou maior teor de unidades G do que *Shmyb58/63* (Tabelas 11A-11B). Os valores da razão S/G apresentados aqui para os transgênicos se relacionam também com os teores dos transcritos observados em *ShF5H*, desde que a redução em *Shmyb85* foi maior (0.92 vezes) do que *Shmyb58/63* (0.77 vezes) enquanto a queda da expressão dos genes que governam a biossíntese de lignina G: *ShCCoAOMT1* (*CCoAOMT A*) foi similar para *Shmyb85* (0.67) como para *Shmyb58/63* (0.61). *ShCCoAOMT2* (*CCoAOMT B*) interessantemente não apresentou queda para nenhum dos transgênicos (Figuras 44A e B). A análise histoquímica de Maüle também revela a razão S/G diferencial entre os transgênicos desde que na região periférica, os feixes vasculares, as fibras ao redor destes e as células parenquimática nas linhagens transgênicas de *Shmyb85* (Figuras. 31 K-S) apresentam um predomínio de coloração amarelada, indicando lignina G. Enquanto *Shmyb58/63* os feixes vasculares da região periférica apresentam fibras com deposição de lignina diferenciada, desde que as camadas de fibras, mas internas dos feixes apresentam coloração amarela lignina G, enquanto as mais externas coloração avermelhada lignina S, indicando também que as células parenquimáticas desta região tem coloração vermelha Lignina S (Figuras 32S, 32K, 32M, 32O e 32Q).

Como mencionado acima, as diferentes proporções dos monômeros, definindo a composição de lignina e em consequência a razão S/G afetam a estrutura física da lignina e portanto o rendimento de sacarificação (Takeda et al., 2017). Em coerência

com isto, na Figura 35B, se observa que as linhagens de *Shmyb58/63*: 1 e 16 que experimentaram aumento significativo da razão S/G tiveram um rendimento percentual da sacarificação de 89.13 e 101.79 respectivamente (É importante deixar em claro que outros fatores tais como a queda de lignina podem ter contribuído naquelas linhagens no que respeita à sacarificação). Em sentido contrário, esta premissa não tem lugar em *Shmyb85* desde que não houve nenhuma relação entre a razão S/G e o percentual de sacarificação. De outro lado parece lógico afirmar que a diminuição do conteúdo de lignina e a configuração anatômica de linhagens celulares com SCW menos espessas e com menor deposição de lignina observada nos testes histoquímicos seriam os únicos fatores determinantes nos percentuais de sacarificação notáveis para *Shmyb85* (Figura 35A) (Wilson and Hatfield, 1997; Chen et al., 2002; Grabber et al., 1992). Porém a acessibilidade pelo coquetéis fúngicos aos polissacarídeos estruturais da parede celular é muito limitada também por compostos fenólicos, aromáticos e particularmente por *p*-ácidos hidroxicinâmicos (ácidos ferúlicos(FA) e ácido *p*-coumárico(*p*CA)) (Zhang et al., 2019). Uma importante característica em gramíneas em comparação com as eudicotiledôneas é as quantidades significativas destes *p*-ácidos hidroxicinâmicos (Molinari et al., 2013; Hatfield et al., 2017) Ambos hidroxicinamatos, especialmente o ferulato, permitem o entrecruzamento entre lignina e hemiceluloses, enquanto que *p*-coumarato principalmente acila as unidades S e a sua acumulação está associada com a lignificação e com a maturidade da parede celular, sendo considerado como um bom marcador bioquímico para prever a deposição de lignina (Zhang et al., 2019; Hatfield et al., 2017). Assim, avaliando o conteúdo de fenóis totais por Folin-Ciocalteu, usando como padrão ácido clorogênico (desde que é um dos fenólicos mais abundantes em cana-de-açúcar (Colombo et al., 2005)), *Shmyb85* apresentou nas linhagens 16 e 17 um aumento de 41.54% e 45.55%, respectivamente, enquanto que *Shmyb58/63* não apresentou mudanças (Figuras 36 A-B). De outro lado observando o perfil fenólico, em destaque na linhagem 16 de *Shmyb85* houve aumento significativo (547.32 %) de ácido clorogênico e redução (-70.66) de ácido coumárico, enquanto no que respeita ao ácido quinico houve aumento nas linhagens, mas só foi significativa esta alteração em *Shmyb85* 17 (125.82%). *Shmyb58/63* só apresentou redução significativa no ácido clorogênico sendo significativa esta variação nas linhagens 12,1,17 e 11 (-85.49, -71.75, -35.87, -59.54, respectivamente) (Tabelas 9A-9B). Nesse sentido observamos que em *Shmyb85* houve aumento de teor de compostos fenólicos, hipotetizando que

estas alterações provavelmente se deva a que um subconjunto de genes da rota dos fenilpropanóides sejam super-expressos (como resposta a um reajuste devido ao silenciamento), aumentando o fluxo através de somente alguns passos da rota e levando à acumulação de intermediários, mas não a uma maior biossíntese de monolignóis ou à incorporação na lignina (Scully et al., 2016), favorecendo ao rendimento em sacarificação. Em sentido contrário em *Shmyb58/63* a redução de ácido clorogênico evidenciaria que derivados deste ácido cinâmico se estariam incorporando na rota biossintética dos monolignóis, favorecendo a lignificação (Silva et al., 2019) afetando a sacarificação. Adicionalmente a análise 2D-RMN, evidenciou que em *Shmyb85 15* e *Shmyb85 14* houve redução de tricina e de *p*-coumarato em todas as linhagens transgênicas elite de *Shmyb85* quando comparando com *Shmyb58/63* onde não houve alterações (Tabelas 11A-11B). Assim, tomando em consideração todas as observações mencionadas acima se explicaria em parte a diferença em rendimento de sacarificação a favor de *Shmyb85*.

No que respeita aos efeitos do silenciamento de *ShMYB85* e *ShMYB58/63* sobre o conteúdo dos polissacarídeos da SCW e outros compostos associados (Figuras 34 e 37), foi observado que ambos transgênicos apresentaram redução do conteúdo de celulose, sendo que em destaque *Shmyb85* apresentou a maior alteração. *Shmyb85* especificadamente apresentou redução no conteúdo de pectinas em sentido contrário ao observado em *Shmyb58/63* onde não houve alterações. O conteúdo de hemiceluloses observado não permite concluir um padrão de alteração por efeito do silenciamento desde que só a *Shmyb85 15* apresentou redução deste polímero enquanto em *Shmyb58/63* só a linhagem 12 apresentou aumento. Além disso, avaliando os carboidratos presentes nas paredes celulares dos transgênicos por análise 2D-NMR, os principais sinais corresponderam aos arabinoxilanos (uns dos componentes principais das hemicelulosas), mas não apresentaram diferenças entre os transgênicos (Figura 42). Em destaque ambos transgênicos apresentaram redução no conteúdo de flavonoides, enquanto só *Shmyb85* experimentou queda nas antocianinas (Figura 37). Zhong et al. (2008) descreveu que o ativador específico de lignina AtMYB85 através de um ensaio de protoplastos para corroborar o seu papel na ativação de genes dos componentes SCW, impulsionou a ativação do promotor do genes *4CL1* da rota biossintética de lignina (Boerjan et al., 2003) mas não para aqueles de *CesA8* (celulose) (Taylor et al., 2004) e *IRX9* (hemicelulose) (Peña et al.,

2007). De forma similar em *Arabidopsis*, AtMYB58/63 ativaram especificadamente genes da rota biossintética de lignina (Zhou et al., 2009), apresentando os promotores destes genes motivos AC, mas não para que aqueles relacionados a celulose (Suzuki et al., 2006; Taylor et al., 2004) e xilanos (hemiceluloses) (Bromley et al., 2013; Rennie & Scheller, 2014). Em monocotiledôneas: Noda et al. (2015) demonstrou usando transativação de protoplastos que em *O. sativa*: OsMYB58/63 e OsMYB42/85 (ortólogo de AtMYB85) são ativadores transcricional que regulam a biossíntese de celulose e lignina por impulsionar a ativação do promotor da *OsCesA7* (Tanaka et al., 2003) e *OsCAldOMT1* (Koshiba et al., 2013). Em *S. bicolor*, a super-expresssão de *SbMYB60* (ortólogo de AtMYB58/63) alterou os conteúdos de lignina, celulose e xilanos (Scully et al., 2016). De outro lado a super-expresssão de *ZmMYB167* (ortólogo de AtMYB85) em *B. distachyon* e *Z. mays* não provocou alterações nos polissacarídeos, mas induziu a biossíntese de lignina e compostos fenólicos associados à SCW e recentemente Rao et al. (2019), descreveu que em internódios de *P. virgatum* a super-expresssão de *PvMYB58/63A* aumentou a expressão de genes relacionados à biossíntese de celulose e hemicelulose tais como: celulose sintase, xilano sintase e da biossíntese de flavonóides. Em sentido contrário a superexpressão de *PvMYB85A* diminuiu a biossíntese de genes da biossíntese da celulose relacionados à parede celular primária, secundária e hemiceluloses, não houve alterações no que respeita aos genes da biossíntese de flavonóides. Tanto *PvMYB58/63A* como *PvMYB85A* aumentou o conteúdo de lignina e genes associados ao polímero. Depois de mencionar os diferentes estudos acima e apesar de que tem sido demonstrado que a função regulatória da SCW entre eudicotiledôneas e gramíneas é conservada por parte de alguns TFs é evidente que existe diferenças na regulação transcricional da SCW entre *Arabidopsis* e monocotiledôneas incluindo o híbrido de cana-de-açúcar aqui silenciado, para MYB85 e MYB58/63, sendo que nas monocotiledôneas estes TFs são amplos reguladores da SCW, hipotetizando que isto pode ser devido a uma diferente composição dos promotores dos distintos genes envolvidos nas diferentes rotas biossintéticas da SCW (Noda et al., 2015), desde que mudanças na composição dos motivos cis regulatórios tem sido considerados como os principais contribuintes para a evolução de caracteres específicos entre distintos tipos de plantas (Yokoyama et al., 2014), entre espécies de um mesmo gênero e monocotiledôneas próximas filogeneticamente. Por exemplo, *Z. mays* e o seu progenitor teosinte apresentaram variação na composição do elementos cis em muitos

promotores entre genes ortólogos (Lemmon et al., 2014). De outro lado os padrões de ligação de um TF em particular ao promotor de um gene alvo e a auto-regulação (ligação ao seu próprio promotor) tem vindo mostrar a ser específicos da espécie indicando uma potencial sub-funcionalização explicando a divergência funcional da regulação entre *S. bicolor*, *Z. mays*, *O. sativa* (Agarwal et al. 2016) e *Arabidopsis*, tal como o exemplo que foi mencionado acima para o caso de MYB58/63 que pode regular diferencialmente a deposição dos polímeros segundo a espécie (Noda et al., 2015).

#### 4. CONCLUSÃO

Aqui neste estudo, nós informamos sobre a caracterização funcional do silenciamento dos homólogos identificados como ativadores específicos de lignina: ShMYB85 e ShMYB58/63 em *Saccharum* spp. cultivar híbrido SP-3280. Apesar da baixa similaridade em termos de sequências entre ShMYB85, ShMYB58/63 e o seus homólogos caracterizados em *Arabidopsis*, a regulação da SCW por parte dos TFs estudados aqui em cana-de-açúcar e o descrito para *Arabidopsis* foi de certa forma conservada embora tenha sido encontradas diferenças fenotípicas particulares para os transgênicos gerados para cana-de-açúcar. De outro lado entre os transgênicos *Shmyb85* e *Shmyb58/63* houve similitudes e diferenças no que respeita a mudanças no metabolismo de lignina, relacionados ao conteúdo e composição de lignina, associados a testes histoquímicos e perfil de expressão de genes representativos da biossíntese deste polímero. O silenciamento de ambos os ativadores em cana-de-açúcar, também teve diferenciais implicações na composição da polissacarídeos da parede celular, assim como também de compostos relacionado a lignina tais como: fenóis, flavonoides e antocianinas. Uma análise 2D-RMN das linhagens elite dos transgênicos reforçou estruturalmente as mudanças na deposição de lignina e polissacarídeos encontradas por efeito do silenciamento de *ShMYB85* e *ShMYB58/63*. Portanto, aqui enfatizamos as oportunidades e desafios de usar estes TFs MYBs como ferramenta de engenharia genética, para redesenhar a SCW com fins de bioenergia, sendo que o conhecimento gerado aqui é vital para ajudar a impulsionar as práticas tradicionais de melhoramento em cana-de-açúcar e abordagens biotecnológicas para a produção de biomassa lignocelulósica e o aproveitamento desta para o desenvolvimento do etanol de segunda geração (E2G).

## 5. REFERENCIAS

1. Agarwal, T., Grotewold, E., Doseff, A. I., & Gray, J. (2016). MYB31/MYB42 syntelogs exhibit divergent regulation of phenylpropanoid genes in maize, sorghum and rice. *Scientific Reports*, *6*(1), 1–17.
2. Aharoni, A., Dixit, S., Jetter, R., Thoenes, E., Van Arkel, G., & Pereira, A. (2004). The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in *Arabidopsis*. *The Plant Cell*, *16*(9), 2463–2480.
3. Altschul, S., & Madden, T. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids ...*, *25*(17), 3389–3402. <http://nar.oxfordjournals.org/content/25/17/3389.short>
4. Altschul, S F, Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402.
5. Altschul, Stephen F, Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410.
6. Ambavaram, M. M. R., Krishnan, A., Trijatmiko, K. R., & Pereira, A. (2011). Coordinated activation of cellulose and repression of lignin biosynthesis pathways in rice. *Plant Physiology*, *155*(2), 916–931.
7. Anderson, W. F., & Akin, D. E. (2008). Structural and chemical properties of grass lignocelluloses related to conversion for biofuels. *Journal of Industrial Microbiology and Biotechnology*, *35*, 355–366.
8. Araújo, P., Tolentino, F. T., Cesarino, I., Gallinari, R. H., Steenackers, W., Mayer, J. L. S., & Mazzafera, P. (2019). Expression of *Eucalyptus globulus* LACCASE48 Restores Lignin Content of *Arabidopsis thaliana* lac17 Mutant. *Plant Molecular Biology Reporter*, *37*(5), 488–498.
9. Bewg, W. P. (2015). *Investigation of lignin biosynthesis in sugarcane for improved lignocellulosic ethanol production*. Queensland University of Technology.
10. Bewg, W. P., Poovaiah, C., Lan, W., Ralph, J., & Coleman, H. D. (2016). RNAi downregulation of three key lignin genes in sugarcane improves glucose release without reduction in sugar production. *Biotechnology for Biofuels*, *9*(1), 270.
11. Bhatia, R., Dalton, S., Roberts, L. A., Moron-Garcia, O. M., Iacono, R., Kosik, O., Gallagher, J. A., & Bosch, M. (2019). Modified expression of ZmMYB167 in

- Brachypodium distachyon* and *Zea mays* leads to increased cell wall lignin and phenolic content. *Scientific Reports*, 9(1), 1–12.
12. Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin biosynthesis. *Annual Review of Plant Biology*, 54, 519–546.
  13. Bomal, C., Bedon, F., Caron, S., Mansfield, S. D., Levasseur, C., Cooke, J. E. K., Blais, S., Tremblay, L., Morency, M.-J., Pavy, N., & others. (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(14), 3925–3939.
  14. Bonawitz, N. D., & Chapple, C. (2010). The genetics of lignin biosynthesis: connecting genotype to phenotype. *Annual Review of Genetics*, 44, 337–363.
  15. Bottcher, A., Cesarino, I., Santos, A. B. dos, Vicentini, R., Mayer, J. L. S., Vanholme, R., Morreel, K., Goeminne, G., Moura, J. C. M. S., Nobile, P. M., Carmello-Guerreiro, S. M., Anjos, I. A. dos, Creste, S., Boerjan, W., Landell, M. G. de A., & Mazzafera, P. (2013). Lignification in sugarcane: biochemical characterization, gene discovery, and expression analysis in two genotypes contrasting for lignin content. *Plant Physiology*, 163(4), 1539–1557.
  16. Bouvier d'Yvoire, M., Bouchabke-Coussa, O., Voorend, W., Antelme, S., Cézard, L., Legée, F., Lebris, P., Legay, S., Whitehead, C., McQueen-Mason, S. J., & others. (2013). Disrupting the cinnamyl alcohol dehydrogenase 1 gene (BdCAD1) leads to altered lignification and improved saccharification in *Brachypodium distachyon*. *The Plant Journal*, 73(3), 496–508.
  17. Brito, M. S., Nobile, P. M., Bottcher, A., dos Santos, A. B., Creste, S., de Landell, M. G. A., Vincentz, M., Vicentini, R., & Mazzafera, P. (2015). Expression profile of sugarcane transcription factor genes involved in lignin biosynthesis. *Tropical Plant Biology*, 8(1–2), 19–30.
  18. Bromley, J. R., Busse-Wicher, M., Tryfona, T., Mortimer, J. C., Zhang, Z., Brown, D. M., & Dupree, P. (2013). GUX 1 and GUX 2 glucuronyltransferases decorate distinct domains of glucuronoxytan with different substitution patterns. *The Plant Journal*, 74(3), 423–434.
  19. Brown, L. C., & Mac Berthouex, P. (2002). *Statistics for environmental engineers*. CRC press.
  20. Brown, L., & Torget, R. (1996). NREL analytical procedure: LAP009 enzymatic saccharification of lignocellulosic biomass hydrolysis. *Golden, CO: National*

*Renewable Energy Laboratory.*

21. Buckeridge, Marcos S, Santos, W. D. dos, & Sousa, A. P. de. (2010). As rotas para o etanol celulósico no Brasil. *Bioetanol Da Cana-de-Açúcar: P&D Para Produtividade e Sustentabilidade*, 365–380.
22. Buckeridge, Marcos Silveira, Cavalari, A. A., & Silva, G. B. (2008). Parede celular. *Fisiologia Vegetal*, 165–181.
23. Burnquist, W. L., Sorrells, M. E., & Tanksley, S. (1992). Characterization of genetic variability in *Saccharum* germplasm by means of restriction fragment length polymorphism (RFLP) analysis. *Proc. Int. Soc. Sugar Cane Technol.*, 21, 355–365.
24. Carroll, A., & Somerville, C. (2009). Cellulosic Biofuels. *Annual Review of Plant Biology*, 60(1), 165–182.
25. Carvalho-Netto, O. V, Bressiani, J. A., Soriano, H. L., Fiori, C. S., Santos, J. M., Barbosa, G. V. S., Xavier, M. A., Landell, M. G. A., & Pereira, G. A. G. (2014). The potential of the energy cane as the main biomass crop for the cellulosic industry. *Chemical and Biological Technologies in Agriculture*, 1(1), 1–8.
26. Cassan-Wang, H., Goué, N., Saidi, M. N., Legay, S., Sivadon, P., Goffner, D., & Grima-Pettenati, J. (2013). Identification of novel transcription factors regulating secondary cell wall formation in *Arabidopsis*. *Frontiers in Plant Science*, 4, 189.
27. Casu, R. E., Jarmey, J. M., Bonnett, G. D., & 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 & Integrative Genomics*, 7(2), 153–167.
28. Cesarino, I., Araújo, P., Domingues Júnior, A. P., & Mazzafera, P. (2012). An overview of lignin metabolism and its effect on biomass recalcitrance. *Brazilian Journal of Botany*, 35(4), 303–311.
29. Cesarino, I., Araújo, P., Sampaio Mayer, J. L., Vicentini, R., Berthet, S., Demedts, B., Vanholme, B., Boerjan, W., & Mazzafera, P. (2013). Expression of SofLAC, a new laccase in sugarcane, restores lignin content but not S:G ratio of *Arabidopsis* lac17 mutant. *Journal of Experimental Botany*, 64(6), 1769–1781.
30. Cesarino, I., Simões, M. S., Brito, M. dos S., Fanelli, A., Silva, T. da F., Romanel, E., & others. (2016). Building the wall: recent advances in understanding lignin metabolism in grasses. *Acta Physiologiae Plantarum*, 38(11).
31. Chabannes, M., Barakate, A., Lapierre, C., Marita, J. M., Ralph, J., Pean, M., Danoun, S., Halpin, C., Grima-Pettenati, J., & Boudet, A. M. (2001). Strong

- decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of *cinnamoyl CoA reductase (CCR)* and *cinnamyl alcohol dehydrogenase (CAD)* in tobacco plants. *The Plant Journal*, 28(3), 257–270.
32. Chang, S., Puryear, J., & Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter*, 11(2), 113–116.
33. Cheavegatti-Gianotto, A., de Abreu, H. M. C., Arruda, P., Bespalhok Filho, J. C., Burnquist, W. L., Creste, S., di Ciero, L., Ferro, J. A., de Oliveira Figueira, A. V., de Sousa Filgueiras, T., Grossi-de-Sá, M. de F., Guzzo, E. C., Hoffmann, H. P., de Andrade Landell, M. G., Macedo, N., Matsuoka, S., de Castro Reinach, F., Romano, E., da Silva, W. J., ... César Ulian, E. (2011). Sugarcane (*Saccharum X officinarum*): A Reference Study for the Regulation of Genetically Modified Cultivars in Brazil. *Tropical Plant Biology*, 4(1), 62–89.
34. Chen, F., & Dixon, R. A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology*, 25, 759–761.
35. Chen, Fang, & Dixon, R. A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology*, 25(7), 759–761.
36. Chen, L., Auh, C., Chen, F., Cheng, X., Aljoe, H., Dixon, R. A., & Wang, Z. (2002). Lignin deposition and associated changes in anatomy, enzyme activity, gene expression, and ruminal degradability in stems of tall fescue at different developmental stages. *Journal of Agricultural and Food Chemistry*, 50(20), 5558–5565.
37. Chow, C.-N., Lee, T.-Y., Hung, Y.-C., Li, G.-Z., Tseng, K.-C., Liu, Y.-H., Kuo, P.-L., Zheng, H.-Q., & Chang, W.-C. (2019). PlantPAN3. 0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Research*, 47(D1), D1155--D1163.
38. Colombo, R., Yariwake, J. H., Queiroz, E. F., Ndjoko, K., & Hostettmann, K. (2005). On-line identification of sugarcane (*Saccharum officinarum* L.) methoxyflavones by liquid chromatography--UV detection using post-column derivatization and liquid chromatography--mass spectrometry. *Journal of Chromatography A*, 1082(1), 51–59.
39. D'Hont, A., Paulet, F., & Glaszmann, J. C. (2002). Oligoclonal interspecific origin of “North Indian” and “Chinese” sugarcane. *Chromosome Research*, 10(3), 253–262.

40. De Oliveira, M. E. D., Vaughan, B. E., & Rykiel, E. J. (2005). Ethanol as Fuel: Energy, Carbon Dioxide Balances, and Ecological Footprint. *BioScience*, *55*(7), 593.
41. de Souza, A. P., Grandis, A., Leite, D. C. C., & Buckeridge, M. S. (2014). Sugarcane as a bioenergy source: history, performance, and perspectives for second-generation bioethanol. *BioEnergy Research*, *7*(1), 24–35.
42. del Río, J. C., Lino, A. G., Colodette, J. L., Lima, C. F., Gutiérrez, A., Martínez, Á. T., Lu, F., Ralph, J., & Rencoret, J. (2015). Differences in the chemical structure of the lignins from sugarcane bagasse and straw. *Biomass and Bioenergy*, *81*, 322–338.
43. dos Santos, A. B., Bottcher, A., Vicentini, R., Mayer, J. L. S., Kiyota, E., Landell, M. A. G., Creste, S., & Mazzafera, P. (2015). Lignin biosynthesis in sugarcane is affected by low temperature. *Environmental and Experimental Botany*, *120*, 31–42.
44. DuBois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., & Smith, F. (1956). Colorimetric Method for Determination of Sugars and Related Substances. *Analytical Chemistry*, *28*(3), 350–356.
45. e Silva, N. V., Mazzafera, P., & Cesarino, I. (2019). Should I stay or should I go: are chlorogenic acids mobilized towards lignin biosynthesis? *Phytochemistry*, *166*, 112063.
46. Eloy, N., Voorend, W., Lan, W., Saleme, M. de L. S., Cesarino, I., Vanholme, R., Smith, R. A., Goeminne, G., Pallidis, A., Morreel, K., & others. (2016). Silencing chalcone synthase impedes the incorporation of tricetin in lignin and increases lignin content. *Plant Physiology*, pp--01108.
47. Eulgem, T., Rushton, P. J., Robatzek, S., & Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends in Plant Science*, *5*(5), 199–206.
48. Farrell, A. E., Plevin, R. J., Turner, B. T., Jones, A. D., O'Hare, M., & Kammen, D. M. (2006). Ethanol can contribute to energy and environmental goals. *Science (New York, N.Y.)*, *311*(5760), 506–508.
49. Feng, J.-X., Liu, D., Pan, Y., Gong, W., Ma, L.-G., Luo, J.-C., Deng, X. W., & Zhu, Y.-X. (2005). An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the Arabidopsis AP2/EREBP transcription factor gene family. *Plant Molecular Biology*, *59*(6), 853–868.

50. Ferreira, S. S., Hotta, C. T., de Carli Poelking, V. G., Leite, D. C. C., Buckeridge, M. S., Loureiro, M. E., Barbosa, M. H. P., Carneiro, M. S., & Souza, G. M. (2016). Co-expression network analysis reveals transcription factors associated to cell wall biosynthesis in sugarcane. *Plant Molecular Biology*, *91*(1–2), 15–35.
51. Ferrer, J.-L., Austin, M. B., Stewart, C., & Noel, J. P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry*, *46*(3), 356–370.
52. Fornalé, S., Capellades, M., Encina, A., Wang, K., Irar, S., Lapierre, C., Ruel, K., Joseleau, J.-P., Berenguer, J., Puigdomènech, P., & others. (2012). Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for *cinnamyl alcohol dehydrogenase*. *Molecular Plant*, *5*(4), 817–830.
53. Fornalé, S., Lopez, E., Salazar-Henao, J. E., Fernández-Nohales, P., Rigau, J., & Caparros-Ruiz, D. (2014). AtMYB7, a new player in the regulation of UV-sunscreens in *Arabidopsis thaliana*. *Plant and Cell Physiology*, *55*(3), 507–516.
54. Fornalé, S., Shi, X., Chai, C., Encina, A., Irar, S., Capellades, M., Fuguet, E., Torres, J.-L., Rovira, P., Puigdomenech, P., & others. (2010). ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *The Plant Journal*, *64*(4), 633–644.
55. Fornalé, S., Sonbol, F.-M., Maes, T., Capellades, M., Puigdomènech, P., Rigau, J., & Caparros-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*(6), 809–823.
56. Franke, R., McMichael, C. M., Meyer, K., Shirley, A. M., Cusumano, J. C., & Chapple, C. (2000). Modified lignin in tobacco and poplar plants over-expressing the *Arabidopsis* gene encoding *ferulate 5-hydroxylase*. *The Plant Journal*, *22*(3), 223–234.
57. Fu, C., Mielenz, J. R., Xiao, X., Ge, Y., Hamilton, C. Y., Rodriguez, M., Chen, F., Foston, M., Ragauskas, A., Bouton, J., & others. (2011). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proceedings of the National Academy of Sciences*, *108*(9), 3803–3808.
58. Gallego-Giraldo, L., Shadle, G., Shen, H., Barros-Rios, J., Fresquet Corrales, S., Wang, H., & Dixon, R. A. (2016). Combining enhanced biomass density with reduced lignin level for improved forage quality. *Plant Biotechnology Journal*,

- 14(3), 895–904.
59. Garsmeur, O., Droc, G., Antonise, R., Grimwood, J., Potier, B., Aitken, K., Jenkins, J., Martin, G., Charron, C., Hervouet, C., & others. (2018). A mosaic monoploid reference sequence for the highly complex genome of sugarcane. *Nature Communications*, 9(1), 1–10.
  60. Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., Lapierre, C., Pollet, B., Verhaegen, D., Chaubet-Gigot, N., & others. (2005). EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *The Plant Journal*, 43(4), 553–567.
  61. Goldemberg, J. (2007). Ethanol for a sustainable energy future. *Science (New York, N.Y.)*, 315(5813), 808–810.
  62. Grabber, J. H. (2005). How do lignin composition, structure, and cross-linking affect degradability? A review of cell wall model studies. *Crop Science*, 45(3), 820–831.
  63. Grabber, J. H., Jung, G. A., Abrams, S. M., & Howard, D. B. (1992). Digestion kinetics of parenchyma and sclerenchyma cell walls isolated from orchardgrass and switchgrass. *Crop Science*, 32(3), 806–810.
  64. Grabber, J. H., Ralph, J., Lapierre, C., & Barrière, Y. (2004). Genetic and molecular basis of grass cell-wall degradability. I. Lignin--cell wall matrix interactions. *Comptes Rendus Biologies*, 327(5), 455–465.
  65. Gray, J., Caparrós-Ruiz, D., & Grotewold, E. (2012). Grass phenylpropanoids: regulate before using! *Plant Science*, 184, 112–120.
  66. Grima-Pettenati, J., Soler, M., Camargo, E. L. O., & Wang, H. (2012). Chapter 6 - Transcriptional Regulation of the Lignin Biosynthetic Pathway Revisited: New Players and Insights. In L. Jouanin & C. Lapierre (Eds.), *Lignins Biosynthesis, Biodegradation and Bioengineering* (Vol. 61, pp. 173–218). Academic Press. <https://doi.org/http://dx.doi.org/10.1016/B978-0-12-416023-1.00006-9>
  67. Grivet, L., Glaszmann, J. C., & D'Hont, A. (2006). Molecular evidence of sugarcane evolution and domestication. *Darwin's Harvest: New Approaches to the Origins, Evolution and Conservation of Crops*. Columbia University Press, New York, 49–66.
  68. Guidelli, G. V, Mattiello, L., Gallinari, R. H., Lucca, P. C. de, & Menossi, M. (2018). pGVG: a new Gateway-compatible vector for transformation of sugarcane and

- other monocot crops. *Genetics and Molecular Biology*, 41(2), 450–454.
69. Gunasekara, C., Subramanian, A., Avvari, J. V. R. K., Li, B., Chen, S., & Wei, H. (2016). ExactSearch: a web-based plant motif search tool. *Plant Methods*, 12(1), 1–4.
70. Ha, C. M., Escamilla-Trevino, L., Serrani Yarce, J. C., Kim, H., Ralph, J., Chen, F., & Dixon, R. A. (2016). An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *The Plant Journal*.
71. Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.
72. Handakumbura, P. P., & Hazen, S. P. (2012). Transcriptional regulation of grass secondary cell wall biosynthesis: playing catch-up with *Arabidopsis thaliana*. *Frontiers in Plant Science*, 3, 74.
73. Hatfield, R. D., Rancour, D. M., & Marita, J. M. (2017). Grass cell walls: a story of cross-linking. *Frontiers in Plant Science*, 7, 2056.
74. Hatfield, R., & Fukushima, R. S. (2005). Can lignin be accurately measured? *Crop Science*, 45(3), 832–839.
75. He, J.-B., Zhao, X.-H., Du, P.-Z., Zeng, W., Beahan, C. T., Wang, Y.-Q., Li, H.-L., Bacic, A., & Wu, A.-M. (2018). KNAT7 positively regulates xylan biosynthesis by directly activating *IRX9* expression in *Arabidopsis*. *Journal of Integrative Plant Biology*, 60(6), 514–528.
76. Hirano, K., Kondo, M., Aya, K., Miyao, A., Sato, Y., Antonio, B. A., Namiki, N., Nagamura, Y., & Matsuoka, M. (2013). Identification of transcription factors involved in rice secondary cell wall formation. *Plant and Cell Physiology*, 54(11), 1791–1802.
77. Hisano, H., Nandakumar, R., & Wang, Z.-Y. (2009). Genetic modification of lignin biosynthesis for improved biofuel production. *In Vitro Cellular & Developmental Biology - Plant*, 45(3), 306–313.
78. Hoang, N. V., Furtado, A., Botha, F. C., Simmons, B. A., & Henry, R. J. (2015). Potential for genetic improvement of sugarcane as a source of biomass for biofuels. *Frontiers in Bioengineering and Biotechnology*, 3, 182.
79. Huang, D., Wang, S., Zhang, B., Shang-Guan, K., Shi, Y., Zhang, D., Liu, X., Wu, K., Xu, Z., Fu, X., & others. (2015). A gibberellin-mediated DELLA-NAC signaling cascade regulates cellulose synthesis in rice. *The Plant Cell*, 27(6), 1681–1696.

80. Huntley, S. K., Ellis, D., Gilbert, M., Chapple, C., & 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*(21), 6178–6183.
81. Jackson, L. A., Shadle, G. L., Zhou, R., Nakashima, J., Chen, F., & Dixon, R. A. (2008). Improving saccharification efficiency of alfalfa stems through modification of the terminal stages of monolignol biosynthesis. *Bioenergy Research*, *1*(3–4), 180–192.
82. Jacobsen, K. R., Fisher, D. G., Maretzki, A., & Moore, P. H. (1992). Developmental changes in the anatomy of the sugarcane stem in relation to phloem unloading and sucrose storage. *Plant Biology*, *105*(1), 70–80.
83. Jiang, W., Zeng, Q., Jiang, Y., Gai, Y., & Jiang, X. (2020). Molecular and functional characterization of *ferulate-5-hydroxylase* in *Populus tomentosa*. *Journal of Plant Biochemistry and Biotechnology*, 1–7.
84. Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B., & Martin, C. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *The EMBO Journal*, *19*(22), 6150–6161.
85. Jin, J., Tian, F., Yang, D.-C., Meng, Y.-Q., Kong, L., Luo, J., & Gao, G. (2016). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research*, gkw982.
86. Johansen, D. A. (1940). *Plant microtechnique*. McGraw-Hill Publishing Company, Ltd., London.
87. Jung, H.-J. G., Samac, D. A., & Sarath, G. (2012). Modifying crops to increase cell wall digestibility. *Plant Science*, *185*, 65–77.
88. Jung, H. G., & Casler, M. D. (2006). Maize Stem Tissues: Cell Wall Concentration and Composition during Development. *Crop Science*, *46*, 1793–1800.
89. Jung, H. G., & Engels, F. M. (2002). Alfalfa stem tissues. *Crop Science*, *42*(2), 524–534.
90. Jung, J. H., Fouad, W. M., Vermerris, W., Gallo, M., & Altpeter, F. (2012). RNAi suppression of lignin biosynthesis in sugarcane reduces recalcitrance for biofuel production from lignocellulosic biomass. *Plant Biotechnology Journal*, *10*(9), 1067–1076.
91. Kawaoka, A., & Ebinuma, H. (2001). Transcriptional control of lignin biosynthesis

- by tobacco LIM protein. *Phytochemistry*, 57(7), 1149–1157.
92. Kawaoka, A., Kaothien, P., Yoshida, K., Endo, S., Yamada, K., & Ebinuma, H. (2000). Functional analysis of tobacco LIM protein Ntlm1 involved in lignin biosynthesis. *The Plant Journal*, 22(4), 289–301.
93. Kim, W.-C., Kim, J.-Y., Ko, J.-H., Kim, J., & Han, K.-H. (2013). Transcription factor MYB46 is an obligate component of the transcriptional regulatory complex for functional expression of secondary wall-associated cellulose synthases in *Arabidopsis thaliana*. *Journal of Plant Physiology*, 170(15), 1374–1378.
94. Kinkema, M., Geijskes, J., Delucca, P., Palupe, A., Shand, K., Coleman, H. D., Brinin, A., Williams, B., Sainz, M., & Dale, J. L. (2014). Improved molecular tools for sugar cane biotechnology. *Plant Molecular Biology*, 84(4–5), 497–508.
95. Kiyota, E., Mazzafera, P., & 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(16), 7015–7020.
96. Ko, J.-H., Kim, W.-C., & Han, K.-H. (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in *Arabidopsis*. *The Plant Journal*, 60(4), 649–665.
97. Koshiba, T., Hirose, N., Mukai, M., Yamamura, M., Hattori, T., Suzuki, S., Sakamoto, M., & Umezawa, T. (2013). Characterization of 5-hydroxyconiferaldehyde O-methyltransferase in *Oryza sativa*. *Plant Biotechnology*, 13–219.
98. Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., Mimura, T., Fukuda, H., & Demura, T. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes & Development*, 19(16), 1855–1860.
99. Lan, W., Lu, F., Regner, M., Zhu, Y., Rencoret, J., Ralph, S. A., Zakai, U. I., Morreel, K., Boerjan, W., & Ralph, J. (2015). Tricin, a flavonoid monomer in monocot lignification. *Plant Physiology*, 167(4), 1284–1295.
100. Lapierre, C., Pilate, G., Pollet, B., Mila, I., Leple, J.-C., Jouanin, L., Kim, H., & Ralph, J. (2004). Signatures of *cinnamyl alcohol dehydrogenase* deficiency in poplar lignins. *Phytochemistry*, 65(3), 313–321.
101. Lemmon, Z. H., Bukowski, R., Sun, Q., & Doebley, J. F. (2014). The role of cis regulatory evolution in maize domestication. *PLoS Genetics*, 10(11), e1004745.

102. Lennartsson, P. R., Erlandsson, P., & Taherzadeh, M. J. (2014). Integration of the first and second generation bioethanol processes and the importance of by-products. *Bioresource Technology*, *165*, 3–8. <https://doi.org/10.1016/j.biortech.2014.01.127>
103. Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., de Peer, Y., Rouzé, P., & Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, *30*(1), 325–327.
104. Li, E., Bhargava, A., Qiang, W., Friedmann, M. C., Forneris, N., Savidge, R. A., Johnson, L. A., Mansfield, S. D., Ellis, B. E., & Douglas, C. J. (2012). The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in *Populus*. *New Phytologist*, *194*(1), 102–115.
105. Li, X., Weng, J.-K., & Chapple, C. (2008). Improvement of biomass through lignin modification. *The Plant Journal: For Cell and Molecular Biology*, *54*(4), 569–581.
106. Li, Y., Kajita, S., Kawai, S., Katayama, Y., & Morohoshi, N. (2003). Down-regulation of an anionic peroxidase in transgenic aspen and its effect on lignin characteristics. *Journal of Plant Research*, *116*(3), 175–182.
107. Lin, C.-S., Hsu, C.-T., Yang, L.-H., Lee, L.-Y., Fu, J.-Y., Cheng, Q.-W., Wu, F.-H., Hsiao, H. C.-W., Zhang, Y., Zhang, R., & others. (2018). Application of protoplast technology to CRISPR/Cas9 mutagenesis: from single-cell mutation detection to mutant plant regeneration. *Plant Biotechnology Journal*, *16*(7), 1295–1310.
108. Liu, Y., Wei, M., Hou, C., Lu, T., Liu, L., Wei, H., Cheng, Y., & Wei, Z. (2017). Functional characterization of *Populus* PsnSHN2 in coordinated regulation of secondary wall components in tobacco. *Scientific Reports*, *7*(1), 1–11.
109. Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta Ct}$  Method. *Methods (San Diego, Calif.)*, *25*(4), 402–408.
110. Llerena, J. P. P., Figueiredo, R., dos Santos Brito, M., Kiyota, E., Mayer, J. L. S., Araujo, P., Schimpl, F. C., Dama, M., Pauly, M., & Mazzafera, P. (2019). Deposition of lignin in four species of *Saccharum*. *Scientific Reports*, *9*(1), 1–19.
111. Martins, A. P. B., dos Santos Brito, M., Mayer, J. L. S., Llerena, J. P. P., Oliveira, J. F., Takahashi, N. G., Carlin, S. D., Borges, D. N. A. F., Andrade, L. M., Peixoto-Júnior, R. F., & others. (2018). Ectopic expression of sugarcane *SHINE* changes cell wall and improves biomass in rice. *Biomass and Bioenergy*, *119*, 322–334.

112. Mäule, C. (1901). *Das verhalten verholzter membranen gegen kaliumpermanganat, eine holzreaktion neuer art*. A. Zimmer's Verlag (Ernst Mohrmann).
113. McCann, M. C., & Carpita, N. C. (2008). Designing the deconstruction of plant cell walls. *Current Opinion in Plant Biology*, 11(3), 314–320.
114. McCarthy, R. L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., Spicer, C., & Ye, Z.-H. (2010). The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. *Plant and Cell Physiology*, 51(6), 1084–1090.
115. McCarthy, R. L., Zhong, R., & Ye, Z.-H. (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(11), 1950–1964.
116. Meyer, K., Shirley, A. M., Cusumano, J. C., Bell-Lelong, D. A., & Chapple, C. (1998). Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, 95(12), 6619–6623.
117. Ming, R., Moore, P. H., Wu, K. K., D'hont, A., Glaszmann, J. C., Tew, T. L., Mirkov, T. E., da Silva, J., Jifon, J., Rai, M., Schnell, R. J., Brumbley, S. M., Lakshmanan, P., Comstock, J. C., & Paterson, A. H. (2006). Sugarcane Improvement through Breeding and Biotechnology. In *Plant Breeding Reviews* (pp. 15–118).
118. Mintz-Oron, S., Mandel, T., Rogachev, I., Feldberg, L., Lotan, O., Yativ, M., Wang, Z., Jetter, R., Venger, I., Adato, A., & others. (2008). Gene expression and metabolism in tomato fruit surface tissues. *Plant Physiology*, 147(2), 823–851.
119. Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., & Ohme-Takagi, M. (2007). NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. *The Plant Cell*, 19(1), 270–280.
120. Mitsuda, N., Seki, M., Shinozaki, K., & Ohme-Takagi, M. (2005). The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. *The Plant Cell*, 17(11), 2993–3006.
121. Molinari, H. B. C., Pellny, T. K., Freeman, J., Shewry, P. R., & Mitchell, R. A. C. (2013). Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in *Brachypodium distachyon*. *Frontiers in Plant Science*, 4.

122. Moore, P. H. (1995). Temporal and spatial regulation of sucrose accumulation in the sugarcane stem. *Functional Plant Biology*, 22(4), 661–679.
123. Moura, J. C. M. S., Bonine, C. A. V., de Oliveira Fernandes Viana, J., Dornelas, M. C., & Mazzafera, P. (2010). Abiotic and biotic stresses and changes in the lignin content and composition in plants. *Journal of Integrative Plant Biology*, 52(4), 360–376.
124. Naidu, K. M., & Sreenivasan, T. V. (1987). Conservation of Sugarcane germplasm. *Copersugar Int. Sugarcane Breeding Workshop*, 3–52.
125. Newman, L. J., Perazza, D. E., Juda, L., & Campbell, M. M. (2004). Involvement of the R2R3-MYB, AtMYB61, in the ectopic lignification and dark-photomorphogenic components of the det3 mutant phenotype. *The Plant Journal*, 37(2), 239–250.
126. Noda, S., Koshiba, T., Hattori, T., Yamaguchi, M., Suzuki, S., & Umezawa, T. (2015). The expression of a rice secondary wall-specific cellulose synthase gene, OsCesA7, is directly regulated by a rice transcription factor, OsMYB58/63. *Planta*, 242(3), 589–600.
127. Ohashi-Ito, K., Oda, Y., & Fukuda, H. (2010). Arabidopsis VASCULAR-RELATED NAC-DOMAIN6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. *The Plant Cell*, 22(10), 3461–3473.
128. Öhman, D., Demedts, B., Kumar, M., Gerber, L., Gorzsás, A., Goeminne, G., Hedenström, M., Ellis, B., Boerjan, W., & Sundberg, B. (2013). MYB 103 is required for *FERULATE-5-HYDROXYLASE* expression and syringyl lignin biosynthesis in Arabidopsis stems. *The Plant Journal*, 73(1), 63–76.
129. Owczarzy, R., Tataurov, A. V., Wu, Y., Manthey, J. A., McQuisten, K. A., Almabrazi, H. G., Pedersen, K. F., Lin, Y., Garretson, J., McEntaggart, N. O., & others. (2008). IDT SciTools: a suite for analysis and design of nucleic acid oligomers. *Nucleic Acids Research*, 36(suppl\_2), W163-W169.
130. Papini-Terzi, F. S., Rocha, F. R., Vêncio, R. Z. N., Oliveira, K. C., Felix, J. de M., Vicentini, R., Rocha, C. de S., Simões, A. C. Q., Ulian, E. C., di Mauro, S. M. Z., da Silva, A. M., Pereira, C. A. de B., Menossi, M., & Souza, G. M. (2005). Transcription profiling of signal transduction-related genes in sugarcane tissues. *DNA Research: An International Journal for Rapid Publication of Reports on Genes and Genomes*, 12(1), 27–38.

131. Pauly, M., & Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant Journal*, *54*(4), 559–568.
132. Peña, M. J., Zhong, R., Zhou, G.-K., Richardson, E. A., O'Neill, M. A., Davill, A. G., York, W. S., & Ye, Z.-H. (2007). *Arabidopsis* irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *The Plant Cell*, *19*(2), 549–563.
133. Poovaiah, C. R., Bewg, W. P., Lan, W., Ralph, J., & Coleman, H. D. (2016). Sugarcane transgenics expressing MYB transcription factors show improved glucose release. *Biotechnology for Biofuels*, *9*(1), 1.
134. Poovaiah, C. R., Nageswara-Rao, M., Soneji, J. R., Baxter, H. L., & Stewart, C. N. (2014). Altered lignin biosynthesis using biotechnology to improve lignocellulosic biofuel feedstocks. *Plant Biotechnology Journal*, *12*(9), 1163–1173.
135. Porto, B. N., Magalhães, P. C., Campos, N. A., Alves, J. D., & Magalhães, M. M. (2011). Otimização de protocolos de extração de rna em diferentes tecidos de milho. *Revista Brasileira de Milho e Sorgo*, *9*(2), 189–200.
136. Preston, J., Wheeler, J., Heazlewood, J., Li, S. F., & Parish, R. W. (2004). AtMYB32 is required for normal pollen development in *Arabidopsis thaliana*. *The Plant Journal*, *40*(6), 979–995.
137. Price, S. (1963). Cytogenetics of modern sugar canes. *Economic Botany*, *17*(2), 97–106.
138. Qin, W., Yin, Q., Chen, J., Zhao, X., Yue, F., He, J., Yang, L., Liu, L., Zeng, Q., Lu, F., & others. (2020). The class II KNOX transcription factors KNAT3 and KNAT7 synergistically regulate monolignol biosynthesis in *Arabidopsis*. *Journal of Experimental Botany*, *71*(18), 5469–5483.
139. Raes, J., Rohde, A., Christensen, J. H., de Peer, Y., & Boerjan, W. (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiology*, *133*(3), 1051–1071.
140. Ralph, J., Akiyama, T., Kim, H., Lu, F., Schatz, P. F., Marita, J. M., Ralph, S. A., Reddy, M. S. S., Chen, F., & Dixon, R. A. (2006). Effects of *coumarate 3-hydroxylase* down-regulation on lignin structure. *The Journal of Biological Chemistry*, *281*(13), 8843–8853.
141. Ralph, J., Grabber, J. H., & Hatfield, R. D. (1995). Lignin-ferulate cross-links in grasses: active incorporation of ferulate polysaccharide esters into ryegrass lignins. *Carbohydrate Research*, *275*(1), 167–178.

142. Ralph, J., Lapierre, C., Marita, J. M., Kim, H., Lu, F., Hatfield, R. D., Ralph, S., Chapple, C., Franke, R., Hemm, M. R., & others. (2001). Elucidation of new structures in lignins of CAD-and COMT-deficient plants by NMR. *Phytochemistry*, 57(6), 993–1003.
143. Ramos, R. L. B., Tovar, F. J., Junqueira, R. M., Lino, F. B., & Sachetto-Martins, G. (2001). Sugarcane expressed sequences tags (ESTs) encoding enzymes involved in lignin biosynthesis pathways. *Genetics and Molecular Biology*, 24(1–4), 235–241.
144. Rao, X., Chen, X., Shen, H., Ma, Q., Li, G., Tang, Y., Pena, M., York, W., Frazier, T. P., Lenaghan, S., & others. (2019). Gene regulatory networks for lignin biosynthesis in switchgrass (*Panicum virgatum*). *Plant Biotechnology Journal*, 17(3), 580–593.
145. Rao, X., & Dixon, R. A. (2018). Current models for transcriptional regulation of secondary cell wall biosynthesis in grasses. *Frontiers in Plant Science*, 9, 399.
146. Rass-Hansen, J., Falsig, H., Jørgensen, B., & Christensen, C. H. (2007). Bioethanol: fuel or feedstock? *Journal of Chemical Technology & Biotechnology*, 82(4), 329–333.
147. Reddy, M. S. S., Chen, F., Shadle, G., Jackson, L., Aljoe, H., & 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(46), 16573–16578.
148. Ren, R. (2012). global status report 2012. Renewable energy policy network for the 21st century. In *REN 21 Secretariat*. <http://www.ren21.net/gsr>
149. Rennie, E. A., & Scheller, H. V. (2014). Xylan biosynthesis. *Current Opinion in Biotechnology*, 26, 100–107.
150. Rezende, C. A., de Lima, M. A., Maziero, P., Ribeiro deAzevedo, E., Garcia, W., & Polikarpov, I. (2011). Chemical and morphological characterization of sugarcane bagasse submitted to a delignification process for enhanced enzymatic digestibility. *Biotechnology for Biofuels*, 4(1), 54.
151. Riaño-Pachón, D. M., & Mattiello, L. (2017). Draft genome sequencing of the sugarcane hybrid SP80-3280. *F1000Research*, 6.
152. Roach, B. T. (1978). Utilization of *Saccharum spontaneum* in sugarcane breeding. *Proc. Int. Soc. Suagr Cane Technol.*, 16, 43–58.
153. Roach, B. T., & Daniels, J. (1987). The *Saccharum* Complex and the genus

- Saccharum. Copersugar Int. Sugarcane Breeding Workshop*, 1–33.
154. Romano, J. M., Dubos, C., Prouse, M. B., Wilkins, O., Hong, H., Poole, M., Kang, K.-Y., Li, E., Douglas, C. J., Western, T. L., & others. (2012). AtMYB61, an R2R3-MYB transcription factor, functions as a pleiotropic regulator via a small gene network. *New Phytologist*, 195(4), 774–786.
  155. Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics methods and protocols* (pp. 365–386). Springer.
  156. Saathoff, A. J., Sarath, G., Chow, E. K., Dien, B. S., & 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(1), e16416.
  157. Santos, A. B. dos, Bottcher, A., Kiyota, E., Mayer, J. L. S., Vicentini, R., Brito, M. dos S., Creste, S., Landell, M. G. A., & Mazzafera, P. (2015). Water Stress Alters Lignin Content and Related Gene Expression in Two Sugarcane Genotypes. *Journal of Agricultural and Food Chemistry*, 63(19), 4708–4720.
  158. Scully, E. D., Gries, T., Sarath, G., Palmer, N. A., Baird, L., Serapiglia, M. J., Dien, B. S., Boateng, A. A., Ge, Z., Funnell-Harris, D. L., & others. (2016). Overexpression of *SbMyb60* impacts phenylpropanoid biosynthesis and alters secondary cell wall composition in *Sorghum bicolor*. *The Plant Journal*, 85(3), 378–395.
  159. Shen, H., Fu, C., Xiao, X., Ray, T., Tang, Y., Wang, Z., & Chen, F. (2009). Developmental control of lignification in stems of lowland switchgrass variety Alamo and the effects on saccharification efficiency. *BioEnergy Research*, 2(4), 233–245.
  160. Shen, H., He, X., Poovaiah, C. R., Wuddineh, W. A., Ma, J., Mann, D. G. J., Wang, H., Jackson, L., Tang, Y., Neal Stewart, C., & others. (2012). Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytologist*, 193(1), 121–136.
  161. Sibout, R., Baucher, M., Gattineau, M., Van Doorselaere, J., Mila, I., Pollet, B., Maba, B., Pilate, G., Lapierre, C., Boerjan, W., & others. (2002). Expression of a poplar cDNA encoding a *ferulate-5-hydroxylase/coniferaldehyde 5-hydroxylase* increases S lignin deposition in *Arabidopsis thaliana*. *Plant Physiology and*

- Biochemistry*, 40(12), 1087–1096.
162. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., & others. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), 539.
  163. Simmons, B. A., Loque, D., & Blanch, H. W. (2008). Next-generation biomass feedstocks for biofuel production. *Genome Biology*, 9(12), 242.
  164. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L., & Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 27(3), 431–432.
  165. Soccol, C. R., Vandenberghe, L. P. de S., Medeiros, A. B. P., Karp, S. G., Buckeridge, M., Ramos, L. P., Pitarelo, A. P., Ferreira-Leito, V., Gottschalk, L. M. F., Ferrara, M. A., Silva Bon, E. P. da, Moraes, L. M. P. de, Araujo, J. de A., & Torres, F. A. G. (2010). Bioethanol from lignocelluloses: Status and perspectives in Brazil. *Bioresource Technology*, 101, 4820–4825.
  166. Somerville, C., Youngs, H., Taylor, C., Davis, S. C., & Long, S. P. (2010). Feedstocks for lignocellulosic biofuels. *Science (New York, N.Y.)*, 329(5993), 790–792.
  167. Sonbol, F.-M., Fornalé, S., Capellades, M., Encina, A., Tourino, S., Torres, J.-L., Rovira, P., Ruel, K., Puigdomenech, P., Rigau, J., & others. (2009). The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in *Arabidopsis thaliana*. *Plant Molecular Biology*, 70(3), 283–296.
  168. Stevenson, G. C. (1965). *Genetics and breeding of sugarcane*. Longman.
  169. Stewart, J. J., Akiyama, T., Chapple, C., Ralph, J., & Mansfield, S. D. (2009). The effects on lignin structure of overexpression of *ferulate 5-hydroxylase* in hybrid poplar. *Plant Physiology*, 150(2), 621–635.
  170. Suzuki, S., Li, L., Sun, Y.-H., & Chiang, V. L. (2006). The cellulose synthase gene superfamily and biochemical functions of xylem-specific cellulose synthase-like genes in *Populus trichocarpa*. *Plant Physiology*, 142(3), 1233–1245.
  171. Swain, T., & Hillis, W. E. (1959). The phenolic constituents of *Prunus domestica*. I.—The quantitative analysis of phenolic constituents. *Journal of the Science of Food and Agriculture*, 10(1), 63–68.
  172. Szczerbowski, D., Pitarelo, A. P., Zandoná Filho, A., & Ramos, L. P. (2014).

- Sugarcane biomass for biorefineries: comparative composition of carbohydrate and non-carbohydrate components of bagasse and straw. *Carbohydrate Polymers*, 114, 95–101.
173. Takeda, Y., Koshiba, T., Tobimatsu, Y., Suzuki, S., Murakami, S., Yamamura, M., Rahman, M. M., Takano, T., Hattori, T., Sakamoto, M., & others. (2017). Regulation of *CONIFERALDEHYDE 5-HYDROXYLASE* expression to modulate cell wall lignin structure in rice. *Planta*, 246(2), 337–349.
  174. Tamagnone, L., Merida, A., Parr, A., Mackay, S., Cullianez-Macia, F. A., Roberts, K., & Martin, C. (1998). The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *The Plant Cell*, 10(2), 135–154.
  175. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28(10), 2731–2739.
  176. Tanaka, K., Murata, K., Yamazaki, M., Onosato, K., Miyao, A., & Hirochika, H. (2003). Three distinct rice cellulose synthase catalytic subunit genes required for cellulose synthesis in the secondary wall. *Plant Physiology*, 133(1), 73–83.
  177. Taylor, N. G., Gardiner, J. C., Whiteman, R., & Turner, S. R. (2004). Cellulose synthesis in the *Arabidopsis* secondary cell wall. *Cellulose*, 11(3), 329–338.
  178. Team, R. C., & others. (2013). *R: A language and environment for statistical computing*.
  179. Tetreault, H. M., Gries, T., Palmer, N. A., Funnell-Harris, D. L., Sato, S., Ge, Z., Sarath, G., & Sattler, S. E. (2020). Overexpression of *ferulate 5-hydroxylase* increases syringyl units in *Sorghum bicolor*. *Plant Molecular Biology*, 1–17.
  180. Tew, T., & Cobill, R. (2008). Genetic Improvement of Sugarcane (*Saccharum* spp.) as an Energy Crop. In W. Vermerris (Ed.), *Genetic Improvement of Bioenergy Crops SE - 9* (pp. 273–294). Springer New York.
  181. Thompson, J. D., Gibson, T., Higgins, D. G., & others. (2002). Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*, 2–3.
  182. Tian, F., Yang, D.-C., Meng, Y.-Q., Jin, J., & Gao, G. (2020). PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Research*, 48(D1), D1104-D1113.
  183. Valdivia, E. R., Herrera, M. T., Gianzo, C., Fidalgo, J., Revilla, G., Zarra, I., &

- Sampedro, J. (2013). Regulation of secondary wall synthesis and cell death by NAC transcription factors in the monocot *Brachypodium distachyon*. *Journal of Experimental Botany*, ers394.
184. Valério, L., Carter, D., Rodrigues, J. C., Tournier, V., Gominho, J., Marque, C., Boudet, A.-M., Maunders, M., Pereira, H., & Teulières, C. (2003). Down regulation of *cinnamyl alcohol dehydrogenase*, a lignification enzyme, in *Eucalyptus camaldulensis*. *Molecular Breeding*, 12(2), 157–167.
185. Van Acker, R., Vanholme, R., Storme, V., Mortimer, J. C., Dupree, P., & Boerjan, W. (2013). Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnology for Biofuels*, 6(1), 46.
186. van der Weijde, T., Alvim Kamei, C. L., Torres, A. F., Vermerris, W., Dolstra, O., Visser, R. G. F., & Trindade, L. M. (2013). The potential of C4 grasses for cellulosic biofuel production. *Frontiers in Plant Science*, 4, 107.
187. Van Heiningen, A. (2006). Converting a kraft pulp mill into an integrated forest biorefinery. *Pulp and Paper Canada*, 107(6), 38–43.
188. Vanholme, R., Cesarino, I., Rataj, K., Xiao, Y., Sundin, L., Goeminne, G., Kim, H., Cross, J., Morreel, K., Araujo, P., & others. (2013). *Caffeoyl shikimate esterase* (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science*, 341(6150), 1103–1106.
189. Vanholme, R., Demedts, B., Morreel, K., Ralph, J., & Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiology*, 153(3), 895–905.
190. Vargas, L., Cesarino, I., Vanholme, R., Voorend, W., Saleme, M. L. S., Morreel, K., & Boerjan, W. (2016). Improving total saccharification yield of *Arabidopsis* plants by vessel-specific complementation of *caffeoyl shikimate esterase* (*cse*) mutants. *Biotechnology for Biofuels*, 9(1), 139.
191. Vettore, A. L., Silva, F. R. da, Kemper, E. L., & Arruda, P. (2001). The libraries that made SUCEST. *Genetics and Molecular Biology*, 24(1–4), 1–7.
192. Vicentini, R., Bottcher, A., dos Santos Brito, M., dos Santos, A. B., Creste, S., de Andrade Landell, M. G., Cesarino, I., & Mazzafera, P. (2015). Large-scale transcriptome analysis of two sugarcane genotypes contrasting for lignin content. *PloS One*, 10(8), e0134909.
193. Vijayan Nair, N., Nair, S., Sreenivasan, T. V., & Mohan, M. (1999). Analysis of genetic diversity and phylogeny in *Saccharum* and related genera using RAPD

- markers. *Genetic Resources and Crop Evolution*, 46(1), 73–79.
194. Vogel, J. (2008). Unique aspects of the grass cell wall. *Current Opinion in Plant Biology*, 11(3), 301–307.
195. Waclawovsky, A. J., Sato, P. M., Lembke, C. G., Moore, P. H., & Souza, G. M. (2010). Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. *Plant Biotechnology Journal*, 8(3), 263–276.
196. Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Geddes, B., Lu, F., & Ralph, J. (2015). Syringyl lignin production in conifers: Proof of concept in a Pine tracheary element system. *Proceedings of the National Academy of Sciences*, 112(19), 6218–6223.
197. Wang, H.-Z., & Dixon, R. A. (2012). On--off switches for secondary cell wall biosynthesis. *Molecular Plant*, 5(2), 297–303.
198. Wang, H., Avci, U., Nakashima, J., Hahn, M. G., Chen, F., & Dixon, R. A. (2010). Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proceedings of the National Academy of Sciences*, 107(51), 22338–22343.
199. Wang, Q., Wang, M., Zhang, X., Hao, B., Kaushik, S. K., & Pan, Y. (2011). WRKY gene family evolution in *Arabidopsis thaliana*. *Genetica*, 139(8), 973–983.
200. Wang, Shaogan, Yang, H., Mei, J., Liu, X., Wen, Z., Zhang, L., Xu, Z., Zhang, B., & Zhou, Y. (2019). Rice homeobox protein KNAT7 integrates the pathways regulating cell expansion and wall stiffness. *Plant Physiology*, 181(2), 669–682.
201. Wang, Shumin, Yamaguchi, M., Grienberger, E., Martone, P. T., Samuels, A. L., & Mansfield, S. D. (2020). The Class II KNOX genes KNAT3 and KNAT7 work cooperatively to influence deposition of secondary cell walls that provide mechanical support to *Arabidopsis* stems. *The Plant Journal*, 101(2), 293–309.
202. Wang, Y.-H., Acharya, A., Burrell, A. M., Klein, R. R., Klein, P. E., & Hasenstein, K. H. (2013). Mapping and candidate genes associated with saccharification yield in sorghum. *Genome*, 56(11), 659–665.
203. Wang, Yin, Chantreau, M., Sibout, R., & Hawkins, S. (2013). Plant cell wall lignification and monolignol metabolism. *Frontiers in Plant Science*, 4, 220.
204. Wang, Youhua, Wan, L., Zhang, L., Zhang, Z., Zhang, H., Quan, R., Zhou, S., & Huang, R. (2012). An ethylene response factor OsWR1 responsive to drought stress transcriptionally activates wax synthesis related genes and increases wax production in rice. *Plant Molecular Biology*, 78(3), 275–288.

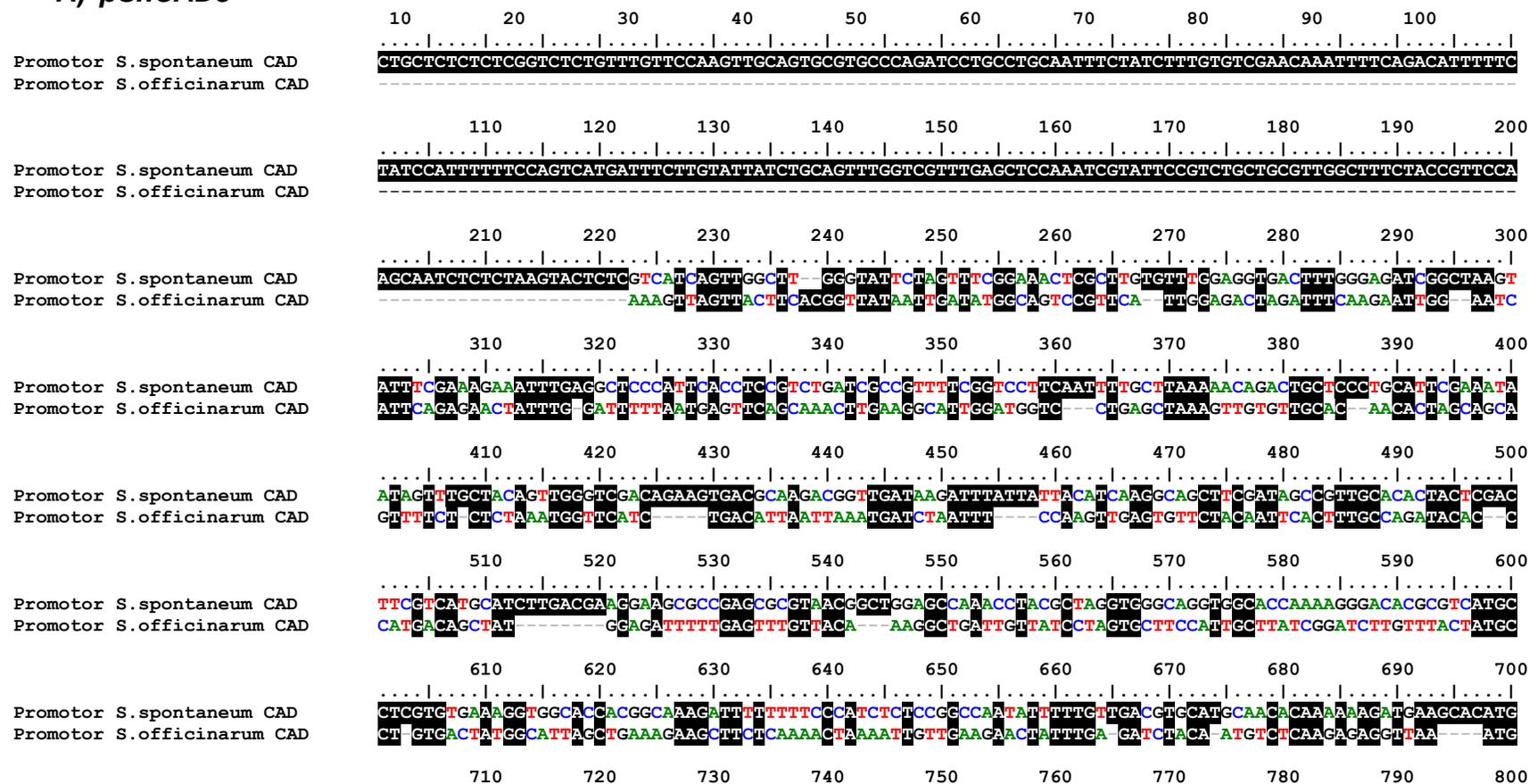
205. Welker, C. M., Balasubramanian, V. K., Petti, C., Rai, K. M., DeBolt, S., & Mendu, V. (2015). Engineering plant biomass lignin content and composition for biofuels and bioproducts. *Energies*, 8(8), 7654–7676.
206. Wilczyński, B., & Dojer, N. (2009). BNFinder: exact and efficient method for learning Bayesian networks. *Bioinformatics*, 25(2), 286–287.
207. Wilkerson, C. G., Mansfield, S. D., Lu, F., Withers, S., Park, J.-Y., Karlen, S. D., Gonzales-Vigil, E., Padmakshan, D., Unda, F., Rencoret, J., & others. (2014). Monolignol *ferulate transferase* introduces chemically labile linkages into the lignin backbone. *Science*, 344(6179), 90–93.
208. Wilson, J. R., & Hatfield, R. D. (1997). Structural and chemical changes of cell wall types during stem development: consequences for fibre degradation by rumen microflora. *Australian Journal of Agricultural Research*, 48, 165–180.
209. Wu, Z., Wang, N., Hisano, H., Cao, Y., Wu, F., Liu, W., Bao, Y., Wang, Z.-Y., & Fu, C. (2019). Simultaneous regulation of F5H in COMT-RNA i transgenic switchgrass alters effects of COMT suppression on syringyl lignin biosynthesis. *Plant Biotechnology Journal*, 17(4), 836–845.
210. Wuddineh, W. A., Mazarei, M., Turner, G. B., Sykes, R. W., Decker, S. R., Davis, M. F., & Stewart Jr, C. N. (2015). Identification and molecular characterization of the switchgrass AP2/ERF transcription factor superfamily, and overexpression of PvERF001 for improvement of biomass characteristics for biofuel. *Frontiers in Bioengineering and Biotechnology*, 3, 101.
211. Yamaguchi, M., Kubo, M., Fukuda, H., & Demura, T. (2008). VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in *Arabidopsis* roots and shoots. *The Plant Journal*, 55(4), 652–664.
212. Yamaguchi, M., Mitsuda, N., Ohtani, M., Ohme-Takagi, M., Kato, K., & Demura, T. (2011). VASCULAR-RELATED NAC-DOMAIN 7 directly regulates the expression of a broad range of genes for xylem vessel formation. *The Plant Journal*, 66(4), 579–590.
213. Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., & Demura, T. (2010). VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in *Arabidopsis*. *The Plant Cell*, 22(4), 1249–1263.
214. Yang, F., Li, W., Jiang, N., Yu, H., Morohashi, K., Ouma, W. Z., Morales-Mantilla, D. E., Gomez-Cano, F. A., Mukundi, E., Prada-Salcedo, L. D., & others. (2017). A

- maize gene regulatory network for phenolic metabolism. *Molecular Plant*, 10(3), 498–515.
215. Yang, L., Zhao, X., Yang, F., Fan, D., Jiang, Y., & Luo, K. (2016). PtrWRKY19, a novel WRKY transcription factor, contributes to the regulation of pith secondary wall formation in *Populus trichocarpa*. *Scientific Reports*, 6(1), 1–12.
  216. Yokoyama, K. D., Zhang, Y., & Ma, J. (2014). Tracing the evolution of lineage-specific transcription factor binding sites in a birth-death framework. *PLoS Computational Biology*, 10(8), e1003771.
  217. Yu, Y. (2019). *OsKNAT7 Bridges Secondary Cell Wall Formation and Cell Growth Regulation*. Am Soc Plant Biol.
  218. Yuan, J. S., Tiller, K. H., Al-Ahmad, H., Stewart, N. R., & Stewart, C. N. (2008). Plants to power: bioenergy to fuel the future. *Trends in Plant Science*, 13(8), 421–429.
  219. Zhang, J., Zhang, S., Li, H., Du, H., Huang, H., Li, Y., Hu, Y., Liu, H., Liu, Y., Yu, G., & others. (2016). Identification of transcription factors ZmMYB111 and ZmMYB148 involved in phenylpropanoid metabolism. *Frontiers in Plant Science*, 7, 148.
  220. Zhang, Y., Legland, D., El Hage, F., Devaux, M.-F., Guillon, F., Reymond, M., & Méchin, V. (2019). Changes in cell walls lignification, feruloylation and p-coumaroylation throughout maize internode development. *PloS One*, 14(7), e0219923.
  221. Zhao, Q., & Dixon, R. A. (2011). Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends in Plant Science*, 16(4), 227–233.
  222. Zhao, Q., Wang, H., Yin, Y., Xu, Y., Chen, F., & Dixon, R. A. (2010). Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. *Proceedings of the National Academy of Sciences*, 107(32), 14496–14501.
  223. Zhong, R., Demura, T., & Ye, Z.-H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*. *The Plant Cell*, 18(11), 3158–3170.
  224. Zhong, R., Lee, C., McCarthy, R. L., Reeves, C. K., Jones, E. G., & Ye, Z.-H. (2011). Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant and Cell Physiology*, 52(10), 1856–1871.
  225. Zhong, R., Lee, C., & Ye, Z.-H. (2010a). Evolutionary conservation of the

- transcriptional network regulating secondary cell wall biosynthesis. *Trends in Plant Science*, 15(11), 625–632.
226. Zhong, R., Lee, C., & Ye, Z.-H. (2010b). Global analysis of direct targets of secondary wall NAC master switches in *Arabidopsis*. *Molecular Plant*, 3(6), 1087–1103.
227. Zhong, R., Lee, C., Zhou, J., McCarthy, R. L., & Ye, Z.-H. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *The Plant Cell*, 20(10), 2763–2782.
228. Zhong, R., McCarthy, R. L., Haghghat, M., & Ye, Z.-H. (2013). The poplar MYB master switches bind to the SMRE site and activate the secondary wall biosynthetic program during wood formation. *PLoS One*, 8(7), e69219.
229. Zhong, R., McCarthy, R. L., Lee, C., & Ye, Z.-H. (2011). Dissection of the transcriptional program regulating secondary wall biosynthesis during wood formation in poplar. *Plant Physiology*, 157(3), 1452–1468.
230. Zhong, R., Richardson, E. A., & Ye, Z.-H. (2007). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis*. *Planta*, 225(6), 1603–1611.
231. Zhong, R., & Ye, Z.-H. (2009). Transcriptional regulation of lignin biosynthesis. *Plant Signaling & Behavior*, 4(11), 1028–1034.
232. Zhong, R., & Ye, Z.-H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant and Cell Physiology*, 53(2), 368–380.
233. Zhong, R., & Ye, Z.-H. (2015). Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. *Plant and Cell Physiology*, 56(2), 195–214.
234. Zhong, R., Yuan, Y., Spiekerman, J. J., Guley, J. T., Egbosiuba, J. C., & Ye, Z.-H. (2015). Functional Characterization of NAC and MYB Transcription Factors Involved in Regulation of Biomass Production in Switchgrass (*Panicum virgatum*). *PLoS One*, 10(8), e0134611.
235. Zhou, J., Lee, C., Zhong, R., & Ye, Z.-H. (2009). MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *The Plant Cell*, 21(1), 248–266.

## 6.MATERIAL SUPLEMENTAR

Figura S1 – Alinhamento de seqüências de nucleotídeos da região regulatória isolada dos genes da via biossintética de lignina *ShCAD8*, *ShCOMT* e *ShF5H* nas espécies: *Saccharum spontaneum* e *Saccharum officinarum*.

A) *pShCAD8*

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

810 820 830 840 850 860 870 880 890 900

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

910 920 930 940 950 960 970 980 990 1000

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

1010 1020 1030 1040 1050 1060 1070 1080 1090 1100

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

1110 1120 1130 1140 1150 1160 1170 1180 1190 1200

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

1510 1520 1530 1540 1550 1560 1570 1580 1590 1600

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

1610 1620 1630 1640 1650 1660 1670 1680 1690 1700

Promotor *S.spontaneum* CAD  
Promotor *S.officinorum* CAD

```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
C T T C A A A A T G C T G T A A T A T T G C T T T T T G G A G G C C G A A A C T A C G T T G G A T T T C C A G C G G A A T C A T T T A G C C A A T T A A C T T G C T A G G A A C A T A C A A A A G A
G T T T G G C A A C A T T A T A C A T T G A ----- G A A G C G A T T G T T G G A T A A G A T C G A T A T T C A T A C C A T C A T C A A T G A C T T C G C A T C G A G G A A T G T T A G A A G

810 820 830 840 850 860 870 880 890 900
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
A A A T A A A A C A T A T A C T A G T A C G T A C T G C C A A T A G T C C T C T G G A C C A G G T C T T C T A T G G T C C G C G C G C T G C C G C A A C C A A A T C A C G A G G A G C C G T G
G A T T T T T A A G A T A T T T A A T A T A G G A A T T T A T A G C T A A T T A A T A T T A T G A A T T G T T C A T A T A A T A T T G C T A T T G C A T C C G A G T T A A G T G T

910 920 930 940 950 960 970 980 990 1000
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
A C A A T C C G G G T C C T C G G T C C T C A A C T C A C A T G C A G C T T T C A G C G A G C A G C C A C A C A T C A G A A A T A T C A A G C C G T G A A T C A G T T C A C A A C A G
T T A T A T A T A T G C A T T T A C A T T A G T A C T A T C A A T G C A T G T A T G T A T A A A G G C C C G G A T T A T ----- G G T T C G C C C C G G C C C T T A A A ----- A A A A T C A G

1010 1020 1030 1040 1050 1060 1070 1080 1090 1100
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
A G C A T A T A G T A C C T T C G A A G A A A A T A C C A C T T T A G G G T G T G T C A C T T C C T T C A A A C T T C C A A A A A T T T T A C A A G A T T T C C G T C A C A T C G A A T C T T T C
G C C G G C C C T G A T T T G G A T T G A A A A G C A C T T T ----- G T G G T T G G T T ----- A G T T A T A T A A G A T C A C A T C G T C S A T C G S C T G C T A T

1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
A A C A C A T G C A C G S A C A T T A A A T A T T T A A T A A A A A A T A A C T A A T T A C A T A G T T T A G T T G T A A T C A C G A G A C G A A T C T T T A A A A C A T A A T T A G T C C A T G A T
G T A G T A T C T A C T C A G G T C G C A T G C A C G C A C A ----- C G T T C G A T C C A A T C C C G G C C T G T G C A G T G C A G G C C A ----- C C T C A G A C C G C ----- T C G A T C G C

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T A G A C A T T A A A T T A C C A A A T A C A A A T A A A A G A G T T A C T C T G C T T C C A A A A A T T T T C C A C C C C T A A C A G A A T C A C G A A C T C G A C A C A T G C T T T T G G T T C C T C
C G C C T C C G A T C C C C ----- T C C A A G ----- C T A C A G G T C T C C ----- A A A C C G T T G A T G T T A G C A G A T A G T T G A G G T G T G T C A T C G T C C T T T G G A A A A

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T C G C T T T G G C A T G T G C G T G G G C C T A T G T C G T T A G G A T A C C C C T C A C T A T C C A T C T A C T A G T A C T T T T A C G T C A C T T C T T T T C C T T T T G C A G C C A C T C T A T
A A A C A T T G G A A C A A A C C C A C C A C C C A C C G T G C C G T G C C C T C A C T C A T C C A T C T A C A G ----- C T T T A C G T C A C T T C T T T T C C T T T T G C A G C C A C T C T A T

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
A A T T T C C T T T T T G A C A A C T G A G T C C T A C G A C A A A G T G A G A G T A T T C C A C T A T C C T A A A T A T A G T T G T T T T T T T A A C T A A T T C A G T C T A C C G G T T C A A
A A T T T C C T T T T T G A C A A C G G A G T C C T A C G A C A A A G T G A G A G T A T T C C A C T A T C C T A A A T A T A G T T G T T T T T T T A A C T A A T T C A G T C T A C C G G T T C A A

1510 1520 1530 1540 1550 1560 1570 1580 1590 1600
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T C A T T T T G T T A T G A A C G G A A A T T T G G A G A A A T C T G C C T C C G G A T C A T A T C C T T G C G T C C T G T G G A T G T C A T G A T G T T T T T C A A A G A G T C A A G G C T T T T T T
T T A T T T T G T T A T G A A C G G A A A T T T G G A G A A A T C T G C C T C C G G A T C A T A T C C T T G C G T C C T G T G G A T G T C A T G A T G T T T T T C A A A G A G T C A A G G ----- T T T T T T T

1610 1620 1630 1640 1650 1660 1670 1680 1690 1700
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
T A G A T A T A G A G G A T A T A A A A T T G T T C T C T A T A T C C A C A T T G A A T A T A T G T A G T C A G A A G G T A T A C A A G C T T T T A G G C T C A A A A C T T C A A C C C G A G G A G
T A G A T A T A G A G G A T A T A A A A T T G T T C T C T A T A T C C A C A T T G A A T A T A T G T A G T C A G A A G G T A T A C A A G C T T T T A G G C T C A A A A C T T C A A A C C C G A G A A G

```

```

      1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
Promotor S.spontaneum CAD CATAAAATAAAACAACAAAAAAAGAGAAAGCTTAAAGGACTAAGAAAAAAACCAACGAAAGACTAGCTTCAGTTTTTTTTTTTTTTTGCCTTGCCTTC
Promotor S.officinarum CAD CATAAAATAAAACAACAAAAAAGAGAAAGCTTAAAGGACTAAGAAAAAAACCAACGAAAGACTAGCTTCAGTTTTTTTTTTTTTTTGCCTTGCCTTC

      1810      1820      1830      1840      1850      1860      1870      1880      1890      1900
Promotor S.spontaneum CAD CAACTTATTTTGTCTTTCGCATAGCAGAACCCTCGCCACATCAAACAAGAACTAAGCCCCCTATCCTTATCAGCGTTTAACGCAAGAAAATCCTGATAC
Promotor S.officinarum CAD CAACTTATTTTGTCTTTCGCATAGCAGAACCCTCGCCACATCAAACAAGAACTAAGCCCCCTATCCTTATCAGCGTTTAACGCAAGAAAATCCTGATAC

      1910      1920      1930      1940      1950      1960      1970      1980      1990      2000
Promotor S.spontaneum CAD ATTGCCACATTGGCCCTCAACCCTCAATAGTCCACAGGTCTCTCAGGGAGCAGCAGTATGTCATATCATTGTTCGGCCCTCGTCTCCACGGTCCATT
Promotor S.officinarum CAD ATTGCCACATTGGCCCTCAACCCTCAATAGTCCACAGGTCTCTCAGGGAGCAGCAGTATGTCATATCATTGTTCGGCCCTCGTCTCCACGGTCCATT

      2010      2020      2030      2040      2050      2060      2070      2080      2090      2100
Promotor S.spontaneum CAD TCTCCACCGCTCTTCTCCCCCGAGCTGCTGCTCCTGCTCCCTCGCTCCGAGTTAGTGTAGGGTTGCCCTCGAGAGCTCTGCCGTTGCCGCGGCCAGC
Promotor S.officinarum CAD TCTCCACCGCTCTTCTCCCCCGAGCTGCTGCTCCTGCTCCCTCGCTCCGAGTTAGTGTAGGGTTGCCCTCGAGAGCTCTGCCGTTGCCGCGGCCAGC

      2110      2120      2130      2140      2150      2160      2170      2180      2190      2200
Promotor S.spontaneum CAD GAGGGCGTCCGTGGAGAAGCGAGAGCAGCAGGTGACGATGGAGGAGCAAGGCGCCAGGCGGCGTTCGGATGGGCGCCAGGGACGATACCGGCGTCTCCTC
Promotor S.officinarum CAD GAGGGCGTCCGTGGAGAAGCGAGAGCAGCAGGTGACGATGGAGGAGCAAGGCGCCAGGCGGCGTTCGGATGGGCGCCAGGGACGATACCGGCGTCTCCTC

      2210      2220
Promotor S.spontaneum CAD TCCCCCTACAACCTTCTCCAG
Promotor S.officinarum CAD TCCCCCTACAACCTTCTCCAG

```

**B) pShCOMT**

```

      10      20      30      40      50      60      70      80      90      100
Promotor S.officinarum COMT GCCCTTGCATATAAGTGTACTAAAGTTTACAAGATTAATAATAATCCACGAAAAATCAACTACTATACACTTGAGCTTTTACAATCTATTTCAATTTCTTAG
Promotor S.spontaneum COMT GCCCTTGCATATAAGTGTACTAAAGTTTACAAGATTAATAATAATCCACGAAAAATCAACTACTATACACTTGAGCTTTTACAATCTATTTCAATTTCTTAG

      110      120      130      140      150      160      170      180      190      200
Promotor S.officinarum COMT ATGAATTATACATTCTTTTATAAAAAATCAGATAGAGCCATGATTTTTGTATTTCATCAGTTATGTTATAATCCGTTCCGGCTGCAGGTGCTGCTGCTGCT
Promotor S.spontaneum COMT ATGAATTATACATTCTTTTATAAAAAATCAGATAGAGCCATGATTTTTGTATTTCATCAGTTATGTTATAATCCGTTCCGGCTGCAGGTGCTGCTGCTGCT

      210      220      230      240      250      260      270      280      290      300

```

```

Promotor S.officinarum COMT GCTGG-----GGATGATAAGCTTCAGATCGATCTGTTAGGCAGAAAGATCATACCT
Promotor S.spontaneum COMT GCTGGTGGCTGGCTGCAGCTAATATAAATCCACTGCACAGCTGCAGCAACGGCCGGATGATAAGCTTCAGATCGATCTGTTAGGCAGAAAGATCATACCT

      310      320      330      340      350      360      370      380      390      400
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT ATGATTTCTGATCGTGACAAATAACCTTTTGTCAAGCTGTGCCAACTAAAAAATGAGCAAGGATTAATTGAGTACTAATTAATAACAGTAAAAATATTACCG
Promotor S.spontaneum COMT ATGATTTCTGATCGTGACAAATAACCTTTTGTCAAGCTGTGCCAACTAAAAAATGAGCAAGGATTAATTGAGTACTAATTAATAACAGTAAAAATATTACCA

      410      420      430      440      450      460      470      480      490      500
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT CGATAAATAATCCTTGTGAAATCCCAACGACAGCTGCCACCGTCATCCCCGACCAACCCGGCTCGGCTGGTCACTGGTGTGGTGACCTCTGTGTTCATG
Promotor S.spontaneum COMT CGATAAATAATCCTTGTGAAATACCAAGACAGCTGCCACCGTCATCCCCGACCAACCCGGCTCGGCTGGTCACTGGTGTGGTGACCTCTGTGTTCATG

      510      520      530      540      550      560      570      580      590      600
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT TCCATGCATGGCCTATTGTGAGTGGGCTCCGGGCTCCGGGCAACCTACTGGCGGGCCACTTACTGCACCCCTCCTACCGGCCTACTCCCAAATTAAA
Promotor S.spontaneum COMT TCCATGCATGGCCTATTGTGAGTGGGCTCCGGGCTCCGGGCAACCTACTGGCGGGCCACTTACTGCACCCCTCCTACCGGCCTACTCCCAAATTAAA

      610      620      630      640      650      660      670      680      690      700
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT GTCCAGCTCATCATTTTATTTCTCGCTGAATTTGTGCATA-----CTAAATTTCTTTGCACGTTGCAAACCCAAGCAAAAATATCGACAACGGTCGGAT
Promotor S.spontaneum COMT GTCCAGCTCATCATTTTATTTCTCGCTGAATTTGTGCATAAATTTGCTAAATTTCTTTGCACGTTGCAAACCCAAGCAAAAATATCG-----ATCG

      710      720      730      740      750      760      770      780      790      800
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT TTTTTTCTGATAAATAATAGTCTGTTATTCAGCATTGAAAAACAACAAGCCAGAACACCGTATACTAGTATGTACTACATCACTGTGCATGGATGG
Promotor S.spontaneum COMT TTTTCTGATAAATAATAGTCTGTTATTCAGCATTGAAAAACAACAAGCCAGAACACCGTATACTAGTATGTACTACATCACTGTGCATGGATGG

      810      820      830      840      850      860      870      880      890      900
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT TAGATGGAGCTAGCATGCACACTGACTTTGTAGCTTAGTGAAGTGAATTTTCTCAGGGCACGGACGGCATCACCGGTCACCACTCACAACCACTTCC
Promotor S.spontaneum COMT TAGATGGAGCTAGCATGCACACTGACTTTGTAGCTTAGTGAAGTGAATTTTCTCAGGGCACGGACGGCATCACCGGTCCTACTCACAACCACTTCC

      910      920      930      940      950      960      970      980      990      1000
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT GCTGTGGTTGGGCGAGCAGCATCAACAAATCTCTACGCACTTGACCCTGACGCAACCTAACGCAATAGATTAATCATCACCAACAAACACCAGCAGAAAT
Promotor S.spontaneum COMT GCTGTGGTTGGGCGAGCAGCATCAACAAATCTCTACGCACTTGACCCTGACGCAACCTAACGCAATAGATTAATCATCACCAACAAACACCAGCAGAAAT

      1010     1020     1030     1040     1050     1060     1070     1080     1090     1100
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT GGGCCCTACTTATCTCTTAACCACTGTCGTCATGCATTATTACGCATAGCATATATATAG-----GAGATTAGTACGCATAGCATATATAGGAGAT
Promotor S.spontaneum COMT GGGCCCTACTTATCTCTTAACCACTGTCGTCATGCATTATTACGCATAGCATATATATAGCCTATAGGAGATTAGTACGCATAGCATATATAGGAGAT

      1110     1120     1130     1140     1150     1160     1170     1180     1190     1200
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Promotor S.officinarum COMT TAGTA-----GCTTTTTCTTAGTGCCATGCATCTTTTCATGTACATTTTCTTCCAAAATTAATAATTCGAAAAGGAGGTGAGAAAAAGCAAAGCAG
Promotor S.spontaneum COMT TAGTACTTTAGTAGCTTTTCTTAGTGCCATGCATCTTTTCATGTACATTTTCTTCCAAAATTAATAATTCGAAAAGGAGGTGAGAAAAAGCAAAGCAG

      1210     1220     1230     1240     1250     1260     1270     1280     1290     1300
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

```

Promotor *S.officinarum* COMT  
 Promotor *S.spontaneum* COMT

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400

Promotor *S.officinarum* COMT  
 Promotor *S.spontaneum* COMT

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500

Promotor *S.officinarum* COMT  
 Promotor *S.spontaneum* COMT

1510

Promotor *S.officinarum* COMT  
 Promotor *S.spontaneum* COMT

C) *pShF5H*

10 20 30 40 50 60 70 80 90 100

Promotor *S.officinarum* F5H  
 Promotor *S.spontaneum* F5H

110 120 130 140 150 160 170 180 190 200

Promotor *S.officinarum* F5H  
 Promotor *S.spontaneum* F5H

210 220 230 240 250 260 270 280 290 300

Promotor *S.officinarum* F5H  
 Promotor *S.spontaneum* F5H

310 320 330 340 350 360 370 380 390 400

Promotor *S.officinarum* F5H  
 Promotor *S.spontaneum* F5H

410 420 430 440 450 460 470 480 490 500

Promotor *S.officinarum* F5H  
 Promotor *S.spontaneum* F5H



```

          1510      1520      1530      1540      1550      1560      1570      1580      1590      1600
Promotor S.officinarum F5H TCCTCTCACCATTAGCCACAATTATCTGATCAAAGTTTTCAACCCAATCTGTATA GTCTTCGTCCTCGGTCCTTACCTACTCCTCCGGCTATATAAAC
Promotor S.spontaneum F5H TCCTCTCACCATTAGCCACAATTATCTGATCAAAGTTTTCAACCCAATCTGTATA GTCTTCGTCCTCGGTCCTTACCTACTCCTCCGGCTATATAAAC

          1610      1620      1630      1640      1650      1660      1670      1680      1690      1700
Promotor S.officinarum F5H CCGTGCTCCTCCCCCTTTCTCCTCAGCTC CACTCGAAGAGCGTTCTACCACCACACCACAAGTCCACAATCTAAAAC CAAACGACAAAAAGAAGGAAGG
Promotor S.spontaneum F5H CCGTGCTCCTCCCCCTTTCTCCTCAGCTC CACTCGAAGAGCGTTCTACCACCACACCACAAGTCCACAATCTAAAAC CAAACGACAAAAAGAAGGAAGG

          1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
}
Promotor S.officinarum F5H AAAAAAAAAAGGTGCTGCCGATCGA T CATG GCGGCCGTGGCCAAGATCGCCATGGAGTGGCTCCAAGACCCTCTGAGCTGCGTGTTCCTTGTCACGCTGGA
Promotor S.spontaneum F5H AAAAAAA GGTGCTGCCGATCGA C ATG GCGGCCGTGGCCAAGATCGCCATGGAGTGGCTCCAAGACCCTCTGAGCTGCGTGTTCCTTGTCACGCTGGA

          ....|
Promotor S.officinarum F5H AGGGC
Promotor S.spontaneum F5H AGGGC

```

**Tabela S1 - Sequências SNBE e SMRE identificadas na região regulatórias dos genes *ShCAD8*, *ShCOMT* e *ShF5H* nas espécies de *Saccharum***

<b>A) <i>ShCAD8</i></b>		
SNBE	ESPECIE	POSIÇÃO
TATTTTGTATGAACGGAA	<i>S.officinarum</i>	+1,166,1184,-453,-435
AAGCTTAAGGACTAAGAAA	<i>S.officinarum</i>	+1,390,1408,-229,-211
TTCTGTTATGCGAAAGACA	<i>S.officinarum</i>	-1,490,1472,-129,-147
TTTCTTAGTCCTTAAGCTT	<i>S.officinarum</i>	-1,408,1390,-211,-229
ATTCTCGATGCGAAGTCA	<i>S.officinarum</i>	-521,503,-1098,-1116
SMRE	ESPECIE	POSIÇÃO
ACCAACC	<i>S.officinarum</i>	-760,754,-859,-865
SNBE	ESPECIE	POSIÇÃO
AAGCTTAAGGACTAAGAAA	<i>S.spontaneum</i>	+1,720,1738,-231,-213
TTTCTTAGTCCTTAAGCTT	<i>S.spontaneum</i>	-1,738,1720,-213,-231
TTGTGTTGCATGCACGTCA	<i>S.spontaneum</i>	-681,663,-1270,-1288
SMRE	ESPECIE	POSIÇÃO
TACCAAAT	<i>S.spontaneum</i>	+1,206,1213,-745,-738
ACCAAAT	<i>S.spontaneum</i>	+1,207,1213,-744,-738
ACCAAAC	<i>S.spontaneum</i>	-147,141,-1804,-1810

<b>B) <i>ShCOMT</i></b>		
SMRE	ESPECIE	POSIÇÃO
ACCAACC	<i>S.spontaneum</i>	+444,450,-919,-913
ACCTACT	<i>S.spontaneum</i>	+532,538,-831,-825
ACCTAAC	<i>S.spontaneum</i>	+921,927,-442,-436
CACCAACC	<i>S.spontaneum</i>	-1,197,1190,-166,-173
CACCAAAC	<i>S.spontaneum</i>	-1,185,1178,-178,-185
TACCAACT	<i>S.spontaneum</i>	-1,168,1161,-195,-202
ACCAACC	<i>S.spontaneum</i>	-1,196,1190,-167,-173
ACCAAAC	<i>S.spontaneum</i>	-1,184,1178,-179,-185
ACCAACT	<i>S.spontaneum</i>	-,1167,1161,-196,-202
SMRE	ESPECIE	POSIÇÃO
ACCAACC	<i>S.officinarum</i>	+397,403,-928,-922
ACCTACT	<i>S.officinarum</i>	+491,497,-834,-828
ACCTAAC	<i>S.officinarum</i>	+892,898,-433,-427
CACCAACC	<i>S.officinarum</i>	-1,157,1150,-168,-175
CACCAAAC	<i>S.officinarum</i>	-1,145,1138,-180,-187
TACCAACT	<i>S.officinarum</i>	-,1128,1121,-197,-204
ACCAACC	<i>S.officinarum</i>	-1,156,1150,-169,-175
ACCAAAC	<i>S.officinarum</i>	1144,1138,-181,-187
ACCAACT	<i>S.officinarum</i>	1127,1121,-198,-204

<b>c) ShF5H</b>		
SNBE	ESPECIE	POSIÇÃO
AATTTTGTAATTAAGCCT	<i>S.officinarum</i>	1650+,938,956,-712,-694
SMRE	ESPECIE	POSIÇÃO
ACCTAAT	<i>S.officinarum</i>	1650+,435,441,-1215,-1209
ACCTACT	<i>S.officinarum</i>	1650+,1500,1506,-150,-144
SNBE	ESPECIE	POSIÇÃO
AATTTTGTAATTAAGCCT	<i>S.spontaneum</i>	1619+,953,971,-666,-648
AATTCTATTGAGTAAGGAA	<i>S.spontaneum</i>	1619,-,113,95,-1506,-1524
SMRE	ESPECIE	POSIÇÃO
ACCTAAT	<i>S.spontaneum</i>	1619+,446,452,-1173,-1167
CCTACT	<i>S.spontaneum</i>	1619+,1490,1496,-129,-123

**Tabela S2 - Resumo das Análises feitas para a identificação de sequencias ortológicas de fatores de transcrição relacionados com a regulação da deposição da SCW em cana-de-açúcar**

TIPO	Extremidade, (posição do domínio do TF)	SEQ. COMPLETA	Nome nas espécies de <i>Saccharum</i>	CLONE EM CANA-DE-AÇÚCAR	CLONE EM SORGO ou MILHO	NOME EM <i>A.thaliana</i> ou <i>Zea mays</i>	Obs.
MYB (R2R3 type)	N-terminal	SIM	<b>ShMYB42</b>	Sh_241N24_t000050 (SGH)	GRMZM2G419239_T01(Milho) Sobic.007G177100.1(Sorgo)	ZmMYB42	Desenho do primer na região 3' Repressor
MYB (R2R3 type)	N-terminal	SIM	<b>ShMYB31</b>	Sh_213J23_t000050 (SGH)	GRMZM2G050305_T01(Milho) Sobic.002G279100.1(Sorgo)	ZmMYB31	Desenho de primer na região 3' Repressor
MYB (R2R3 type)	N-terminal e C-terminal	SIM	<b>ShMYB61</b>	Lócus 5776 (Vicentini et al.2015)	Sobic.003G136600.1	AtMYB61	Desenho de primers no meio da sequência codificante, sem comprometer domínios MYB Ativador
MYB (R2R3 type)	N-terminal	SIM	<b>ShMYB85</b>	Sh_250M06_t000170 (SGH)	Sobic.002G275500.3	AtMYB85	Desenho de primers na região 3'. O tblastx no Sucest revela que tanto AtMYB42 e AtMYB85 voltam como um mesmo clone Ativador
MYB (R2R3 type)	N-terminal	SIM	<b>ShMYB46/83</b>	SCUTRZ3103F02 (SUCEST) lc SP803280_c78017_g1_i1 (CTBE)	Sobic.008G112200.1	AtMYB46/83	O tblastx no Sucest revela que tanto AtMYB46 e AtMYB83 voltam como um mesmo clone em cana. Duas sequencias são necessárias para montar um clone completo Ativador

MYB (R2R3 type)	N-terminal	SIM	<b>ShMYB58/63</b>	Sh_227K15_t000030 (SGH)	Sobic.004G273800.1	AtMYB58/63	O tblastx no Sucest revela que tanto AtMYB58 e AtMYB63 voltam como um mesmo clone em cana. Uma sequencia é necessária para montar um clone completo Ativador
MYB (R2R3 type)	N-terminal	SIM	<b>ShMYB52/54</b>	Lócus 4916 (Vicentini et al.2015)	Sobic.001G110900	AtMYB52/54	O tblastx no Sucest revela que tanto AtMYB52 e AtMYB54 voltam como um mesmo clone em cana. Uma sequencia é necessária para montar um clone completo Ativador
NAC	N-terminal	SIM	<b>ShNST1</b>	SCSBST3096G06 (SUCEST) ctg75497 (NCBI)	Sobic.007G0181005	AtNST1/NST2 AtNST3/SND1	Desenho do primer entre as regiões 5'UTR e CDS.O tblastx no Sucest revela que tanto AtNST1,AtNST2e AtNST3 voltam como um mesmo clone em cana. Duas sequencias são necessárias para montar um

							clone completo Ativador
NAC	N-terminal	SIM	<b>ShVNI2</b>	Sh_212D06_t000050 (SGH)	Sobic.008G094700.1	AtVNI2	Desenho primer só na região 5'UTR. É um repressor. Uma sequência é necessária para montar um clone completo
KNOX	C-terminal	SIM	<b>ShKNAT7</b>	lcl SP803280_c97318_g1_i2 (CTBE) ctg30250(NCBI)	Sobic.001G526200.1	AtKNAT7	Duas sequencias são necessária para montar um clone completo. Desenho de primers entre as regiões 3'UTR e CDS. Repressor.
AP2	N-terminal	SIM	<b>ShSHN1</b>	Lócus 27034 (Vicentini et al.2015)	Sobic.004G084600.1	AtSHN1	Sequência completa. Desenho de primers na região 5' UTR. Repressor
WRKY	C-terminal	SIM	<b>ShWRKY12</b>	Sh_226P19_t000020(SGH)	Sobic.006G166300.2	AtWRKY12	Desenho do primer na região 5'. É um Repressor.

**Tabela S3A - Sequências dos pares de primers utilizados na amplificação por RT-qPCR e tamanho do amplicon de Fatores de transcrição relacionados com a regulação da SCW em espécies de *Saccharum***

<b>Gene</b>	<b>Forward 5'--&gt;3'</b>	<b>Reverse 5'--&gt;3'</b>	<b>Amplicon (bp)</b>
<b><i>ShMYB42</i> (RT-qPCR)</b>	TGACCAACATCTCGTTCCAG	CAGAGGTCGAGGTTGAGGTC	<b>136</b>
<b><i>ShMYB31</i> (RT-qPCR)</b>	GCTGGAGGAGCACAACAAG	CTGATGCAGAGGTCCAGGTT	<b>122</b>
<b><i>ShMYB61</i> (RT-qPCR)</b>	TGGCAACAAGAATGGTGGTA	TGAGCCCTGTTGATGTTGAG	<b>125</b>
<b><i>ShMYB85</i> (RT-qPCR)</b>	TTGATGTGGACGAGTTCAGC	TCATGTAGCCATCGACCAAG	<b>132</b>
<b><i>ShMYB46/83</i> (RT-qPCR)</b>	TCATCAGCCAAGTGAACAGC	TTTCATCAGCTCCTCCAGGT	<b>119</b>
<b><i>ShMYB58/63</i> (RT-qPCR)</b>	ACCCTGACGTACTGGACCAC	CACCACCAGGAGTTCAGGTT	<b>125</b>
<b><i>ShMYB52/54</i> (RT-qPCR)</b>	CATGATAGCAACCAGCAGGA	CAATGCTCGAACCACACTTG	<b>124</b>
<b><i>ShNST1</i> (RT-qPCR)</b>	CGGTCTCGCCTTCTACTCC	AAGCTCCACAGGTCGTCGT	<b>90</b>
<b><i>ShVNI2</i> (RT-qPCR)</b>	TGGGGTTCATCGACTTCTTC	CCTCTGCTGGTGGTCTCCT	<b>143</b>
<b><i>ShKNAT7</i> (RT-qPCR)</b>	GACCACCTCCCCCTCATC	GAGCACAGCAGCATCAGGTA	<b>146</b>
<b><i>ShSHN1</i> (RT-qPCR)</b>	CACAGTCCACATCAGCAACA	GCTCAGCAACTCTTCGATCA	<b>103</b>
<b><i>ShWRKY12</i> (RT-qPCR)</b>	TGCCTTCCTGCTAGCCTCTA	CATCACGCCATGATTATTCG	<b>138</b>

**Tabela S3B - Sequências dos pares de primers utilizados na amplificação por RT-qPCR e tamanho do amplicon de genes da rota biosintética da lignina em espécies de *Saccharum***

Gene	Forward 5'-->3'	Reverse 5'-->3'	Amplicon (bp)
<b><i>ShCOMT</i> (RT-qPCR)</b>	GAGGACAAGGACGGCAAGTA	ACCGCGTCCTTGAGGTAGTA	<b>154</b>
<b><i>ShHCT</i> (RT-qPCR)</b>	TCAGACGACACCGCCTTC	GTCCGCCACGAGTTGAT	<b>137</b>
<b><i>ShF5H</i> (RT-qPCR)</b>	AGACGCAGGACGGAGTGTT	AAGAGCTTCATCACGCACAG	<b>138</b>
<b><i>Sh4CL</i> (RT-qPCR)</b>	AGCCGTTCCAGGTCAAGTC	ACTCGGGGTCGTTCAAGTA	<b>161</b>
<b><i>ShCAD A</i> (RT-qPCR)</b>	TCAAGAACGACTGGGGAAAC	GCAGGAGCCGACGAAGTA	<b>140</b>
<b><i>ShCAD B</i> (RT-qPCR)</b>	ATCAGCTCGTCGTCCAAGAA	CGATGATGTAGTCCAGCGAGT	<b>120</b>
<b><i>ShC3H</i> (RT-qPCR)</b>	TAGTGCGGAACCACCTTTCT	TTCGTCAATGTCACCGTTTG	<b>94</b>
<b><i>ShC4H</i> (RT-qPCR)</b>	CGTTCCTCCGTGGGTATCT	CATCACCTTCTTGCGTTCCT	<b>97</b>
<b><i>ShCCoAOMT A</i> (RT-qPCR)</b>	CTCGTGACCGACAAGCAC	AGGGAGTAGCCCGTGAACA	<b>130</b>
<b><i>ShCCoAOMT B</i> (RT-qPCR)</b>	ACGCCGACAAGGACAACACTAC	GCGGTAGAAGCGGATGTACT	<b>151</b>
<b><i>ShCCR</i> (RT-qPCR)</b>	ACTGTCAAGGGAACCGTCAG	CAGATGGCGTCGTAGTCCAG	<b>121</b>
<b><i>ShGAPDH</i> (RT-qPCR)</b>	TTGGTTTCCACTGACTTCGTT	CTGTAGCCCCACTCGTTGT	<b>100</b>

**Figura S2 – Alinhamento de sequências de aminoácidos de Fatores de transcrição relacionados com a regulação da deposição SCW (e lignina) nas espécies *Saccharum* spp., *Panicum virgatum* (Pv) e *Sorghum bicolor* (Sb), *Brachypodium distachyon* (Bd), *Zea mays* (Zm), *Oryza sativa* (Os), *Arabidopsis thaliana* (At) e *Populus trichocarpa* (Pt)**

### ShMYB61

BdMYB61	MGRHS CCYKQKLRKGLWSPEEDEKLLNHINKHGHGCWSTVPKLAGLQRCGKSCRLRWMNY	60
OsMYB61	MGRHS CCYKQKLRKGLWSPEEDEKLMNHITKHGHGCWSTVPKLAGLQRCGKSCRLRWINY	60
PvMYB61	MGRHS CCYKQKLRKGLWSPEEDEKLMNHITKHGHGCWSTVPKLAGLQRCGKSCRLRWINY	60
SbMYB61	MGRHS CCYKQKLRKGLWSPEEDEKLMNHITKHGHGCWSTVPKLAGLQRCGKSCRLRWINY	60
<b>ShMYB61</b>	MGRHS CCYKQKLRKGLWSPEEDEKLMNHITKHGHGCWSTVPKLAGLQRCGKSCRLRWINY	60
ZmMYB61	MGRHS CCYKQKLRKGLWSPEEDEKLMNHITKHGHGCWSTVPKLAGLQRCGKSCRLRWINY	60
AtMYB61	MGRHS CCYKQKLRKGLWSPEEDEKLLTHITNHGHGCWSSVPKLAGLQRCGKSCRLRWINY	60
PtMYB61	MGRHS CCYKQKLRKGLWSPEEDEKLLNYITKHGHGCWSSVPKQADLQRCGKSCRLRWINY	60
	*****:.*:*****:*** * .*****:.*	
BdMYB61	LRPDLKRGAFQEEEDLIELHAVLGNKWSQIATRLPGRDNEIKNLWNNSICKKLRQKG	120
OsMYB61	LRPDLKRGAFSQEEEDLIELHAVLGNRWSQIATRLPGRDNEIKNLWNNSICKKLRQKG	120
PvMYB61	LRPDLKRGAFSQEEEDLIELHAVLGNRWSQIATRLPGRDNEIKNLWNSSICKKLRQKG	120
SbMYB61	LRPDLKRGAFSQEEEDLIELHAVLGNRWSQIATRLPGRDNEIKNLWNSSICKKLRQKG	120
<b>ShMYB61</b>	LRPDLKRGAFSQEEEDLIELHAVLGNRWSQIATRLPGRDNEIKNLWNSSICKKLRQKG	120
ZmMYB61	LRPDLKRGAFSEEEEDLIELHAVLGNRWSQIATRLPGRDNEIKNLWNSSICKKLRQKG	120
AtMYB61	LRPDLKRGAFSPEEENLIELHAVLGNRWSQIASRLPGRDNEIKNLWNSSICKKLRQKG	120
PtMYB61	LRPDLKRGAFSQEENLIELHAVLGNRWSQIAAQLPGRDNEIKNLWNNSICKKLRQKG	120
	***** . :*:*:*****:*****:* ********** .*****:.*:	
BdMYB61	LDPNTHKPLADEADR-RKVAPTMSTERTSESSDQSSAGALGNLSHLLT-ETAQSPELL	178
OsMYB61	IDPNTHKPL-AEVDR-SKATPTISNDRTSESSDVPSSGVALHNLSHLLS-ETAQSSELL	177
PvMYB61	IDPNTHKPLAE--VD-SKAAPTISTERTSESSDDPSSGGAIRNLSHILS-ETAQSPELL	176
SbMYB61	IDPNTHKPL-AEVEH-SKAAPTISTERTSESSDVPSSGGALGNLSHLLS-ETAQSPELL	177
<b>ShMYB61</b>	IDPNTHKPL-AEVEH-SKAAPTISTERTSESSDVPSSGGALGNLSHLLS-ETAQSPELL	177
ZmMYB61	IDPNTHKPLLEVEER-SKAAPTTSTERTSESSDVPSSGGALGNSSHLLL-SETAQSPQL	178
AtMYB61	IDPNTHKPISEVESFSDKDKPTTSNNKRS-----GN-----	151
PtMYB61	IDPNTHKPLSEVENCKEQOPTADKNEK-----VSNVSNELNLEAATLQPPAIS	171
	:*****: * ** . . . . .	
BdMYB61	---PVL-GKQRAKTPSLTCVKVPPKEFFLDQLASGHESLP-SCHSSLPMPNFPFQPPCY	233
OsMYB61	---PVKVKPRTQAPGLARLKVPPKELFLDQLTSGHENLP-SCRSSGPIPFPFQQLCY	233
PvMYB61	---PGL-GKHRKETTLGHLRVPKELFLDQLVSGHDNLP-SCRSTGPIPFPFQQLICY	231

SbMYB61	---PVL-GKHRKETTSLAHLRVPPKELFLDQLVSGHENLT-SCRSTGPIPNFPFQQLMCY	232
ShMYB61	---PVL-GKHRKETTSLAHLRVPPKELFLDQLVSGHENLT-SCRSTGPIPNFPFQQLMCY	232
ZmMYB61	---PAL-GKHRRETTSLAHLRVPPKELFLDQLVSGHENLT-GCRTAGPVPNFPFQQLMCY	233
AtMYB61	-----DHKSPSSSSATNQDFFL-----ERPSDLSDFGFQKLNFN	186
PtMYB61	SSSKINNSKDRNSSSSNMNTNPTTKEFFLDRFGTSHESPASCSPDLMGYFPFQKLDYK	231
	. . . : : * * * :	
BdMYB61	SNGFSSRDGGSKNSLWFWNQNESSGSA---IST--VMPPVSPSTLSTST-----V----L	279
OsMYB61	NNDFNSMDVGNRNSL-WYNQNESSSST---IST--VMPPVSPSTLSTST-----G----L	278
PvMYB61	SNELGSKPGGSTNSL-WFNQNESSCST---VST--VMAPVSPSTLSTST-----R----L	276
SbMYB61	NNEFGHKNGGSNSNF-WFNQNESSGST---IST--VMPPVSPSTLSTST-----G----L	277
<b>ShMYB61</b>	NDEFGNKNGGSNSNF-WFNQNESSGST---IST--VMPPVSPSTLSTST-----G----L	277
ZmMYB61	SNEFGNKNGASNSL-WFNHNETS GST---IST--VMPPVSPSTLSTST-----G----L	278
AtMYB61	SNLGLSVTT---DSSLCSM-----IPQFSPGNMVGSVLQTPVCVKPSI	227
PtMYB61	PSIGLSMNP---NTTLCFNPNSSSEMI SHEFNCSMTPPTILP-SVSTSMFQTPIRVKPSV	287
	. : . * . : * :	
BdMYB61	----NLSPDNPHSGGTGIHSAQFYWDTANPSSSCSTRSSGNNNLGFELQSTSSLENSVF	335
OsMYB61	----NPSPDNANSRGTGIHNSQFYWDTNPNSSSSSTGSSGNNLGFELQSTSSLETNIF	334
PvMYB61	----NRSPEKPHSGGTGIQSTQFYWDTANPSSSSSKGSSGNNLGFELQSTSSILENSIF	332
SbMYB61	----NGSPDNPHSGGTGIQNTQFYWETANPSSSSSRGSSGSNSLGFERQSTSSLENSVF	333
<b>ShMYB61</b>	----NRSFDNPHSGGTGLQNTQCYWETANPSSSSSRGSSGSNGLGFELQSTSSLENSVF	333
ZmMYB61	----NRSFDNPHSGGTGIQSTQFYWDTANPSSCSSRGSSGSNGLGFELQSTSSLENSVF	334
AtMYB61	SLPPDNSSSPI SGGDHVKLAAPNWEFQTN-----NNNTSNFFDNGGF	270
PtMYB61	SLPSDHNPSVGS CD----VNGVQNW EASSFSNNGS--GSNGSSS IELQSN TNFYESSAF	341
	: . . . * : . : . . . : . . *	
BdMYB61	PWAELAQDKDS-----QVHLEELKWPPELLHGTFFPETSTAMQNLSQSLYEDVVKAES	387
OsMYB61	PWSDLAPEKDS-----QAQLEELKWPDLLHGTFFSEMPAPMQNLSQSLYEDVVKAES	386
PvMYB61	PWTDLSPDKNS-----HLEELKWPDLLHGTFFDTPATMQNLSQSLYEDVVKAES	382
SbMYB61	PWTDLTPDKNS-----QVHLEELKWPDLLHGTFFDTPATMQNLSQSLYEDVIKAE S	385
<b>ShMYB61</b>	PWTDLTPDKNS-----QVHLEELKWPDLLHGTFFDTPATMQNLSQSLYEDVIKAE S	385
ZmMYB61	PWTELTDPDKNS-----QVHLGEELKWPDLLHGTFFDTPATMQNLSQSLYEDVIKAE S	386
AtMYB61	SWSIPNSSTSSSQ----VKPNHNFEELKWEYLNT PFFIGSTVQS QTSQPIYIK-SETDY	325
PtMYB61	SWGLANCGKSGEESHRLRSLENDTAEDIKWEYLNTFFL GSTIQNQTSQH VYSE-VKPET	400
	* . . . . * : : * * : * : . : * * : * . : :	
BdMYB61	QFNMEGLCAAWSQNLQPQHLQVLVTDMYDKDLQRMSSPFENI-	429
OsMYB61	QFNMEGLCAAWSQNLQPQHLQVVS DMYDKDLQRM SLSFENI-	428
PvMYB61	QFNMEGLCAAWSQNLQPQHLQVVS DLYDKDLQRM SLSFENI-	424
SbMYB61	QFNIEGLCAAWSQNLQPQHLQVVS DLYDKDLQRM SLSFENI-	427
<b>ShMYB61</b>	QFNIEGLCAAWSQNLQPQHLQVVS DMYDKDLQRM SLSFENI-	427
ZmMYB61	QFNIEGLCAAWSQNLQPQHLQV-SDLYDKDLQRM SLSFENI-	427
AtMYB61	LANVSNMTPWSQENLG--TTETS DVFSKDLQRM AVSFGQSL	366



AtMYB85 LIDNIYDDDELFSYLWSDETTKD-----EASWSDSNFVGGTLYDHNISGADADFPWSP 243  
PtMYB85 LD-SICNDEMILLNSLWMEPP-----LVDASWNNIIPPAAN-----TNDDTGYPSWEE 232  
BdMYB85 NGGGSHDQDPLVKWLLLEEDDLPTGGGIGEKWLDFTAGDDVDDLGSIAAT-----MPWDV 241  
OsMYB85 SSSSSHQDPLVKWLLLEEDLL-PTG---DEPWLNFASNDVDEFSSIAATGATPALPWDV 250  
PvMYB85 ---SSHDQDPLVKWLLLEED-P-PTG---DEPWLNFSGVDVDELSSIAAG--PE-LPWDG 238  
ZmMYB85 SGGGGHDQDPLVKWLLLEE-P-ATG---DEAWLNFTGSVDVDEFSSIAAG--PELLPWDG 241  
SbMYB85 SSSGGHDQDPLVKWLLLEE-P-PTG---DEPWLNFSGVDVDEFSSIAAG--PELLPWDG 278  
**ShMYB85** SSGGGHDQDPLVKWLLLEE-P-PTG---DEPWLNFSGVDVDEFSSIAAG--PELLPWDG 244  
. .:: \*.. \* :: \* : :

AtMYB85 ERINDEKMFLDYCQDFGVHDFGF----- 266  
PtMYB85 -----NYTWLSDCQDFGVHDFGLECFDSMELSALST-LEMEHKH 270  
BdMYB85 G----ATDWLLDYQDFGVGDSSLLSM-SMVDSSN--GSNF---- 274  
OsMYB85 G---MTDWLLDYQDFGMGDSSLVVDASMVNSSNGS--NF---- 285  
PvMYB85 -----ATDWLLDYQDFGLGDSS-LVDGYMINSSNG--AN----- 269  
ZmMYB85 -----ATDWLLDYQDFGLGDSS-LVDGYMVNNSNGAKF---- 275  
SbMYB85 -----ATDWLLDYQDFGLGDSATLVDGYMVNN-SSNGAKL---- 312  
**ShMYB85** -----ATDWLLDYQDFGLGDSATLVDGYMVNN-SSNGAKF---- 278  
:\* \*\*\*: \* .

## ShMYB46/83

PvMYB46/83 MRKPEGPAAANS-----CNNA--GAAAKLRKGLWSPEEDEKLVAYMLRSGQGSWS 48  
ZmMYB46/83 MRKPECPAANSS-----N-A--GAAAKLRKGLWSPEEDERLVAYMLRSGQGSWS 47  
SbMYB46/83 MRKPECPAAANS-----GNAGGAAAATKLRKGLWSPEEDERLVAYMLRSGQGSWS 50  
**ShMYB46/83** MRKPECPAAANS-----GNAGV--AAVKLRKGLWSPEEDERLVSYMLRSGQGSWS 48  
OsMYB46/83 MRKPDCGGGGAAKGGGVLGVAGGNAAVVGKVRKGLWSPEEDEKLVAYMLRSGQGSWS 60  
BdMYB46/83 MRKPECPTTKAASGPGN----SCSNVVAASAAKLRKGLWSPEEDERLVAYMLRSGQGSWS 56  
AtMYB46/83 MRKPEVAIAAST-----HQVKKMKKGLWSPEEDSKLMQYMLSNQGQGSWS 44  
PtMYB46/83 MRKPEASGKNNV-----NNINKFRKGLWSPEEDDKLMNYMLNNGQGSWS 44  
\*\*\*\*: \*.:\*\*\*\*\*.:\*: \*\* .\*\*\*.\*\*

PvMYB46/83 DVARNAGLQRCGKSCRLRWINYLRPDLKRGAFSPQEEELIVSLHAILGNRWSQIAARLPG 108  
ZmMYB46/83 DVARNAGLQRCGKSCRLRWINYLRPDLKRGAFSPQEEELIVSLHAILGNRWSQIAARLPG 107  
SbMYB46/83 DVARNAGLQRCGKSCRLRWINYLRPDLKRGAFSPQEEELIVSLHAILGNRWSQIAARLPG 110  
**ShMYB46/83** DVARNAGLQRCGKSCRLRWINYLRPDLKRGAFSPQEEELIVSLHAILGNRWSQIASRLPG 108  
OsMYB46/83 DVARNAGLQRCGKSCRLRWINYLRPDLKRGAFSPQEEELIVNLHAILGNRWSQIAARLPG 120  
BdMYB46/83 DVARNAGLLRCGKSCRLRWINYLRPDLKRGAFSPQEEELIVNLHAILGNRWSQIAARLPG 116  
AtMYB46/83 DVAKNAGLQRCGKSCRLRWINYLRPDLKRGAFSPQEEELIRFHSILGNRWSQIAARLPG 104  
PtMYB46/83 DVARNAGLQRCGKSCRLRWINYLRPDLKRGAFSPQEEEMI IHLHSLGNRWSQIAARLPG 104  
\*\*\*:\*\*\* \*\*\*\*\*:.\*: .\*:\*\*\*\*\*:\*\*\*

PvMYB46/83 RTDNEIKNFWNSTIKKRIKNSSSASSPAAATDCASPEPNNKGVAGFDISGAASCPDLAGL 168  
ZmMYB46/83 RTDNEIKNFWNSTIKKRLKNSSAASSPA-ATDCASPEPNNKVAA-----AGSCPDL SVL 160

SbMYB46/83 RTDNEIKNFWNSTIKKRLKNTSATSSPA-ATECASPEPNN-KVA-----AGSCPDLAGL 162  
**ShMYB46/83** RTDNEIKNFWNSTIKKRLKNSSAASSPA-ATDCASPEPNNKVV-----AGSCPDLAGL 161  
OsMYB46/83 RTDNEIKNFWNSTIKKRLKISSSSASPATTTDCASPPEHKLGAUV-DLAGGGGATDD--- 176  
BdMYB46/83 RTDNEIKNFWNSTIKKRLKVSSAASSPATTTTECASPEPKLDVGCCLDLA-----S--- 166  
AtMYB46/83 RTDNEIKNFWNSTIKKRLKMSDTSNLINSSSSSPNTASDSSS-----NSASSLDIK 156  
PtMYB46/83 RTDNEIKNFWNSTIKKRLKNLQSSNASPNTSDSSS-----EPK 143  
\*\*\*\*\*:\*. : :..:

PvMYB46/83 DRQVDGGHHHA---M-VTTGLWM---VDSSSSTSSSTSPMQGRPPSSAM-----VAAV 215  
ZmMYB46/83 DHQDGGHH---HAMTT-TTAGLWM---VDSSSCTSSSTSPMHQFQRPTTMM-----AAAVA 209  
SbMYB46/83 DHQDGGHHHHHHLMTT-TTTGLWM---VDSSSCTSSSTSPMHQRPPTTAIMAAAATA 218  
**ShMYB46/83** DHQDGGHHH---HAMTT-VTTGPWM---VDSSSCTSSSTSMHQRPPLPTATMAAAVAAA 215  
OsMYB46/83 -VVVGTAN-----A-AMKSMWV---DSSS-SSSSSSSMQSRPS-----IMAAA-AA 216  
BdMYB46/83 -LEDGGHH-----A-MIKSTWR---MMDSSSSTSSSSSMQTRPS-----SAMAAASR 208  
AtMYB46/83 -DIIGSF-----MS-LQEQGFVNPSLTH----- 177  
PtMYB46/83 -DVMGGL-----MSTMQEQGIFSMNMDPSMSSSSSLATSMK-----AMIL 182

PvMYB46/83 TRSYGGLLPLPDQLRGMAADASPAGFFHGHAAPFKHQAVVASLHGGYYGGSAPHHHGMM 275  
ZmMYB46/83 SGSYGGLVFPFDQVRGVVAD---TGGFFHGHAAPAFKHQVAALHGGYYGGSAPRHHGMT 266  
SbMYB46/83 TRSYGGLVFPFDQLRGVMADASPGRFFHGHAAPFKHQVAALHGGYFYGSTPPHHHGMM 278  
**ShMYB46/83** SRSYGGLVFPFDQLRGVVADASPGGFFHGHAAPAFKHQVAALHGGYYGGSSTPPHHHGMM 274  
OsMYB46/83 GRSYGGLLPLPDQVCGVDTSPP---PPFFHDHSI-SIKQAYYGS-----TGAHH---- 261  
BdMYB46/83 SYSGGLLPLPDQVCGGGGETQA-PPLFFHDHHPLSFKHAS-AL-----HGGSY---- 254  
AtMYB46/83 -----IQTNPF-----PTG-NMI 190  
PtMYB46/83 NTMMDPLLPMLDYDHGL-----NMYGGASGYE-SIT 212  
:

PvMYB46/83 AT-----MEGGGCFMRGEGL-----F-G--VPPLLEAMSSQD--QDHHGQALMASSG 318  
ZmMYB46/83 TTTTVALEGGGCFISGEGM-----L-G-VPPLLEPMSAAL-EQDQGG-TLMASSG 315  
SbMYB46/83 AT-----MEGGGCFMRGEDM-----FVG-VVPPLLDPMSAAAQEQQGQGLMASSG 324  
**ShMYB46/83** MA-----MEGGGCFMRGEGM-----F-G-VVPPLLEPMSAAAQEQQGQA-LMASSG 318  
OsMYB46/83 HHHAIATMDGS--SLIGDHHHH--SSSILFGGASVPPLLDHQTLDDDDHHP---NKTG-- 313  
BdMYB46/83 YHHGVAMEDGGAGCFMGEVVAGEESALF---NVPPPLQLEPAMAAAQDQT---LMASR- 307  
AtMYB46/83 SHP---CNDDFTPYV---DGIYGVNA-GVQGEYFPPLE---CEEGDWYNA--NI----- 233  
PtMYB46/83 APP---CMAQVGVLSGDHGFYGEI-FEGINVEIPPLESVSCMEENAKTQ--NIQ---- 262  
\*:

PvMYB46/83 GGNNNNPKNNSSNTTEATATTTTTVSNNESNITDN---TKENINTMS--LVNSSSNVA 373  
ZmMYB46/83 NNNP--KNN-SSNNT--DTTTLTSLNNEVTDTTT-KDNTNTIS--QVNS-GSN-- 364  
SbMYB46/83 SNNA--KNNNSNNT--ETTTTTLTLNNEVTDTTT-KDNTNTIS--QVNS-GSNVA 377  
**ShMYB46/83** NNP---N-NSSNNT--TETTTTTLTLNNEVTDTTT-KDNTNTIS--QVNT-GSNVA 367  
OsMYB46/83 -----SNTTAATLSSNITDNTSNKNNSDNNNNISSCCISLMNSSSN 356

BdMYB46/83 -----NKS-----TTPEATTLSSNNGSNITDNDNSK-----NNNVSSS 341  
 AtMYB46/83 -----NNH-----LDELNTNGS 245  
 PtMYB46/83 -----DNNTDKYS--YSSPVNSLYHKNCNITSNK-----TDSIAADQM 299

PvMYB46/83 AVYWEGAH-----QQYMSRNVMHGEWDLEELMKDVSSLPFLDFQVE- 414  
 ZmMYB46/83 NVYWEGAR-----QQYMSRNVMHGEWDLEELMKDVSSLPFLDFQVE- 405  
 SbMYB46/83 AVFWEGAH-----QQYMSRNVMHGEWDLEELMKDVSSLPFLDFQVE- 418  
**ShMYB46/83** AVFWEGAH-----QQYMDRNVMHGEWDLEELMKDVSSLPFLDFQVE- 408  
 OsMYB46/83 MIYWEGHHQQQQQHQLQQQQQHMSRNVN-GEWDLEELMKDVSSLPFLDFQVE- 409  
 BdMYB46/83 VVYWPEQL-----HGHQQQQQHMSRNVN-GEWDLEELMKDVSSLPFLDFQVE- 388  
 AtMYB46/83 GNAPEGM-----R-PVEEFWDLQLMNTEVPSFYFNFKQSI 280  
 PtMYB46/83 GNLWHGS-----EELKVGGEWDLEELMKDVSAFPFLDFQ--- 332

\*\*\*::\*\* :::\*

**ShMYB58/63**

BdMYB58/63 MGKGRAPCCAKVGLNRGSWTPQEDMRLIAYIQKHGHANWRALPKQAGLLRCGKSCRLRWI 60  
 OsMYB58/63 MGRGRAPCCAKVGLNRGSWTPQEDMRLIAYIQKHGHANWRALPKQAGLLRCGKSCRLRWI 60  
 PvMYB58/63 MGRGRAPCCAKVGLNRGSWTPQEDMRLVAYIQKHGHSNWRALPKQAGLLRCGKSCRLRWI 60  
 ZmMYB58/63 MGRGRAPCCAKVGLNRGSWTPQEDMRLIAYIQKHGHTNWRALPKQAGLLRCGKSCRLRWI 60  
**ShMYB58/63** MGRGRAPCCAKVGLNRGSWTPQEDMRLIAYIQKHGHSNWRALPKQAGLLRCGKSCRLRWI 60  
 SbMYB58/63 MGRGRAPCCAKVGLNRGSWTPQEDMRLIAYIQKHGHANWRALPKQAGLLRCGKSCRLRWI 60  
 PtMYB58/63 MGKGRAPCCDKS QVKRGPWS PAEDLRLIAFIQKHGHENWRALPKQAGLLRCGKSCRLRWI 60  
 AtMYB58/63 MGKGRAPCCDKTKVKRGPWSHDEDLKLISFIHKNGHENWRS LPKQAGLLRCGKSCRLRWI 60

\*\*.\*\*\*\*\* \* :::\* \* : \*\*:::\*. \* \*\* \*\*\*.\*\*\*\*\*

BdMYB58/63 NYLRPDLKRGNFTAEEEEETLIKHLHAMLGNKWSKIAACLPGRTDNEIKNVWNTHLKKRVAA 120  
 OsMYB58/63 NYLRPDLKRGNFTADEEDTIIKHLHGLLGNKWSKIASCLPGRTDNEIKNVWNTHLKKRVSQ 120  
 PvMYB58/63 NYLRPDLKRGNFTVEEEETIIRLHGMLGNKWSKIAACLPGRTDNEIKNVWNTHLKKKVAQ 120  
 ZmMYB58/63 NYLRPDLKRGNFTDEEEAIIRLHGLLGNKWSKIAACLPGRTDNEIKNVWNTHLKKKV-- 118

**ShMYB58/63** NYLRPDLKRGSF TAE E E E T I I R L H G L L G N K W S K I A A C L P G R T D N E I K N V W N T H L K K K --V 118  
 SbMYB58/63 NYLRPDLKRGNF TAE E E E T I I R L H G L L G N K W S K I A A C L P G R T D N E I K N V W N T H L K K K K Q V 120  
 PtMYB58/63 NYLRPDVKRGNF SEVE E E D T I I K L H Q T L G N K W S K I A S H L P G R T D N E I K N V W N T H L K K K L A C 120  
 AtMYB58/63 NYLRPDVKRGNF SAE E E E D T I I K L H Q S F G N K W S K I A S K L P G R T D N E I K N V W H T H L K K R L S S 120  
 \*\*\*\*\*:\*\*\*.\*: \*\*::\*:\*\* :\*\*\*\*\*: \*\*\*\*\*:\*\*\*\*\*:

BdMYB58/63 ASGG--EKKNKNTKNNKESGAKRQKKPDLPPAPSPSQSSSTTTTTTTNCSSGDSAGEQ 178  
 OsMYB58/63 REKPGDTKKKGKAADASDDA-----DAHSPSSSASS--STT-TAANNNSGDTAGE 168  
 PvMYB58/63 --RE--QKKKAGAAK--NDGAASGGDAGTPGTDSSSA-SSS-TTT-TTTNNSGGSDSGD 171  
 ZmMYB58/63 AQRE--KK-----KAGAGSGDAGTPATAPLSSATSS---T-TTHNSS-GGSDSGD 161  
**ShMYB58/63** APRE--KQKAGAA---DKNDGAASGDAGTPATAASSSVSS---T-TTHNSSGGSDNSGD 169  
 SbMYB58/63 APRE--EQKAGAAADKNDGGAASGEAGTPATAASSSASSSTTTT-TTNNSAGSDNSGD 177  
 PtMYB58/63 KDDGE--HS-----EGDES---KSSSTSSSSS-SSS-T-I-----MSSGKRALEM 158  
 AtMYB58/63 ETNLN--AD-----EAG-S---KGSLEEN----- 140

BdMYB58/63 -----QSNTSKETDALDNIMEIPTLDALGFDFDMLVQDTPVPAQT-----SYCVPPP 226  
 OsMYB58/63 -----QCGTSKEPENVDV--SFFEQDI-D- ISDMLVDAPT---EA-----PLVA--A 206  
 PvMYB58/63 -----QCGTSKEPDAAA---VSPLHDI-G-VWEMLVDAPA---AA-----Q 204  
 ZmMYB58/63 -----QCGTSREPD-ATDVCPLQLEDM-D-VSDMLVDGAP---PAA-----QPMP--S 201  
**ShMYB58/63** -----QCGTSREPDDAIDVSLRLEDI-D-VSDMLVDAPP---A-A-----QPCQ--A 209  
 SbMYB58/63 -----QCGTSREPD-VIDASLLQLEDI-D-VSDMLVDAPP---AAA-----QPCQ--A 217  
 PtMYB58/63 ELDEQKNQGFSTQKPRILE-----NEEDSSPRGVSNNQFKPAMKPN 199  
 AtMYB58/63 -----SQESSPNA----- 148

BdMYB58/63 AAAASPCSSA--S-----PPCA-----ADGELLDLP 250  
 OsMYB58/63 PMPPSPCSSS--S-----LTTTTCVG-----AVSDELLDLP 235  
 PvMYB58/63 PMLSSSCSSS--S-----L-TTCV-----GGAEDLIELP 230  
 ZmMYB58/63 PS-----SSS--S-----L-TTCV-----GGVEELIELP 222  
**ShMYB58/63** PMLS-PCSSS--S-----LT--TCV-----GGVEELIELP 234  
 SbMYB58/63 PMLSSPCSSS--S-----LT-TTCVG-----GGVEELIELP 245  
 PtMYB58/63 ELSSSSFSNNSSITNSSQADVSEPDGEKTGSFFN-FRGRYNVRNSLEEVNKPEEIVTEI 258  
 AtMYB58/63 ---SMSFAGSNIS-----SKDDDAQISQMFHILTYSEFTGMLQEVDKPE--LLEM 194  
 :. \* :

BdMYB58/63 ELDIVPELWSIIDGDAEAPW----SNA AAC-----HVEAGEDGKEWWLEDLERE 296  
 OsMYB58/63 EIDIEPDIWSIIDGYGGDEP--GDGDATVPCTASP-----GEEGAEWWVENLEKE 283  
 PvMYB58/63 VMDIEPDIWSIIDDESADGSGARHGDAAPCTGAA--VSTSEEEAGEAADDWWLENLEKE 288  
 ZmMYB58/63 VIDIEPEIWSIIDGESAVA--RHGGDAAAPCTGTGTAVSTSEAE-EAAANDWWLENLEKE 279  
**ShMYB58/63** VIEIEPEIWSIIDGESADAPDASGGDATAPCTGTAVVS-TSE--AEEAANDWWLENLEKE 291  
 SbMYB58/63 VIEIEPEIWSIIDGESATDAPDASGGGTAPCTGTAADVSTSEAEAAAAANDWWLENLEKE 305  
 PtMYB58/63 PFESDYDFWNMLDSLSSFQTSIGIQL---QNVEAGQSSRFGDAYNMGEVENKKWLRYLENE 315  
 AtMYB58/63 PFDLDPDIWSFIDGSDSFQQPENRA---L-----QESEDEVDKWFKHLESE 238





BdMYB42 LRPDLKRGNFSEEEDELI IKLHSLGKNSLIAGRLPGRTDNEIKNYWNTHIRKLLSRG 120  
 PvMYB42 LRPDLKRGNFTADEDLIVKLHSLGKNSLIAARLPGRDTDNEIKNYWNTHIKRLLSRG 120  
 OsMYB42 LRPDLKRGNFTADEDLIVKLHSLGKNSLIAARLPGRDTDNEIKNYWNTHIRKLLGGRG 120  
**ShMYB42** LRPDLKRGNFTADEDLIVKLHSLGKNSLIAARLPGRDTDNEIKNYWNTHIRKLLGGRG 120  
 SbMYB42 LRPDLKRGNFTADEDLIVKLHSLGKNSLIAARLPGRDTDNEIKNYWNTHIRKLLGGRG 120  
 ZmMYB42 LRPDLKRGNFTADEDLIVKLHSLGKNSLIAARLPGRDTDNEIKNYWNTHIRKLLGSG 120  
 \*\*\*\*\*: :\*: \*\*:\*\*\*\*\*.\*\*\*\*\*:\*\*\*\*. \*

AtMYB42 IDPATHRGINEAKISDLKTKDQIVKDVSVFKFEETDK-----SGDQKQNKYIRNGLVC 175  
 PtMYB42 IDPATHRPLNEPAQEAST-----TISFSTTTSVKEESLSSVKEESNKEKIIISAAAFI 172  
 BdMYB42 IDPVSHRPTNEH--VSNV-----TISFEAAREEKGS MFR-LDEPKPA-----IGHD 163  
 PvMYB42 IDPVTHRPIADA--ARNV-----TISFQPDAPSQQQLSD--DAEAPP-----PPPP 162  
 OsMYB42 IDPVTHRPNVAAA---AA-----TISFHPQPPPT-----K 147  
**ShMYB42** IDPVTHRPIAGA--GAVT-----NISFQPSNAAGAVAA--QTPQHQ-----PIR 161  
 SbMYB42 IDPVTHRPIADAGAGTVT-----TISFQPNKNAA-VAA--QAPQHQ-----PIK 162  
 ZmMYB42 IDPVTHRRVAG---GAAT-----TISFQPSNTAVAAAA--ETAAQA-----PIK 160  
 \*\*\*.:\*\* :\*\*

AtMYB42 KEERVVVEEKIGPDLNLELRISPPWQNRQREISTCTAS-----RFYME----- 217  
 PtMYB42 CKEEKTVPQERC PDLNLELRISLPCQNQPDRHQA-----FK--TGGSTS 214  
 BdMYB42 PVDWGGKPLKCPDLNLDLCLISPPFQEDPMK-----PVKR-----EAGVG 203  
 PvMYB42 QQQQQLKPPPRCPDLNLDLCLISPPCHKEEED-----QELVKPAAVKREMLQAGHGT-LG 215  
 OsMYB42 EEQLILSKPPKCPDLNLDLCLISPPSCQEEEDDDYE-----AKPAMIVRAPELQR---RRGG 199  
**ShMYB42** AEAT-AVKVPRCPDLNLDLCLISPPCQQEVEEED--ELEL-KPAVVKREVLQAGHCHGGG 217  
 SbMYB42 AVATAVVKVPRCPDLNLDLCLISPPCQQKED-----EELDLKPAVVKREVLQAGH--GGS 215  
 ZmMYB42 AEETAAVKAPRC PDLNLDLCLISPPCQHEDDGE EEEELDLIKPAVVKREALQAGHGHHG 220  
 \*\*\*\*\*: \* \* \* \* . .

AtMYB42 NDMECSSETVKCQTENSSSISYSS---IDISSNVGYDFLGLKTRILDF--RSLEMK 269  
 PtMYB42 LCFACSLGLQNSK-DCSCSVIVGTIGSSSSAGSKTGYDFLGMKSGVLDY--RGLEMK 268  
 BdMYB42 VCFCSLGLPRSS-ECKCS-----NFLGLRTAMLDLFR--SLEMK 239  
 PvMYB42 LCFGCSLGLQKGAAGCTCS-----S-----NSHFLGLRVGMLL-DFRGLEMK 256  
 OsMYB42 LCFGCSLGLQKE---CKS-----GGGAGAGANNFLGLRAGMLDFRS--LPMK 243  
**ShMYB42** ALLRLQPGHPEGRARVQLQ-----Q---QQ-PPPLLGAQGRHARLQRPRDEVK 261  
 SbMYB42 LCFGCSLGIQKGAAGCSCS-----S----SNSHRFLGLRSGMLDF--RGLEMK 258  
 ZmMYB42 LCLGCGLGGQKGAAGCSCS-----N-----GHHFLGLRTSVLDF--RGLEMK 260  
 : . . :\*\* : \*

## ShMYB31

AtMYB31 MGRSPCCEKAHTNKGAWTKEEDERLVAYIKAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60  
 PtMYB31 MGRSPCCEKAHTNKGAWTKEEDRLIAYIRTHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60  
 ZmMYB31 MGRSPCCEKAHTNKGAWTKEEDERLVAHIRAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60  
 OsMYB31 MGRSPCCEKAHTNKGAWTKEEDRLIAYIKAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60  
 BdMYB31 MGRSPCCEKAHTNKGAWTKEEDRLTAYIKAHGEGCWRSLPKAAGLLRCGKSCRLRWINY 60



AtMYB31 TSLLGFRSLEMK 282  
 PtMYB31 SGVLDYRGLEMK 268  
 ZmMYB31 TAMLDFRSLEMK 273  
 OsMYB31 TAMLDFRSLEMK 251  
 BdMYB31 TAMLDFRSLEMK 239  
 PvMYB31 TAMLDFRSLEMK 245  
**ShMYB31** TAMLDFRSLEMN 258  
 SbMYB31 TAMLDFRSLEMK 264  
 :.:\*.:\*.\*\*\*:

## ShVNI2

PvVNI2 MD--AAKEPAVAKPRLPPGFRFRPTDEELVVHYLRRRALASPLPAAVDIPDVRILAHDFS 58  
 ZmVNI2 ME----VKKEGAKATLPPGFRFRPTDEELIVHYLRRRALASPLPAAVDIPDVRILAHDFS 56  
**ShVNI2** ME----V-KDAAKPRLPPGFRFRPTDEELIVHYLRRRALASPLPPAVDIPDVRILAHDFS 55  
 SbVNI2 ME----V-KDAAKPRLPPGFRFRPTDEELIVHYLRRRALASPLPPAVDIPDVRILAHDFS 55  
 BdVNI2 METKESTVAADASRRLPPGFRFRPTDEELVVHYLRRRALSSPLPAAVDIPDVRLLAHDFS 60  
 OsVNI2 -----METTAACKLPPGFRFRPTDEELVVHYLRRRALGSPPLPAVDIPDVRLLAHDFS 53  
 PtVNI2 MEKL--SFVKNQVLRLLPPGFRFHPTDEELVVQYLKRKVFACPLPASI-IPEVDVCKSDPW 57  
 AtVNI2 MDNV--KLVKNQVLRLLPPGFRFHPTDEELVVQYLKRKVCSSPLPASI-IPEFDVCRADPW 57  
 . \*\*\*\*\*:\*\*\*\*\*:\*.\*\*:\*:. . .\*\*\* :. \*\*: . : \*\*

PvVNI2 DLLPPGWSEQERYFFTCKEAKYVKGRANRATGAGYWKATGKEMPVAVAVPARGG----- 113  
 ZmVNI2 DLLPPGFGEQERYFFTCKEAKYVKGRANRATGAGYWKATGKEKPVAVAVRGVQG----- 111  
**ShVNI2** DLLPPGFSEQERYFFTCKEAKYVKGRANRATGAGYWKATGKEKPVAVAI PAAARGGQGQ 115  
 SbVNI2 DLLPPGFSEQERYFFTCKEAKYVKGRANRATGAGYWKATGKEKPVAVAVPATR----- 109  
 BdVNI2 DLLPPGWSEEEERYFFTCESKYVKGRANRATGAGYWKATGKEKPVAVSVPA----- 112  
 OsVNI2 DLLPPGWSEQERYFFTCKEAKYVKGRANRATGAGYWKATGKEKPVAVSVAAAPR----- 108  
 PtVNI2 DL--PGDLEQERYFFSTREAKYPNGNRSNRATGSGYWKATGIDKQIVTSK----- 105  
 AtVNI2 DL--PGNLEKERYFFSTREAKYPNGNRSNRATGSGYWKATGIDKRVTTSR----- 105  
 \*\* \*\* \*:\*\*\*\*\*: :\*:\*\* :\* \*:\*\*\*\*\*:\*\*\*\*\* : :...:

PvVNI2 GGQAQAVLVGMKRSLVFYRGKPPTGSKTDWVMHEYRLAGAGLGPCRG-----AA 162  
 ZmVNI2 QGQGQVVLVGMKRSLVFYRGKPPTGSKTEWVMHEYRLAGAGLAPCRRRHADADGD--GVA 169  
**ShVNI2** GQQGQAVLVGMKRSLVFYRGKPPTGSKTDWVMHEYRLAGAGLSPCRRAAA-QDGDADAAV 174  
 SbVNI2 GGQGQAVLVGMKRSLVFYRGKPPTGSKTDWVMHEYRLAGAGLAPCRRATAQQDGDGDAAV 169  
 BdVNI2 -AKGAAVVVGMKRSLVFYRGKAPSGKKTWVMHEYRLAGAGLAPCRLAQAA-----AN 164  
 OsVNI2 -SQAAVVVGMKRSLVFYRGKPPTGKKTWVMHEYRLAGAGLAPCRRATA-----DHP 161  
 PtVNI2 ----GHQVVGKMTLVFYRGKPPHGTTRTDWIMHEYRLASTETTACNTLKNK--NSTQGPV 159  
 AtVNI2 ----GNQIVGLKMTLVFYKGPVHGSRTDWMIMHEYRLSSSPPS-----S 145  
 :\*\*:\*.:\*\*\*\*\*:\*\* \* \* .\*:\*\*\*\*\*:..:

PvVNI2 ARPAEGWVLCRVFRKKGSASANAAAAAEEQG----SSGDVE-EAEEAEEEDGAAGGARTF 217  
 ZmVNI2 RAPAEGWVLCRVFRKKGSAAVARRP---GDR----SDGEGD-ESAG----GGGCGGRGF 217  
**ShVNI2** SRPAEGWVLCRVFRKKGSAAASAAA---SPG----E--DRS-DGEEA--RESESAGGPGF 222  
 SbVNI2 SRPAEGWVLCRVFRKKGSAAAARP---GDR----SDGDGD-EGEEAGESESAGAGGPGF 221  
 BdVNI2 ANPEEGWVLCRVFRKKGGAAAAADTGD-----I-QDPAPEDDGGGAQSGVRF 210  
 OsVNI2 ARPAEGWVLCRVFRKKGSAAASTASPTA-----DAD-DDDATERADDAAGVRF 210  
 PtVNI2 VVPMENWVLCRIFLKKRGTKNEEENIQV-----GNDNRLP-KLR-----ATEPFS 203  
 AtVNI2 MGPTQNWVLCRIFLKKRAGNKNDDDGDSRNLRHNNNNSSDQIEIITTDQTDKTKPIF 205

\* :.\*\*\*\*\*:\* \*\* .

PvVNI2 IDFFARADADAAGRREEQGQRASSPVVSSSCLTDASHE-QHGREQETTSSRG----- 268  
 ZmVNI2 IDFFARADAAGR-----RRRAPSPAVSSRWLADASPQPEQ----EATSRGP----- 259  
**ShVNI2** IDFFARADAAGR-----RRRAASPVVSSSCLTDASPERQQGREQETTSSRGA----- 268  
 SbVNI2 IDFFARADAAGR-----RRRAASPVVSSSCLTDASPERQQGREQETTSSRGA----- 267  
 BdVNI2 IDFFARAEASR-----RRRA-ASPVSSSCVTDASAEH--CREQQETTSSRGC----- 252  
 OsVNI2 IDFFARADAR-----RRRA-ASPVSSSCVTDASAEH--CREQETTSSRNGGAAAGDA 258  
 PtVNI2 LR----- 205  
 AtVNI2 FDFMRKERTTDLNLLPSSPSSDHASSGVTTTEIFSSSDEETSSCNSFR----- 252

:

PvVNI2 -- 268  
 ZmVNI2 -- 259  
**ShVNI2** -- 268  
 SbVNI2 -- 267  
 BdVNI2 -- 252  
 OsVNI2 SD 260  
 PtVNI2 -- 205  
 AtVNI2 -- 252

**ShWRKY12**

OsWRKY12 -----MYACMEG--SQLETACLPAALYAPLCPYTPPS 30  
 BdWRKY12 -----MEG--TSRLEACLPSLYALDPYTAPPPP 26  
 PvWRKY12 -----MEG--SSQFETCLPSLF-----SPCAL 20  
 ZmWRKY12 MPPNTTLTHPIDQLITTTTHTPHGRSILLYACMEGSSQLLETCLPASLYAVT-PPPCAH 59  
**ShWRKY12** -----MEGS-SQLETCLPAS-LYAL-SPCAPP 25  
 SbWRKY12 -----MHACMEGS-SQLETCLPASSLYAL----SP 26  
 PtWRKY12 -----MHQISQQEHEEEAQEGQRS---FLIRMDHGERDV----PNYELHVS----- 39  
 AtWRKY12 -----MEGGRRVF----SNYDLQQV----- 17

.

OsWRKY12 PPSFLAPLPSLQHKLPQLPQLVHDHAAATGTNHGVMFSSDH-----GCLYPL-LPGIPFC 84  
 BdWRKY12 LPPFLAPLPNQHKLLQMPLVVQ----EQSVNHGVMFSSDH---GGGGLYPL-LPGIPFC 78





SbNST1	-----MSISVNGQSCVPPGFRFHPTEEELLNYYLRKKVASQEIDLVDVIRDVLDLNKLEPWD	55
ZmNST1	-----MSISVNGQSCVPPGFRFHPTEEELLNYYLRKKVASQEIDLVDVIRDVLDLNKLEPWD	55
	***** :***** :*****	
PtNST1	IQERCKIGTTPQNDWYFFSHKDKKYPTGTRTNRATAAGFWKATGRDKVIYSTGKRIGMRK	119
AtNST1	IQEMCKIGTTPQNDWYFFSHKDKKYPTGTRTNRATAAGFWKATGRDKIYISNGRRIGMRK	120
BdNST1	IQERCKIGSGPQNDWYFFSHKDKKYPTGTRTNRATAAGFWKATGRDKAIYNAVSRIGMRK	115
OsNST1	IQERCKIGSGPQNDWYFFSHKDKKYPTGTRTNRATAAGFWKATGRDKAIYNAVHRIGMRK	115
PvNST1	IQEKCKIGSGPQNDWYFFSHKDKKYPTGTRTNRATAAGFWKATGRDKAIYNAVKRIGMRK	115
<b>ShNST1</b>	IQEKCKIGSGPQNDWYFFSHKDKKYPTGTRTNRATAAGFWKATGRDKAIYNAVKRIGMRK	115
SbNST1	IQEKCKIGSGPQNDWYFFSHKDKKYPTGTRTNRATAAGFWKATGRDKAIYNAVKRIGMRK	115
ZmNST1	IQEKCKIGSGPQNDWYFFSHKDKKYPTGTRTNRATAAGFWKATGRDKAIYNAVKRIGMRK	115
	*** ***: ***** **	
PtNST1	TLVIFYKGRAPHGQKSDWIMHEYRLDDSTSDTNVSN-----VMEEEAQ	162
AtNST1	TLVIFYKGRAPHGQKSDWIMHEYRLDDNIISPEDVT-----VHEVVSIGEASQ	168
BdNST1	TLVIFYKGRAPHGLKSDWIMHEYRLDDSSAATAAMVT-----ASSVAASEAAGQQGP	170
OsNST1	TLVIFYKGRAPHGQKSDWIMHEYRLDDPATDTAAATPTVT----SAAAAAAMAAAADGGQ	171
PvNST1	TLVIFYKGRAPHGQKSDWIMHEYRLDDPGAVVSGDD-----AATVAAAAAASDAGQ	167
<b>ShNST1</b>	TLVIFYKGRAPHXQKSDWIMHEYRLDDPAAAGFRDAAGAAGRTVAAACCVCG--CVXGPRA	173
SbNST1	TLVIFYKGRAPHGQKSDWIMHEYRLDDPAAAGDAATAA---AATVAAAAAASDGGQE	172
ZmNST1	TLVIFYKGRAPHGQKSDWIMHEYRLDDPAAAAAAGSGDAVANDDAATVSK-----A	166
	***** ***** *	
PtNST1	EEGWVCRIFKKNLNLKTLDKPFSSSPIAD-----TRNQML	199
AtNST1	DEGWVCRIFKKNLHKTNLNSPVGASLSGGGDT-----PKTTSS	208
BdNST1	EDGWVCRVFKKKHHKDTNSGSGSGG-----NKKAAA-----LRRSSSSPLY	214
OsNST1	EDGWVCRVFKKKHHKKEAGGGGKHGGDGS-----G-----AKAAHA	210
PvNST1	EDGWVCRVFKKKHHKESGSGGGGRHSGRV---GDGGK-----SAAAAHRGLQ	214
<b>ShNST1</b>	RGPLGGVQVVQKEAPPQGVFGPEAGPPNTGSNTKHGGGAKAAAAAAAHQHGHGGLQ	233
SbNST1	D-GWVCRVFKKKHHKESGSGGGGSKHGGSNNEHGGGKAAAAAA--AAQHGHGGLQ	229
ZmNST1	T-TLIAVNLFSAAPPVHRHH-----	185
	...	
PtNST1	SSCDEGTIDQTFHYMGRT---CKEENV-----ADNSATARYLRPVDTAINYV	243
AtNST1	QIFNEDTLDQFLELMGRS---CKEELN-----	232
BdNST1	SSGDAAALDQILHYMGRSSAACKQEHDSPRPAQAQ---QAQARPTSRYLRPIETALA--	269
OsNST1	YSSDDALDQILQYMGRS---CKQEHELPSQASGGGAGAGSRPASRYLRPIDTVL---	264
PvNST1	YSSDDTLDQILQYMGRS---FKQEHEQVLSPO-----GGGAGPRYLRPIDTVL---	260
<b>ShNST1</b>	YSSDDTLDQILQYMGRS---CKQEHELLSPPPAAS--AAGPGRAPSRYLRPIETVL---	285
SbNST1	YSSDDALDQILQYMGRS---CKQEHELLSPPP-----PGRAASRYLRPIETVL---	275
ZmNST1	-----	185

PtNST1	HHDGFMKLPSPLESPNSISSQNCYQPMITDNEGS-----ITNQMSYPLDPGLDNWAT	294
AtNST1	-LDPFMKLPNLESPNSQAINNCHVSSPDTN-----HNIHVSNNVVDTSFVTSWAA	280
BdNST1	GGHGFMKLPPLLESPSSAAAAAPPNT-----TPVPETTMDWAM	306
OsNST1	GGHGFMKLPPLLESPSAAATLSSTPSTGGDAASS--AAAAAADHLLH----HHHRTDWAM	318
PvNST1	GGHGFMKLPPLLESPSAAALTPHLS----AAGDAEAA-SAGDD--LLHGAGTANGITDWAM	313
<b>ShNST1</b>	GGHRFMKLPPLLESPSAAAAALTPPHAVSGDAAAAEEVVDVGLHRAG-GIGITDWAM	344
SbNST1	GGHGFMKLPPLLESPSAAAAAM----TPQAVSGDA----GVVDDLLGLHRGGIGNGITDWAM	327
ZmNST1	-----	185

PtNST1	LDRLFAYQLNGQTETSRQLPC---IDPTITYCTPSTDHLHDLRLPTLRSSFPLPSN---R	348
AtNST1	LDRLVASQLNGPYSITAVNESHVGH-DHLALPSVRSFY-PSLNRS-ASYHAGL-----	332
BdNST1	MDRLVASHLNGQLHDDHASTAVVDDHRLCSAFDDGAGED-NDDGGL-AFYSAATRLLG	364
OsNST1	MDRLVASHLNGANSDA-----PDDQLCFDA-----ADDDGL-AYYSAAATRLLG	361
PvNST1	MDRLVASHLNGQAADPAAPPA-----SHHQLCFDDGPGADD-DADGL-AFYAAATRLHA	366
<b>ShNST1</b>	MDRLVASHLNGQ-APDVAPGA-----DHLSSCFDDATGADDADAAGL-AFYSAANRL-L	396
SbNST1	MDRLVASHLNGQEAPDVAPAA-----DQLGSCFDDATGADDADAAGL-AFYSAANRLLV	381
ZmNST1	-----	185

PtNST1	SYHGTQDYNNEIDLWNFTTRSSPDT-----LCQLSNTGA	382
AtNST1	----TQEYTPEMELWNTTSSLSSSPGP-----FCHVSNNGSG	365
BdNST1	AA----AGAVDDDLWSFARSAERLSHHGSQ-----	390
OsNST1	GA----NAGTDDDLWSFARSAAPPPPPSSATPERLSHVAL---	400
PvNST1	G----GGSSDDDLWSFTRSSARP--G---AATSTERLSHVSL---	399
<b>ShNST1</b>	GSAGSSGAGSDDDLWSFTRSSASAAAA---AATPTERLSHVSL---	436
SbNST1	GSAGSSGAGSDDDLWSFTRSSAAA--A---AATSTERLSHVSL---	419
ZmNST1	-----	185

## ShKNAT7

BdKNAT7	MLQ----GDHQLIEGGTMEGMSF--GGGGGECSSSSAAAAA-----AAVTAEEAEAE	47
OsKNAT7	MQG-----GDHGGMEMGVGSFTGGGGGECSSSSATAAAAA---AAAAAAAAAEAE	49
PvKNAT7	---MQ--GGGDQGGSLGMDVGF-----AGGAQCSSSSAA-----AAAAAEAE	39
ZmKNAT7	---MQRGGDQAGASLGEMGVGY-AGGGGGGECSSSSAAA-----AAAAAEGEE	46
<b>ShKNAT7</b>	--MQGGGGDQAGGSLMDMGVGY-AGG--GGECSSSTSAAAAAA-AAAAAAAAAEAE	54
SbKNAT7	MQGGGGGGDQAGGSLGMDMGVGY-AAGSGGAECSSSTSAAAAAAAAAAAAAAAAAAEAE	59
PtKNAT7	-----MQEPNLGMMGSGSGALGGLSS-----GEMSVLSGDQ	33
AtKNAT7	-----MQEAALGMMGATVGGD-----GDTAVVAEQ	25

\* .: .:.

BdKNAT7	RQLLKEMAVHPLCEQLVAAHVGCRLRVATPIDHLP IIDAQLAQSGLLHSYAAHH---RP	104
OsKNAT7	RQLLKGEI AVHPLCEQLVAAHVGCRLRVATPIDHLP IIDAQLAQSSGLLHSYAAHH---RP	106

PvKNAT7	RQLLKGEI AVHPLCEQLVTAHVGCLRVATPIDHPLIDAQLAQSSGLLHSYAAHH---RP	96
ZmKNAT7	RQLLKGEI AVHPLCEQLVAAHVGCLRVATPIDHPLIDAQLAQSSGLLHSYAAHH---RP	103
<b>ShKNAT7</b>	RQLLKGEI AVHPLCEQLVAAHVGCLRVATPIDHPLIDAQLAQSSGLLHSYAAHH---RP	111
SbKNAT7	RQLLKGEI AVHPLCEQLVAAHVGCLRVATPIDHPLIDAQLAQSSGLLHSYAAHH---RP	116
PtKNAT7	SRQLKAEIATHPLYEQLLSAHVSCLRVATPIDQLPLIDAQLSQSHLLRSYASQHNQHG	93
AtKNAT7	NRQLKGEIATHPMYEQLLAAHVACL RVATPIDQLPI IEAQLSQSHLLRSYASTAVG---	82
	: **.*:*.***: ***:***.*****:*.**:***:* ***:***:	
BdKNAT7	FLSPHDKQDLDSFLAQYLMLLCSFREQLQQHVRVHAVEAVMACREIEQSLQDLTGATLEE	164
OsKNAT7	FLSPHDKQELDSFLAQYMMLLCSFREQLQQHVRVHAVEAVMACREIEQSLQDLTGATLEE	166
PvKNAT7	FLSPHDKHDLDSDFLAQYLMLLCSFREQLQQHVRVHAVEAVMACREIEQSLQDLTGATLEE	156
ZmKNAT7	FLSPHDKHDLDSDFLAQYLMLLCSFREQLQQHVRVHAVEAVMACREIEQSLQDLTGATLEE	163
<b>ShKNAT7</b>	FLSPHDKHDLDSDFLAQYLMLLCSFREQLQQHVRVHAVEAVMACREIEQSLQDLTGATLEE	171
SbKNAT7	FLSPHDKHDLDSDFLAQYLMLLCSFREQLQQHVRVHAVEAVMACREIEQSLQDLTGATLEE	176
PtKNAT7	SLSPHERQDLDNFLAQYLIILCTFKDQLQQHVRVHAVEAVMACREIETTLQALTGVTLGE	153
AtKNAT7	--YHHRHELDNFLAQYVMVLC SFKEQLQQHVRVHAVEAVMACREIENNLHSLTGATLGE	140
	*:::***.*****:::***:***:*****.*****.*** ***:*** *	
BdKNAT7	GTGATMSEDEEPEQTIMEAAAAMDMSNGDHMMGFGLVPTDSERSLMERVRQELKIELK	224
OsKNAT7	GTGATMSEDEDETA PMLE--GPMDMGSDGHDLMGFGPLMPTDSERSLMERVRQELKIELK	224
PvKNAT7	GSGATMSEDEDE- PPMLE--GALDMGSDGHDMMGFGPLLPTDSERSLMERVRQELKIELK	213
ZmKNAT7	GTGATMSEDEDE-APMLE--VGLDMGSDGHDMMGFGPLMPTDSERSLMERVRQELKIELK	220
<b>ShKNAT7</b>	GTGATMSEDEDE- PPMLE--AGLDMGSDGHDMMGFGPLLPTDSERSLMERVRQELKIELK	228
SbKNAT7	GTGATMSEDEDE- PPMLE--GALDMGSDGHDMMGFGPLLPTDSERSLMERVRQELKIELK	233
PtKNAT7	GTGATMSDEDDQLQDFSLDQSS---ADGHDMMGFGPLLPTESERSLMERVRQELKIELK	210
AtKNAT7	GSGATMSEDEDDLPMD FSSDN SVD FSGGHDMTGFGPLLPTESERSLMERVRQELKLELK	200
	*:****:***: :. :.***: *****:***:*****.*****:***	
BdKNAT7	QGFKSRIGDVREEILRKR RAGKLP GD TTTILKQWWQQH SKW PYPTEDDKAKLVEETGLQL	284
OsKNAT7	QGFKSRIEDVREEILRKR RAGKLP GD TTTILKQWWQQH SKW PYPTEDDKAKLVEETGLQL	284
PvKNAT7	QGFKSRIEDVREEILRKR RAGKLP GD TTSILKQWWQQH SKW PYPTVSAL-----	262
ZmKNAT7	QGFKSRIEDVREEILRKR RAGKLP GD TTSILKQWWQEH SKW PYPTEDDKAKLVEETGLQL	280
<b>ShKNAT7</b>	QGFKSRIEDVREEILRKR RAGKLP GD TTSILKQWWQQH SKW PYPTEDDKAKLVEETGLQL	288
SbKNAT7	QGFKSRIEDVREEILRKR RAGKLP GD TTSILKQWWQQH SKW PYPTEDDKAKLVEETGLQL	293
PtKNAT7	QGFKSRIEDVREEILRKR RAGKLP GD TTVLKNWWQQH SKW PYPTEDDKAKLVEETGLQL	270
AtKNAT7	QGFKSRIEDVREEIMRKR RAGKLP GD TTVLKNWWQQH CKW PYPTEDDKAKLVEETGLQL	260
	***** *****.*****.*****:***:***.*.***** .	
BdKNAT7	KQINNWF INQRKR NWHNNSQTST-LKSKRKR	314
OsKNAT7	KQINNWF INQRKR NWHNNSQTST-LKSKRKR	314
PvKNAT7	-----	262
ZmKNAT7	KQINNWF INQRKR NWHNNSQTST-LKSKRKR	310
<b>ShKNAT7</b>	KQINNWF INQRKR NWHNNSQTST-LKSKRKR	318

SbKNAT7	KQINNWFINQRKRNWHNNSQTST-LKSKRKR	323
PtKNAT7	KQINNWFINQRKRNWHNSQSVTSLKSKRKR	301
AtKNAT7	KQINNWFINQRKRNWHNNSHSLTSLKSKRKH	291

## 7. ANEXOS



# INFORMAÇÃO

Informamos que o projeto **CIBio/IB No. 2011/02 – “Controle da biossíntese de lignina em cana-de-açúcar e eucalipto”**, cujo pesquisador responsável é o Prof. Dr. Paulo Mazzafera, sub-projeto, “Identificação e caracterização de fatores de transcrição que regulam a biossíntese de lignina em *Saccharum spp.*” do pós graduando Juan Pablo Portilla Llerena, encontra-se devidamente aprovado e regularizado junto a CIBIO/IB-UNICAMP e a CTNBio, conforme legislação vigente. O professor responsável se compromete a incluir o nome do aluno no próximo relatório da CIBio/IB/Unicamp.

Cidade Universitária “Zeferino Vaz”,  
31 de agosto de 2021

Prof. Dr. José Luiz Proença Módena  
Presidente da CIBio/Instituto de Biologia – Unicamp

---

**CIBio/IB-UNICAMP**

Comissão Interna de Biossegurança  
Instituto de Biologia – Unicamp  
Caixa Postal 6109 – 13083-970 Campinas SP  
Tel.: (19) 3521-6359 – email: [comisib@unicamp.br](mailto:comisib@unicamp.br)

---

Documento assinado eletronicamente por **JOSÉ LUIZ PROENÇA MODENA, PRESIDENTE DA CIBIO/IB/UNICAMP**, em 31/08/2021, às 15:23 horas, conforme Art. 10 § 2º da MP 2.200/2001 e Art. 1º da Resolução GR 54/2017.

---



A autenticidade do documento pode ser conferida no site:  
[sigad.unicamp.br/verifica](http://sigad.unicamp.br/verifica), informando o código verificador:  
**7A43196B 6CE34B53 89AAA693 F9A01EF8**





COORDENADORIA DE PÓS-GRADUAÇÃO  
INSTITUTO DE BIOLOGIA  
Universidade Estadual de Campinas  
Caixa Postal 6109. 13083-970, Campinas, SP, Brasil  
Fone: (19) 3521-6378. email: cpgib@unicamp.br



## DECLARAÇÃO

Em observância ao **§5º do Artigo 1º da Informação CCPG-UNICAMP/001/15**, referente a Bioética e Biossegurança, declaro que o conteúdo de minha tese de Doutorado, intitulada **“Identificação e caracterização de fatores de transcrição que regulam a biossíntese de lignina em *Saccharum spp.*”**, desenvolvida no Programa de Pós Graduação em Biologia Vegetal do Instituto de Biologia da Unicamp, não versa sobre pesquisa envolvendo seres humanos ou animais.

Assinatura: \_\_\_\_\_

Nome do(a) aluno(a): Juan Pablo Portilla Llerena

Assinatura: \_\_\_\_\_

Nome do(a) orientador(a): Paulo Mazzafera

Data: 24/08/2021

## Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **IDENTIFICAÇÃO E CARACTERIZAÇÃO DE FATORES DE TRANSCRIÇÃO QUE REGULAM A BIOSÍNTESE DE LIGNINA EM SACCHARUM SPP.**, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 6 de Novembro de 2021

Assinatura : \_\_\_\_\_

Nome do(a) autor(a): **Juan Pablo Portilla Llerena**  
RG n.º V993310-C-MJ

Assinatura : \_\_\_\_\_

Nome do(a) orientador(a): **Prof. Dr. Paulo Mazzafera**  
RG n.º 8336840-1