UNIVERSIDADE DE SÃO PAULO ESCOLA DE ENGENHARIA DE LORENA

AMANDA FANELLI DE SOUZA

Análise filogenética e caracterização funcional de genes BAHD de cana-deaçúcar potencialmente envolvidos com a incorporação de ácidos hidroxicinâmicos na parede celular

> Lorena 2019

AMANDA FANELLI DE SOUZA

Análise filogenética e caracterização funcional de genes BAHD de canade-açúcar potencialmente envolvidos com a incorporação de ácidos hidroxicinâmicos na parede celular

> Tese apresentada à Escola de Engenharia de Lorena da Universidade de São Paulo para obtenção do título de Doutor em Ciências do Programa de Pós-Graduação em Biotecnologia Industrial na área de concentração de Conversão de Biomassa

Orientador: Dr. Elisson Antonio da Costa Romanel

Versão Original

Lorena 2019 NÃO AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, SERÁ DISPONIBILIZADO AUTOMATICAMENTE APÓS 2 ANOS DA PUBLICAÇÃO

> Ficha catalográfica elaborada pelo Sistema Automatizado da Escola de Engenharia de Lorena, com os dados fornecidos pelo(a) autor(a)

Souza, Amanda Fanelli de Análise filogenética e caracterização funcional de genes BAHD de cana-de-açúcar potencialmente envolvidos com a incorporação de ácidos hidroxicinâmicos na parede celular / Amanda Fanelli de Souza; orientador Elisson Antonio da Costa Romanel - Versão Original. - Lorena, 2019. 146 p.

Tese (Doutorado em Ciências - Programa de Pós Graduação em Biotecnologia Industrial na Área de Conversão de Biomassa) - Escola de Engenharia de Lorena da Universidade de São Paulo. 2019

1. Cana-de-açúcar. 2. Ácidos hidroxicinâmicos. 3. Bahd. I. Título. II. Romanel, Elisson Antonio da Costa, orient.

Aos meus pais, que pelo exemplo, me ensinaram a sempre trabalhar pelos meus sonhos e sempre farão parte de todas as minhas conquistas

AGRADECIMENTOS

Gostaria de agradecer especialmente à minha mãe Cleonice, que sempre se esforçou e se dedicou muito para que eu pudesse estudar, a quem devo toda a minha formação. Por todo apoio me ajudando a superar os obstáculos, me acompanhando ao laboratório, me levando para os lugares sempre e me incentivando a seguir em frente em busca dos meus objetivos. Ao meu irmão, Gabriel, que me ensina desde criança, também me apoiou muito e sempre me incentiva, inclusive dando contribuições a este texto.

Ao meu orientador, prof. Dr. Elisson Romanel, pelas oportunidades desde quando eu ainda estava na graduação. Por todos os ensinamentos, incentivo, e dedicação a este projeto, desde quando o iniciamos, contribuindo muito para minha carreira e formação.

Ao meu co-orientador no exterior, PhD Ronald D. Hatfield, por ter me recebido tão bem em seu país e compartilhado sua sabedoria e conhecimento de uma carreira inspiradora. Obrigada por ter me aceitado no DFRC e por trabalhar conosco próximo de sua aposentadoria.

À prof. Dra Tatiane da Franca, pelos ensinamentos e ajuda desde quando comecei a trabalhar no laboratório, contribuindo muito para este trabalho e minha formação. Ao prof. Dr. André Ferraz, por também compartilhar seu conhecimento e pelas contribuições.

Aos colaboradores PhD David Rancour e PhD Michael D. Sullivan, pelas contribuições e apoio nas questões e desafios desse projeto. Aos pesquisadores Steven D. Karlen e John Ralph pelas análises DFRC e acidólise branda.

A todos do grupo do Laboratório de Genômica de Plantas e Bioenergia, desde que ele foi fundado, em especial Denise, Daniele, Thaís, Paula, Luís e Fernanda, pela amizade e contribuições ao projeto. Aos técnicos José Carlos Tavares e José Carlos Moreira pela amizade e assistência.

A todos do USDA Dairy Forage Research Center (Madison, WI), pelo acolhimento e apoio (Geoffrey Brink, Cindi Birch, Lila Waters, Jim Richmond, Kenneth Kalscheur, Diane Amundson, Jan Pitas, Laurie Reinhardt, Wayne Zeller, George Aldrich, Kristan Reed). Agradeço principalmente as técnicas Malory Schultz e Miranda Sikora, pela amizade e ajuda quando eu estava longe de casa, e apoio técnico também no laboratório. A Lisa Koch pela assistência técnica e apoio.

Ao CNPq pela bolsa de doutorado (142474/20150) e a CAPES pela bolsa do período sanduíche no exterior (PDSE-CAPES 88881.135309/2016-01). À FAPESP ((2014/17486-6, 2014/06923-6, 2016/24391-7, and 2016/25785-9) e o CNPq (444912/2014-2) pelo financiamento do projeto de pesquisa.

À Escola de Engenharia de Lorena (EEL-USP) pela minha formação ao longo dos últimos oito anos, e ao USDA Dairy Forage Research Center pela colaboração, em um ano que contribuiu muito para este trabalho e meu crescimento pessoal e profissional.

RESUMO

SOUZA. A.F. Análise filogenética e caracterização funcional de genes BAHD de canade-açúcar potencialmente envolvidos com a incorporação de ácidos hidroxicinâmicos na parede celular. 2019. 146 p. Tese (Doutorado em Ciências) – Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, 2019.

O uso eficiente da biomassa lignocelulósica de cana-de-açúcar em biorefinarias é um desafio dada a baixa digestibilidade desses materiais. Em gramíneas, o conteúdo dos ácidos hidroxicinâmicos é um dos fatores que contribui para a recalcitrância da parede celular. Genes da superfamília BAHD envolvidos na adição dos ácidos hidroxicinâmicos ferúlico (FA) e p-cumárico (pCA) (subclado de Mitchell) foram caracterizados funcionalmente em gramíneas, e têm potencial de aplicação no aumento da digestibilidade da parede celular. Dessa forma, os objetivos desse trabalho foram a identificação dos genes da família BAHD em cana-de-açúcar, identificação de genes candidatos com potencial função relacionada à incorporação de ácidos hidroxicinâmicos e caracterização funcional dos genes selecionados com potencial biotecnológico ScAt10 e ScAt4. Utilizando métodos de bioinformática foi feita uma análise filogenética da família BAHD em gramíneas. Usando bancos de ESTs e RNA-Seq de cana-de-açúcar, foram identificados transcritos pertencentes à família BAHD e a análise in silico do padrão de expressão dos genes do clado de Mitchell deu suporte ao provável papel de alguns genes. A análise de expressão dos genes do clado de Mitchell ao longo do desenvolvimento do colmo de híbridos de cana-de-açúcar (H89 e H321) por RT-qPCR, revelou que de uma forma geral esses genes foram mais expressos no entrenó mais jovem. Os genes ScAt10 e ScAt4 foram inseridos em vetores de super-expressão (pzp221b:Ox) e silenciamento por RNAi (pzp221b:RNAi), os quais foram transformados na gramínea modelo milho (Zea mays). As linhagens de superexpressão Ox:ScAt10 apresentaram um aumento muito significativo (~170 vezes) no teor de pCA, especificamente na hemicelulose, e redução no teor de FA (~10x). Ademais, pelo menos um evento de silenciamento ScAt4: RNAi apresentou diminuição de ~3x de pCA especificamente na lignina, porém nas linhagens de super-expressão Ox:ScAt4 não foram detectadas alterações no conteúdo de pCA/FA da parede celular. Os resultados sugerem que ScAt10 possui atividade relacionada à incorporação de pCA na hemicelulose, que pode ser explorada para aumento da digestibilidade.

Palavras chave: Cana-de-açúcar. Ácidos Hidroxicinâmicos. BAHD

ABSTRACT

SOUZA. A.F. Phylogenetic analysis and functional characterization of sugarcane BAHD genes potentially involved with hydroxycinnamic acid incorporation into cell wall. 2019. 146 p. Thesis (Doctoral of Science) – Escola de Engenharia de Lorena, Universidade de São Paulo, Lorena, 2019.

The efficient use of sugarcane lignocellulosic biomass is challenging, especially due to the low digestibility of these materials. In grasses, hydroxycinnamic acids content is one of several factors that contributes to cell recalcitrance. BAHD family genes involved in ferulic acid (FA) and p-coumaric acid (pCA) incorporation (Mitchell's subclade) have already been identified and functionally characterized in many grasses and have potential to increase cell wall digestibility. In this sense, the goals of this work were the identification of BAHD genes in sugarcane, identification of candidate genes putatively involved with hydroxycinnamic acid incorporation and functional characterization of selected genes ScAt10 and ScAt4, with potential in biotechnological applications. Using bioinformatics approaches, we performed a phylogenetic analysis of BAHD in grasses. Using ESTs and RNA-seq data, sugarcane transcripts belonging to BAHD family were identified, and also in silico expression analysis of Mitchell's clade genes supported their predicted role in many cases. Expression analysis of Mitchell's clade genes using RTqPCR throughout culm development of sugarcane hybrids (H89 and H321) showed a higher expression in the youngest internode. ScAt10 and ScAt4 were cloned into (pzp221b:Ox) and RNAi silencing (pzp221b:RNAi), which were overexpression transformed into model grass maize (Zea mays). Transgenic overexpression lines (Ox:ScAt10) showed a very significant increase (~170x, p<0.001) in pCA content, specifically bound to hemicellulose. In addition, at least one silencing event ScAt4:RNAi showed $\sim 3x$ decrease in pCA content, specifically attached to lignin. However, lines overexpressing ScAt4 did not show significant differences in pCA content. This work data suggests that ScAt10 has activity related with incorporation of pCA into hemicellulose, which could be explored for biotechnological applications to increase cell wall digestibility.

Keywords: Sugarcane. Hydroxycinnamic Acid. BAHD

LISTA DE ILUSTRAÇÕES

Figura 1 – Corte transversal de cana-de-açúcar, visualizado através de microscópio óptico, demonstrando feixes vasculares rodeados por células de parênguima......27 Figura 2 - Estrutura química simplificada do ácido ferúlico (FA) ligado à Figura 3 – Esquema simplificado das principais unidades que constituem a lignina.......30 Figura 4 – Filogenia de proteínas BAHD aciltransferases bioquimicamente caracterizadas e putativas identificadas nos genomas de Populus (Álamo), Arabidopsis, Medicago (Luzerna Figura 5 - Filogenia do clado de Mitchell inferida pelo método bayesiano utilizando sequências identificadas em arroz e enzimas BAHD aciltransferases caracterizadas Figura 6 – Linhagens transgênicas de milho em diferentes estágios de crescimento.......63 Figura 7- Alinhamento de regiões das sequências codificantes clonadas de cana-de-açúcar (ScAt10 e ScAt4) com as sequências genômicas dos ortólogos em milho (ZmAt10 e ZmAt4), mostrando a posição e especificidade dos iniciadores para genotipagem das linhagens de Figura 8- Região de anelamento dos iniciadores para genotipagem das linhagens de Figura 9– Posição dos iniciadores para análise dos níveis de expressão de ScAt10 e ScAt4 Figura 10- Cromatograma das soluções padrão contendo padrão interno ácido 2hidroxicinâmico e os padrões ácido p-cumárico e ácido ferúlico na concentração de 1mg/ml......73 Figura 11 - Mecanismo das etapas do processo DFRC (Derivatization Followed by Figura 12 – Distribuição entre os clados filogenéticos das sequências BAHD identificadas no genoma das espécies e nos bancos genômicos, de sequências expressas e transcritos de Figura 13 - Análise filogenética do clado A de Mitchell das BAHD aciltransferases84 Figura 14 – Análise in silico do padrão de expressão dos genes ScAtl a ScAtl0 em Figura 15 – Análise do padrão de expressão por RT-qPCR dos genes ScAt1 a ScAt10 em Figura 14 – Análise do padrão de expressão por RT-qPCR dos genes ScAt1 a ScAt10 em diferentes tecidos de híbridos de cana-de-açúcar H89 e H321......91 Figura 17 – Análise da tendência de expressão por RT-qPCR dos genes ScAt1 a ScAt10 s ao longo do desenvolvimento do colmo dos híbridos de cana-de-açúcar H89 e H32192 Figura 18 – Confirmação dos vetores de entrada *pDONR:RNAi:ScAt10* e pDONR:RNAi:ScAt4 por digestão com a enzima de restrição PvuI......95 Figura 19 – Confirmação dos vetores de entrada pDONR:Ox:Sc:At10 por digestão com enzima de restrição KpnI......95 Figura 20 – Confirmação dos vetores de entrada pDONR:Ox:ScAt4 por dupla digestão com as enzimas de restrição *KpnI* e *PvuII*......96 Figura 21 - Confirmação dos vetores de destino pzp221b:RNAi:ScAt10 e Figura 22 - Confirmação dos vetores de destino pzp221b:Ox:ScAt10 por digestão com a enzima de restrição KpnI......97

Figura 23 – Confirmação dos vetores de destino pzp221b:Ox:ScAt4 por digestão com a Figura 24 - Confirmação dos vetores de destino pzp221b:RNAi:ScAt10 e Figura 25 - Confirmação dos vetores de destino pzp221b:Ox:ScAt10 por digestão com a Figura 26 - Confirmação dos vetores de destino pzp221b:Ox:ScAt4 por digestão com a enzima de restrição BamHI.....100 Figura 27 - Genotipagem das linhagens A871 (Ox:ScAt10), A872 (Ox:ScAt4) A873 Figura 28 - Análise dos dados de expressão de milho B73 provenientes do atlas de RNAseq......104 Figura 29 - Expressão Relativa de ZmAt10 (RT-qPCR) nas amostras da nona folha de dois indivíduos de uma linhagem de milho selvagem104 Figura 30 – Expressão Relativa de ZmAt10 (RT-qPCR) nas linhagens de milho transgênicas A873 (ScAt4:RNAi) e nos controles linhagens escape e selvagem105 Figura 31 – Expressão Relativa de ScAt10 (RT-qPCR) nas linhagens de milho transgênicas A871 (Ox:ScAt10) e nos controles linhagens escape e selvagem106 Figura 32- Expressão Relativa de ScAt4 (RT-qPCR) em folha nas linhagens de milho transgênicas A885 (Ox:ScAt4) e nos controles linhagens escape e selvagem107 Figura 33- Expressão Relativa de ZmAt4 (RT-qPCR) nas linhagens de milho transgênicas A885 (Ox:ScAt4) e nos controles linhagens escape e selvagem108 Figura 34 – Análise de expressão por RT-qPCR dos genes ScAt4 e ZmAt4 nos eventos Figura 35- Quantidade de ácido p-cumárico liberada pelo tratamento brando no colmo das linhagens A872 (At10:RNAi) e controle linhagens escape e selvagem B73.....110 Figura 36- Quantidade de ácido ferúlico liberada pelo tratamento brando no colmo das linhagens A872 (At10:RNAi) e controle linhagens escape e selvagem B73.....111 Figura 37- Quantidade de ácido p-cumárico liberada pelo tratamento brando no colmo das linhagens A873 (At4:RNAi) e controle linhagens escape e selvagem B73.....112 Figura 38- Quantidade de ácido ferúlico liberada pelo tratamento brando no colmo das linhagens A873 (At4:RNAi) e controle linhagens escape e selvagem B73.....113 Figura 39- Quantidade de ácido p-cumárico liberada pelo tratamento brando no colmo das linhagens A871 (Ox:ScAt10) e controle linhagens escape e selvagem B73.....114 Figura 40- Quantidade de ácido ferúlico liberada pelo tratamento brando no colmo das linhagens A871 (Ox:ScAt10) e controle linhagens escape e selvagem B73.....115 Figura 41– Razão entre as quantidades pCA/FA liberadas pelo tratamento brando no colmo das linhagens A871 (Ox:ScAt10) e controle linhagens escape e selvagem B73116 Figura 42- Quantidade de ácido p-cumárico liberada pelo tratamento brando no colmo das linhagens A885 (Ox:ScAt4) e controle linhagens escape e selvagem B73.....117 Figura 43- Quantidade de ácido ferúlico liberada pelo tratamento brando para as linhagens A885 (Ox:ScAt4) e controle linhagens escape e selvagem (B73)117 Figura 44- Quantidade de ácido p-cumárico especificamente unido à hemicelulose, liberado pela hidrólise ácida branda no colmo das linhagens A872 (At10:RNAi) e controle Figura 45– Quantidade de ácido ferúlico especificamente unida à hemicelulose, liberado pela hidrólise ácida branda no colmo das linhagens A872 (At10:RNAi) e controle linhagens escape e selvagem B73.....121

Figura 46- Quantidade de ácido p-cumárico especificamente unido à hemicelulose,
liberado pela hidrólise ácida branda, no colmo das linhagens A873 (At14:RNAi) e controle
linhagens escape e selvagem B73122
Figura 47- Quantidade de ácido ferúlico especificamente unido à hemicelulose, liberado
pela hidrólise ácida branda, no colmo das linhagens A873 (At14:RNAi) e controle
linhagens escape e selvagem B73123
Figura 48- Quantidade de ácido p-cumárico especificamente unido à hemicelulose,
liberado pela hidrólise ácida branda, no colmo das linhagens A871 (Ox:ScAt10) e controle
linhagens escape e selvagem B73124
Figura 49- Quantidade de ácido ferúlico especificamente unido à hemicelulose, liberado
pela hidrólise ácida branda, no colmo das linhagens A871 (Ox:ScAt10) e controle
linhagens escape e selvagem B73125
Figura 50– Quantidade de ácido p-cumárico conjugado a unidades S (S-pCA), proveniente
da análise DFRC, no colmo das linhagens A872 (At10:RNAi) e controle linhagens escape e
selvagem (B73)
Figura 51- Quantidade de ácido p-cumárico conjugado a unidades S (S-pCA), proveniente
da análise DFRC, no colmo das linhagens A873 (At4:RNAi) e controle linhagens escape e
selvagem (B73)
Figura 52- Quantidade de ácido p-cumárico conjugado a unidades S (S-pCA), proveniente
da análise DFRC, no colmo das linhagens A871 (Ox:ScAt10) e controle linhagens escape e
selvagem B73
Figura 53- Quantidade de Glicose, Arabinose e Xilose na parede celular do colmo das
linhagens A871 (Ox:ScAt10) e controle linhagens escape e selvagem B73130

LISTA DE TABELAS

Tabela 1 – Genes identificados na família BAHD relacionados à incorporação de ácidos
hidroxicinâmicos
Tabela 2- Padronização dos iniciadores dos unigenes de cana-de-açúcar da família BAHD,
grupo A do clado de Mitchell, para uso em RT-qPCR
Tabela 3– Sequência de iniciadores para clonagem de Ox:ScAt10, Ox:ScAt4, At10:RNAi,
At4:RNAi e temperatura de anelamento predita pelo software primer3plus
Tabela 4– Condições de amplificação dos produtos para clonagem por PCR, comparando
as temperaturas de anelamento utilizadas com a temperatura predita pelos softwares primer
<i>3 plus e Neb calculator</i>
Tabela 5– Condições de amplificação dos produtos attB PCR, comparando as temperaturas
de anelamento utilizadas com a temperatura predita pelos softwares primer 3 plus e Neb
calculator
Tabela 6 – Iniciadores específicos para genotipagem por PCR das linhagens A871
(Ox:ScAt10), A872 (ScAt10:RNAi), A873 (ScAt4:RNAi) e A885(Ox:ScAt4)65
Tabela 7 – Iniciadores específicos para análise do nível de expressão por RT-qPCR das
linhagens A871 (Ox:ScAt10), A872 (ScAt10:RNAi), A873 (ScAt4:RNAi) e
A885(<i>Ox:ScAt4</i>)
Tabela 8 – Padronização dos iniciadores para análise de expressão dos genes ScAt10,
ScAt4, ZmAt10 e ZmAt4 nas linhagens transgênicas de milho por RT-qPCR70
Tabela 9- Quantidade total de proteínas BAHD aciltransferases identificadas nas diferentes
espécies vegetais e comparação com os dados da literatura79
Tabela 10- Número de sequências identificadas em cada banco de dados de cana de
açúcar: Genoma (DNAg), SUCEST, RNA-seq a partir de plântula, entrenó, folha e raiz79
Tabela 11- Análise estatística dos dados de expressão dos genes ScAt1 a ScAt10 por RT-
qPCR em diferentes tecidos de colmo dos híbridos H321 e H8991
Tabela 12– Quantidades de unidades S, G, H, S-pCA e G-pCA detectadas pelo método
DFRC para as linhagens A871 e controles escape e selvagem B73131

LISTA DE ABREVIATURAS E SIGLAS

- BAC Bacterial Artificial Chromosome
- WGP Whole Genome Profiling
- GAX Glucuronoarabinoxilana
- FA Ácido Ferúlico
- pCA Ácido p-Cumárico
- Ara Arabinose
- G Guaiacil
- S Siringil
- H *p*-Hidroxifenil
- MWL Miled Wood Lignin
- RT-qPCR Real Time Reverse Transcription Quantitative Polymerase Chain Reaction
- UBI Promotor de Ubiquitina de milho
- RNAi RNA de interferência
- MUSCLE Multiple Sequence Comparison by Log Expectation
- BLAST Basic Local Alignment Search Tool
- HMMER Hidden Markov Model
- EST Expressed Sequence Tag
- RNA-seq Sequenciamento de RNA
- cDNA DNA complementar
- H89 Híbrido de cana-de-açúcar número 89
- H321- Híbrido de cana-de-açúcar número 321
- Ct Treshold Cycle
- siRNA small interference RNA
- PCR Polymerase Chain Reaction
- Tm Temperatura de melting do DNA
- LB Luria Bertani
- SOB Super Optimal Broth
- GLC/FID Gas Liquid Chromatography with Flame Ionization Detector
- HPLC High Pressure Liquid Chromatography
- MeAra-*p*CA Metil-Arabinose-*p*-Cumarato
- MeAra-FA Metil-Arabinose-Ferulato

DFRC - Derivatization Followed by Reductive Cleavage GC/MS - Gas Chromatography coupled with Mass Spectrometry ANOVA - One Way Analysis of Variance DNAg - DNA genômico Ox - Super-expressão

CDS - Sequência de DNA codificante

SUMÁRIO

1 INTRODUÇÃO23
1.1 CANA-DE-AÇÚCAR: ESPÉCIE ALVO PARA CONVERSÃO DE BIOMASSA 23
1.2 EVOLUÇÃO E GENÔMICA DA CANA-DE-AÇÚCAR
1.3 ANATOMIA DOS TIPOS CELULARES DA BIOMASSA DE CANA-DE-AÇÚCAR
1.4 PAREDE CELULAR EM GRAMÍNEAS: ESTRUTURA E INCORPORAÇÃO DOS ÁCIDOS HIDROXICINÂMICOS27
1.5 RECALCITRÂNCIA DA PAREDE CELULAR E SUA RELAÇÃO COM OS ÁCIDOS HIDROXICINÂMICOS
1.6 COMPOSIÇÃO DA PAREDE CELULAR E RECALCITRÂNCIA EM CANA-DE- AÇÚCAR
1.7 FAMÍLIA GÊNICA BAHD ACILTRANSFERASES
1.8 ENZIMAS ENVOLVIDAS NA INCORPORAÇÃO DOS ÁCIDOS HIDROXICINÂMICOS NA PAREDE CELULAR
2 JUSTIFICATIVAS
3 OBJETIVOS
3.1 OBJETIVOS GERAIS
3.2 OBJETIVOS ESPECÍFICOS
4 MÉTODOS
4.1 IDENTIFICAÇÃO DOS GENES DA FAMÍLIA BAHD ACILTRANSFERASES EM CANA-DE-AÇÚCAR E GRAMÍNEAS RELACIONADAS48
4.2 ANÁLISE <i>IN SILICO</i> DO PADRÃO DE EXPRESSÃO DAS SEQUÊNCIAS IDENTIFICADAS EM CANA-DE-AÇÚCAR50
4.3 COLETA DE AMOSTRAS DOS HÍBRIDOS DE CANA-DE-AÇÚCAR COM CONTEÚDO CONTRASTANTE DOS COMPONENTES DA PAREDE CELULAR 50
4.4 EXTRAÇÃO DE RNA DOS TECIDOS DE CANA-DE-AÇÚCAR, TRATAMENTO COM DNASE E SÍNTESE DE cDNA51
4.5 ANÁLISE DO PADRÃO DE EXPRESSÃO DOS GENES SELECIONADOS EM DIVERSOS ÓRGÃOS E TECIDOS EM CANA-DE-AÇÚCAR POR RT-qPCR
4.5.1 Desenho de iniciadores específicos para os transcritos de cana-de-açúcar identificados na família BAHD52
4.5.2 Reação RT-qPCR e análise da curva de <i>melting</i> 53
4.6 CLONAGEM DE GENES DE CANA-DE-AÇÚCAR SELECIONADOS COM POTENCIAL BIOTECNOLÓGICO54
4.6.1 Desenho de iniciadores para amplificação do cDNA54
4.6.2 Síntese do cDNA e amplificação por reação de PCR para as construções de silenciamento
4.6.3 Síntese das sequências codificantes para as construções de super-expressão 58

4.6.4 Preparo de células competentes de Escherichia coli	.59
4.6.5 Reação de recombinação Gateway e inserção em Escherichia coli	.59
4.7 TRANSFORMAÇÃO DAS LINHAGENS DE MILHO, CULTIVO DAS PLANTAS TRANSGÊNICAS E GENOTIPAGEM POR PCR	5 .62
4.7.1 Transformação das linhagens de milho pela Plant Transformation Facility, Io State University	wa .62
4.7.2 Cultivo das plântulas T ₀ , cruzamento e geração de sementes T ₁	.62
4.8 ANÁLISE DO PADRÃO DE EXPRESSÃO DOS GENES <i>AT10</i> E <i>AT4</i> NAS LINHAGENS DE MILHO TRANSGÊNICAS POR RT-qPCR	.67
4.8.1 Extração de RNA dos tecidos foliares, tratamento com DNAse e síntese cDNA	de .67
4.8.2 Desenho de iniciadores específicos para as sequências ScAt10, ScAt4, ZmAt1 ZmAt4	0 e .68
4.8.3 Reação RT-qPCR e análise da curva de melting	.70
4.9 ANÁLISE FENOTÍPICA DA COMPOSIÇÃO QUÍMICA DA PAREDE CELULA DAS LINHAGENS DE MILHO TRANSGÊNICAS E SELVAGEM B73	ג .71
4.9.1 Isolamento da parede celular do colmo das plantas de milho	.71
4.9.2 Determinação do conteúdo total de ácidos hidroxicinâmicos éster da paro celular	ede .71
4.9.3 Determinação do conteúdo de ácidos hidroxicinâmicos unidos especificament hemicelulose GAX	e à .74
4.9.4 Determinação do conteúdo de ácidos hidroxicinâmicos unidos especificament lignina	e à .75
4.9.5 Determinação do conteúdo de açúcares neutros totais da parede celular	.76
4.9.6 Análise estatística	.77
5 RESULTADOS E DISCUSSÃO	.78
5.1 IDENTIFICAÇÃO, ANOTAÇÃO E ANÁLISE FILOGENÉTICA DOS GENES DE BIOSSÍNTESE DOS ÁCIDOS HIDROXICINÂMICOS	2 .78
5.2 ANÁLISE DE EXPRESSÃO <i>IN SILICO</i> DOS GENES DE CANA-DE-AÇÚCAR D GRUPO A DE MITCHELL EM DIFERENTES TECIDOS)O .86
5.3 ANÁLISE DE EXPRESSÃO DOS GENES DE CANA-DE-AÇÚCAR DO GRUPO DE MITCHELL AO LONGO DO DESENVOLVIMENTO DO COLMO DE HÍBRIDO CONTRASTANTES	A 9S .89
5.4 CLONAGEM DOS GENES DE BIOSSÍNTESE DE p CA NOS VETORES DE SUPER-EXPRESSÃO (OX) E SILENCIAMENTO GÊNICO (RNAi)	.94
5.5 TRANSFORMAÇÃO DAS CONSTRUÇÕES DE SUPER-EXPRESSÃO E SILENCIAMENTO EM LINHAGENS DA PLANTA MODELO ZEA MAYS	101
5.6 CONFIRMAÇÃO DA PRESENÇA DO INSERTO POR GENOTIPAGEM DAS LINHAGENS DE MILHO TRANSGÊNICAS	101
5.7 ANÁLISE DA EXPRESSÃO GÊNICA DAS LINHAGENS DE MILHO TRANSGÊNICAS POR RT-qPCR	102

5.8 ANÁLISE FENOTÍPICA DA COMPOSIÇÃO QUÍMICA DA PAREDE CELULAR DAS LINHAGENS DE MILHO TRANSGÊNICAS109
5.8.1 Análise do impacto do silenciamento ou super-expressão no conteúdo de ácidos hidroxicinâmicos das linhagens transgênicas109
5.8.2 Análise do impacto do silenciamento ou super-expressão gênica no conteúdo de ácidos hidroxicinâmicos especificamente unido à hemicelulose (GAX)
5.8.3 Análise do impacto do silenciamento ou super-expressão gênica no conteúdo de ácidos hidroxicinâmicos especificamente unido à lignina
5.8.4 Análise do impacto do silenciamento ou super-expressão gênica na composição da hemicelulose (GAX) de linhagens A873 e A871129
5.8.5 Análise do impacto da super-expressão gênica de ScAt10 na composição da lignina 131
6 CONCLUSÃO
REFERÊNCIAS
ANEXO A – MAPAS DOS VETORES <i>GATEWAY® pDONR221, pDONR:OX:ScAt10, pDONR:OX:ScAt4, pDONR:RNAi:ScAt10, pDONR:RNAi:ScAt4</i> . OS SÍTIOS DE RESTRIÇÃO DAS ENZIMAS USADAS PARA CONFIRMAÇÃO DOS VETORES FORAM INDICADOS NA FIGURA
ANEXO B – MAPAS DOS VETORES DE DESTINO <i>GATEWAY® PZP221B:OX</i> , <i>PZP221B:RNAi</i> , <i>PZP221B:OX:ScAt10</i> , <i>PZP221B:OX:ScAt4</i> , <i>PZP221B:RNAi:ScAt4</i> E <i>PZP221B:RNAi:ScAT10</i> . OS SÍTIOS DE RESTRIÇÃO DAS ENZIMAS USADAS PARA CONFIRMAÇÃO FORAM INDICADOS NA FIGURA
ANEXO C – CURVA DE MELTING DOS GENES BAHD DE CANA-DE-AÇÚCAR EM ANÁLISES DE RT-qPCR NOS HÍBRIDOS H89 E H321 DE CANA- DE - AÇÚCAR
ANEXO D – CURVA DE MELTING DOS GENES ENDÓGENOS DE MILHO E OS TRANSGENES DE CANA-DE-AÇÚCAR NAS ANÁLISES DE RT-qPCR DAS LINHAGENS TRANSGÊNICAS DE MILHO A871, A873 E A885146

1 INTRODUÇÃO

1.1 CANA-DE-AÇÚCAR: ESPÉCIE ALVO PARA CONVERSÃO DE BIOMASSA

A cana-de-açúcar (*Saccharum* spp.) tem sido utilizada há muitas décadas para produção de açúcar e de biocombustível, por meio de processos fermentativos a partir do caldo rico em sacarose (AMORIM et al., 2011). A maior parte do bagaço é utilizada para geração de energia elétrica através da queima, enquanto folha e palha em geral são deixadas no campo durante a colheita (DIAS et al., 2013).

A produção de cana-de-açúcar é de extrema importância para a economia brasileira, e principalmente para o Estado de São Paulo. De acordo com a última pesquisa Produção Agrícola Municipal do IBGE, no ano de 2017 foram produzidos cerca de 760 milhões de toneladas de cana-de-açúcar no país, representando um valor de produção médio de 54 bilhões de reais, contribuindo com 21% do valor total da produção agrícola nacional. São Paulo foi o estado brasileiro com maior produção de cana-de-açúcar, com mais de 450 milhões de toneladas, gerando um valor total de 30 bilhões de reais que corresponde a 75% do valor total da produção agrícola do estado (IBGE, 2017).

No entanto, cerca de 2/3 de toda a biomassa produzida, correspondentes ao bagaço e a palha, poderiam ser utilizados no contexto das denominadas biorefinarias para a produção de combustíveis e bioprodutos. A produtividade de biocombustíveis poderia aumentar consideravelmente se os açúcares da parede celular fossem convertidos em etanol de segunda geração. Além disso, diversos bioprodutos de maior valor agregado poderiam ser obtidos, tais como insumos químicos, metabólitos, enzimas, ou ainda alimentação animal (CHANDEL et al., 2012; DEMIRBAS, 2009; SOCCOL et al., 2010).

Nos últimos anos, essa utilização da biomassa vem ganhando destaque devido à crescente busca pela substituição dos combustíveis fósseis, à preocupação com as mudanças climáticas, bem como à necessidade de melhor distribuição geográfica das reservas energéticas. Ademais, existe ainda a vantagem de não competição com a produção de alimentos, e a possibilidade de aumento da produtividade de etanol sem a necessidade de expansão das terras cultivadas. No caso da cana-de-açúcar, o aproveitamento da palha e do bagaço seria ainda favorecido pela facilidade de integração com os processos já utilizados para a produção do etanol de primeira geração (DE SOUZA et al., 2014; RAGAUSKAS et al., 2006; SIMS et al., 2010).

Entretanto, existem alguns desafios para a viabilidade econômica da conversão do material lignocelulósico. Para a conversão dos açúcares da parede celular em produtos por meio de processos fermentativos, é preciso uma etapa de hidrólise dos polissacarídeos. Entretanto, o acesso das enzimas hidrolíticas à celulose é dificultado pela estrutura complexa e recalcitrante da parede celular, o que exige etapas de pré-tratamento que encarecem o produto final (BEHERA et al., 2014; SINGH; SUHAG; DHAKA, 2015; WYMAN, 2008). Assim, existem apenas duas usinas de etanol de segunda geração em operação no Brasil, com capacidade de produção ainda abaixo do esperado (DOS SANTOS et al., 2016)

Muitos esforços têm sido feitos para superar esses obstáculos, tais como a modificação de enzimas hidrolíticas, aperfeiçoamento do processo industrial, bem como a seleção e engenharia genética de plantas, visando diminuir a recalcitrância da matériaprima (MACRELLI; MOGENSEN; ZACCHI, 2012). Nesse contexto, há um grande interesse no desenvolvimento de novos cultivares de cana-de-açúcar com características desejáveis para a produção de biocombustíveis e biomateriais. Para isso, podem ser utilizadas técnicas como a modificação da expressão gênica, além da identificação de marcadores moleculares (SOUZA et al., 2011).

Nesse sentido, estudos evolutivos e genômicos da cana-de-açúcar fornecem ferramentas que permitem identificação de genes com potenciais aplicações biotecnológicas. Para possibilitar o estudo e anotação desses genes, também é fundamental a compreensão da composição, estrutura e dos componentes relacionados com a recalcitrância da parede celular de cana-de-açúcar.

1.2 EVOLUÇÃO E GENÔMICA DA CANA-DE-AÇÚCAR

A cana-de-açúcar é uma cultura perene pertencente à família das *Poaceas* ou gramíneas, à qual também pertencem espécies como Sorgo (*Sorghum bicolor*), Milho (*Zea mays*), Brachypodium (*Brachypodum distachyon*), arroz (*Oryza sativa*) e Setaria (*Setaria spp.*). Juntamente com sorgo, milho e Setaria, as espécies do gênero *Saccharum* fazem parte de um grupo com metabolismo C4, enquanto Brachypodium e arroz pertencem a um grupo C3 (KELLOGG, 2001). As análises de sequencias genômicas comparativas entre gramíneas revelaram que a cana-de-açúcar divergiu de sorgo há cerca de 7.8 <u>m</u>ilhões de <u>anos atrás</u> (m.a.a.), há 11.5 m.a.a de milho e há 26 m.a.a de setaria, enquanto que

Brachypodium e arroz divergiram de cana-de-açúcar há cerca de 52 m.a.a. (BENNETZEN et al., 2012; KIM; TANG; PATERSON, 2009)

As variedades comerciais modernas de cana-de-açúcar são resultado da hibridização interespecífica entre *S. officinarum* (2n=80) e *S. spontaneum* (2n=138) com menores contribuições *de S. robustum, S. sinense, S. barberi, Erianthus* e *Miscanthus* (PATERSON; MOORE; TEW, 2013). A estrutura genômica poliploide e aneuploide, com cerca de 10 a 12 cópias de genes homólogos (SOUZA et al., 2011), além do tamanho e da natureza polimórfica do genoma de cana-de-açúcar, representaram um desafio para o seu sequenciamento. Assim, apenas recentemente um genoma de referência foi anotado e disponibilizado (GARSMEUR et al., 2018)

No entanto, considerando-se a dificuldade de obter um genoma completo, grandes esforços já vinham sendo feitos para a obtenção de sequências expressas e transcriptômicas, que pudessem contribuir disponibilizando sequências de DNA de canade-açúcar. Nesse sentido, o programa de sequências expressas EST (Expressed sequence tags) do SUCEST (VETTORE et al., 2001) permitiu a obtenção de cerca de 240000 EST provenientes de 26 bibliotecas de cDNA construídas a partir de distintos órgãos e tecidos, em diversos estágios de desenvolvimento. Essas sequências foram montadas em mais de 40 000 contigs, fornecendo sequências codificantes completas ou parciais de cana-deaçúcar (VETTORE, 2003). Ademais, com o avanço da tecnologia do sequenciamento de RNA em profundidade (deep-sequencing), diversas análises de transcriptomas de cana-deaçúcar foram feitas (CARDOSO-SILVA et al., 2014; MATTIELLO et al., 2015; VICENTINI et al., 2015). Uma das grandes vantagens no uso do RNA-Seq é a possibilidade de ampliar a descoberta de novos genes, comparado com as EST (CARDOSO-SILVA et al., 2014), além da possibilidade de correlacionar o padrão de expressão dos genes com o órgão/tecido analisado, período do desenvolvimento e/ou sob estresse biótico ou abiótico (STRICKLER; BOMBARELY; MUELLER, 2012).

Mais recentemente, em 2017, um genoma parcial (*draft genome*) foi disponibilizado para a variedade SP80-3280 (RIAÑO-PACHÓN; MATTIELLO, 2017), utilizando a tecnologia do sequenciamento de *reads* longas. Apenas em 2018 um genoma monoploide de referência foi disponibilizado para o híbrido R570 (GARSMEUR et al., 2018). Para superar as dificuldades do sequenciamento, sequências de BAC (cromossomo artificial de bactéria) representando a parte do genoma monoplóide mais rica em genes, a eucromatina, foram analisadas pelo método WGP (*Whole Genome Profiling*). Como a

região da eucromatina participa mais intensamente das recombinações, considerou-se que era a mais importante para programas de desenvolvimento de cultivares. Outro genoma de referência também foi disponibilizado para a espécie ancestral *S. spontaneum*, utilizando um clone haploide AP 85-441. (ZHANG et al., 2018). A haploidia desse clone facilitou a montagem de 32 pseudo-cromossomos compreendendo 8 grupos homólogos.

Assim, o desenvolvimento das tecnologias de sequenciamento e o uso de diferentes abordagens permitiram muito recentemente a obtenção de genomas de espécies do gênero *Sacharum*. Entretanto, os estudos transcriptômicos disponibilizados anteriormente continuam sendo importantes, podendo trazer informações adicionais, além de padrões de expressão gênica. Por exemplo, existem transcriptomas analisados provenientes de cultivares distintas, que podem diferir de R570. Também é possível que algumas sequências codificantes estejam presentes em regiões não incluídas na análise de Garsmeur et al., 2018.

1.3 ANATOMIA DOS TIPOS CELULARES DA BIOMASSA DE CANA-DE-AÇÚCAR

A cana-de-açúcar, assim como as demais monocotiledôneas, apresenta uma anatomia celular complexa. Os entrenós apresentam feixes vasculares rodeados por células de parênquima. Cada feixe é composto de um floema, células de vasos e fibras (Figura 1).

As células de parênquima armazenam material de reserva, principalmente sacarose. Os vasos tem principal função de transportar água, enquanto as fibras estão relacionadas com o suporte estrutural (FERRAZ et al., 2014). Figura 1 – Corte transversal de cana-de-açúcar, visualizado através de microscópio óptico, demonstrando feixes vasculares rodeados por células de parênquima.



Legenda: Ph –Células de floema, V- Vasos, P- parênquima e Fibras nos feixes vasculares Fonte: FERRAZ et al., 2014.

O entrenó de cana-de-açúcar pode ser dividido em regiões, da mais externa para a mais interna, denominadas córtex e medula, respectivamente. O córtex contém pequenas quantidades de células de parênquima e maior parte dos feixes vasculares, com grande quantidade de fibras. Já na medula predominam células de parênquima, embora apresentem pequenas quantidades de feixes vasculares (SANJUÁN et al., 2001).

1.4 PAREDE CELULAR EM GRAMÍNEAS: ESTRUTURA E INCORPORAÇÃO DOS ÁCIDOS HIDROXICINÂMICOS

A parede celular das gramíneas (*Poaceas*) possui características únicas quando comparada com as demais angiospermas. Embora a estrutura seja similar, a composição e estrutura dos polissacarídeos não celulósicos, bem como os tipos de ligações cruzadas, o conteúdo de compostos fenólicos e a composição da lignina são distintos (BURTON; FINCHER, 2012; VOGEL, 2008).

A parede celular primária das gramíneas, classificada como de tipo II, constitui-se de celulose envolvida por <u>G</u>lucurono<u>a</u>rabino<u>x</u>ilana (GAX, tipo de hemicelulose) e altos níveis de ácidos hidroxicinâmicos (principalmente ácidos ferúlico e *p*-cumárico). Em eudicotiledôneas e monocotiledôneas não comelinídeas, por sua vez, a parede celular

primária é do tipo I e constitui-se de celulose envolvida por Xiloglucana e altos níveis de proteínas estruturais e pectina (VOGEL, 2008).

Nas gramíneas, a parede celular secundária é depositada no interior da parede primária principalmente nas fibras e nos vasos conforme as células amadurecem (FERRAZ et al., 2014; VOGEL, 2008). A deposição ocorre apenas em pequena quantidade em algumas células de parênquima, que apresentam menor espessura. Assim como nas eudicotiledôneas, a parede secundária de gramíneas é composta principalmente de celulose, hemicelulose e lignina. Em gramíneas, a principal hemicelulose é a GAX, diferentemente das eudicotiledôneas, em que se encontram xilanas com estrutura distinta (VOGEL, 2008).

A Glucuronoarabinoxilana (GAX) consiste em uma cadeia de xiloses, unidas por ligações β -1,4, substituída por resíduos de arabinofuranose unidos por ligações α -1,2 ou α -1,3. Ocorre também substituição por ácido glucurônico ou 4-O-metilglucorônico (SCHELLER; ULVSKOV, 2010), por meio de ligações α -1,2. Os resíduos de arabinose podem ainda ser substituídos por unidades adicionais de xilose unidas por ligações β -1,2 (WENDE; FRY, 1997), ou ainda unidos a outra arabinose por ligações α -1,3 (PAULY et al., 2013). Um aspecto único de gramíneas é a presença de ácido ferúlico (FA) (FAIK, 2010), unido por ligaçõe éster à posição O5 da arabinofuranose, bem como de ácido *p*-cumárico (*p*CA), embora em menores níveis (BARTLEY et al., 2013; HATFIELD et al., 2009; MUELLER-HARVEY et al., 1986). (Figura 2).

Uma característica importante do ácido ferúlico é a sua capacidade de formar dímeros (diferulatos), por meio de acoplamento oxidativo, mediado por peroxidases. Por meio desse mecanismo, são formadas ligações 5-5, 8-O-4, 8-8, 8-5. A maior parte dos dehidrodímeros de ferulato é formada durante a lignificação, quando há maior disponibilidade de peroxidases e a quantidade de ferulato atinge níveis máximos (GRABBER et al., 1995). Esses dímeros formam ligações cruzadas que interligam cadeias de GAX (HATFIELD et al., 2009; WENDE; FRY, 1997) (Figura 2), o que provavelmente está relacionado ao fortalecimento da parede celular e desaceleração do crescimento (MACADAM; GRABBER, 2002).

Adicionalmente, por meio do acoplamento oxidativo, os monômeros e os dímeros de ferulato podem unir cadeias de GAX à lignina (Figura 2), por meio de ligações éter ou carbono-carbono (GRABBER et al., 1995). Na parede primária, existem evidências de que os ferulatos atuam como sítios de iniciação ou nucleação, onde se inicia a deposição de lignina (RALPH; GRABBER; HATFIELD, 1995). Dessa forma, foi proposto que essas

ligações cruzadas promovem o fim da fase de alongamento, e mudança do desenvolvimento da parede celular primária para a secundária. Após o crescimento, a deposição de ferulatos e diferulatos na lignina continua até os estágios mais avançados de desenvolvimento da parede secundária.

Figura 2 - Estrutura química simplificada do ácido ferúlico (FA) ligado à glucurono<u>a</u>rabino<u>x</u>ilana (GAX) da parede celular de gramíneas



Legenda: A) Ácido ferúlico esterificado à arabinofuranose de GAX. B) Ácido diferúlico unindo por ligações cruzadas FA a GAX. C) Ácido ferúlico ligando a lignina à GAX. Fonte: Adaptado de Oliveira et al., 2014.

Embora *p*-cumaratos sejam encontrados unidos por ligação éster a GAX da mesma forma que os ferulatos, a proporção é bem menor (1 *p*CA para cada 15 FA em palha de cevada) (MUELLER-HARVEY et al., 1986). Diversos estudos em milho falharam em identificar *p*CA ligados a arabinose em tecidos de colmo, utilizando hidrólise ácida branda com TFA para quebrar as ligações da arabinose à xilana, liberando *p*CA-arabinose e FAarabinose conjugados (HATFIELD; MARITA; FROST, 2008; MARITA et al., 2003). Apenas recentemente, o desenvolvimento de um método de acidólise usando uma mistura dioxano/H₂O e HCL 2M permitiu a identificação de pequenos níveis de *p*CA-Ara em colmo de milho (*Zea mays*) (2.5% do total *p*CA-Ara mais FA-Ara) e níveis um pouco maiores em folha (11.7% do total *p*CA-Ara e FA-Ara). Essa análise também encontrou *p*CA em hemicelulose de outras gramíneas como *Brachypodium distachion*, arroz (*Oryza sativa*), *Miscanthus sinensis*, dentre outras, com destaque para Brachypodium, em que cerca de 30% do total de *p*CA-Ara e FA-Ara nas folhas correspondeu a *p*CA-Ara (LAPIERRE et al., 2018). Ainda assim, a maior parte do pCA das gramíneas está ligado a lignina, e não à hemicelulose.

A lignina de gramíneas, assim como nas demais angiospermas, é um polímero composto primariamente das 3 unidades: guaiacil (G), derivadas do monolignol álcool coniferílico; siringil (S), derivadas do álcool sinapílico, e menores quantidades de unidades *p*-<u>h</u>idroxifenil (H), derivadas do álcool <u>p</u>-cumarílico. Os precursores (monolignois) são sintetizados no citoplasma, por meio da via dos fenilpropanoides, e depois são exportados para a parede celular, onde sofrem polimerização radicalar (Figura 3). O polímero é ligado principalmente por ligações éter e carbono-carbono (BAUCHER et al., 1998; BOERJAN; RALPH; BAUCHER, 2003; MOTTIAR et al., 2016).

Figura 3 – Esquema simplificado das principais unidades que constituem a lignina



Fonte: Adaptado de MOTTIAR et al., 2016.

Existem algumas particularidades na composição e estrutura da lignina de gramíneas, além das ligações cruzadas promovidas pelos ferulatos. Uma característica é que de um modo geral, há presença de quantidades significativas de unidades H quando comparada às eudicotiledôneas. Outro aspecto único é a acilação por ácido *p*-cumárico (*p*CA), o que não ocorre nas eudicotiledôneas (GRABBER et al., 2004).

O pCA é primariamente associado a unidades S, embora também possa ser incorporado ao guaiacil (LU; RALPH, 1999), e permanece principalmente como uma extremidade fenólica livre, praticamente não promovendo nenhuma ligação cruzada mediada por acoplamento oxidativo de radical (HATFIELD et al., 2009; RALPH, 2010). Isso porque, embora o pCA seja um composto fenólico, suas reações de acoplamento (*cross coupling*) são ineficientes, e a transferência de radical para os monolignois e extremidades fenólicas livres de lignina é rápida. No entanto, o p-cumarato pode participar de reações de ciclodimerização, resultante de fotoativiação. Esses ciclo-dímeros formados (derivados de ciclobutano) provavelmente promovem ligações cruzadas na lignina, por ser onde a maioria do pCA se encontra. Além disso, ciclo-dímeros também podem ser formados entre pCA e FA e, considerando que o FA está ligado à xilana, podem unir lignina à hemicelulose(GRABBER et al., 2004). No entanto, a significância desses ciclodímeros é difícil de ser estimada, uma vez que as quantidades detectadas são de fato muito baixas, além do fato de que tais dímeros foram detectados após hidrólise alcalina, não sendo possível determinar aonde eles estavam originalmente ligados (HATFIELD; RALPH; GRABBER, 2008).

Embora a função do pCA na lignina não seja completamente compreendida, existem evidências de que atue como um agente de transferência de radical. Isso porque o pCA é rapidamente oxidado por peroxidases e lacases, enquanto o ácido sinapílico, precursor das unidades S, é dificilmente oxidado. Radicais livres de pCA podem rapidamente transferir seu estado oxidado para o álcool sinapílico, que então pode ser incorporado na lignina (HATFIELD; RALPH; GRABBER, 2008; TAKAHAMA; ONIKI, 1997).

1.5 RECALCITRÂNCIA DA PAREDE CELULAR E SUA RELAÇÃO COM OS ÁCIDOS HIDROXICINÂMICOS

Diversos trabalhos relatam que a recalcitrância da parede celular de plantas é afetada por vários fatores, como a cristalinidade da celulose; a quantidade, a composição e a ramificação da hemicelulose; a quantidade e composição de lignina (BUANAFINA, 2009; HIMMEL et al., 2007; JEOH et al., 2007; PAULY; KEEGSTRA, 2008; YANG; WYMAN, 2004). Além desses fatores, tem sido demonstrado nas gramíneas que o conteúdo de ácidos hidroxicinâmicos afeta a recalcitrância. As ligações cruzadas que eles promovem entre cadeias de hemicelulose e lignina contribuem para a baixa digestibilidade das gramíneas por espécies ruminantes e também para a dificuldade de degradação por coquetéis enzimáticos de celulases e xilanases (BUANAFINA, 2009).

Desde a primeira investigação do efeito da proporção *p*CA/FA na gramínea *Lolium* (*ryegrass*) com significativa correlação com a digestibilidade da parede celular (HARTLEY, 1972), outros estudos confirmam a relação direta entre FA e recalcitrância da parede celular (BUANAFINA, 2009; OLIVEIRA et al., 2015).

Em modelos sintéticos contendo hemicelulose ligada à pCA e FA, a concentração de ácidos hidroxicinâmicos foi inversamente proporcional à digestibilidade *in vitro* (JUNG; RALPH; HATFIELD, 1991). Ademais, estudos utilizando modelos de parede celular demonstraram que o aumento nos níveis de dímeros de ferulato reduz a degradação de polissacarídeos estruturais (GRABBER et al., 1995; GRABBER; HATFIELD; RALPH, 1998). Adicionalmente, em paredes celulares de milho sinteticamente lignificadas, foi demonstrado que a redução nas ligações cruzadas entre FA e lignina aumenta significativamente (cerca de 70%) a hidrólise enzimática, sobretudo de xilanas (GRABBER; RALPH; HATFIELD, 1998).

Embora esses trabalhos utilizem modelos sintéticos da parede celular, eles refletem com precisão a natureza da parede celular de gramíneas. Em entrenós maduros das gramíneas *Lolium perenne* e *Phalaris aquatica*, o conteúdo de ácidos hidroxicinâmicos (determinado por extração alcalina quente, sendo principalmente ácido ferúlico) correlacionou negativamente com a digestibilidade *in vitro* (LAM; IIYAMA; STONE, 2003). Ademais, gramíneas da espécie *Bromus inermis Leyss* com menores níveis de FA unido por ligações éter (ou seja, com menos ligações FA-lignina) apresentaram maior digestibilidade dos polissacarídeos da parede celular(CASLER; JUNG, 1999). Esses estudos claramente indicam que o ácido ferúlico e as ligações cruzadas que promovem na parede possuem papel crítico na digestibilidade de gramíneas.

Enquanto o conteúdo de FA e as ligações cruzadas que ele promove tem sido relacionado com à baixa digestibilidade, tem sido demostrado que uma maior incorporação de pCA e menor de FA à hemicelulose é favorável para a conversão. Linhagens transgênicas com aumento na quantidade de pCA especificamente na hemicelulose e redução de FA apresentaram aumento na digestibilidade enzimática (BARTLEY et al., 2013, LI et al., 2018). Ademais, um estudo com linhagens mutantes de Brachypodium verificou que algumas delas possuíam maior digestibilidade, mesmo com maior teor de lignina. É possível que, dentro das diversas alterações que as linhagens apresentaram, o aumento no conteúdo de pCA unido à GAX possa ter contribuído, junto com outros fatores como a diminuição da cristalinidade da celulose, para o aumento na sacarificação (TIMPANO et al., 2015).

A quantidade de ácido *p*-cumárico unida à lignina também pode influenciar na digestibilidade de gramíneas. Como supracitado, o *p*CA é incorporado preferencialmente à unidades S. Estudos com entrenós de milho em diferentes estágios de desenvolvimento demonstraram que os níveis de *p*CA crescem com a lignificação, e que o aumento de *p*CA é acompanhado pelo aumento de unidades siringil liberadas por oxidação com nitrobenzeno (HATFIELD; RALPH; GRABBER, 2008). Dessa forma, o aumento na incorporação de *p*CA pode estar associado com aumento da proporção S:G, o que pode aumentar o rendimento de alguns tipos de pré-tratamentos, como alcalino ou sulfito ácido, devido às diferenças na estrutura desses monolignóis. Como as unidades S são substituídas por metoxila em C5 (Figura 4), ligninas enriquecidas com S são menos entrecruzadas, além de serem menos susceptíveis às reações secundárias de condensação que competem com a degradação (BOERJAN; RALPH; BAUCHER, 2003). Além disso, a incorporação de ácido *p*-cumárico pode aumentar a solubilidade da lignina em pré-tratamentos alcalinos, uma vez que aumenta a quantidade de extremidades fenólicas livres.

Em suma, esses estudos sugerem que a seleção ou engenharia de plantas com reduzidas ligações cruzadas ferulato-hemicelulose e ferulato-lignina, ou aumento de pCA ligado a hemicelulose ou ao álcool sinapílico, pode reduzir a recalcitrância das plantas.

1.6 COMPOSIÇÃO DA PAREDE CELULAR E RECALCITRÂNCIA EM CANA-DE-AÇÚCAR

O estudo da composição da parede celular (celulose, hemicelulose, lignina e ácidos hidroxicinâmicos) de colmo maduro em 11 híbridos de cana-de-açúcar demonstrou uma composição de ~35%-45% de celulose, ~25-32% hemicelulose, ~16-25% lignina; ~3.9-5.8% de pCA; ~0.8%-1.2% de FA. Em todas as amostras houve a predominância de pCA em relação a FA, em uma proporção pCA/FA de 3.5 (híbrido #89) a 6.4 (híbrido #121) (MASARIN et al., 2011). Proporção semelhante foi encontrada para a espécie ancestral *Saccharum spontaneum* e o híbrido comercial RB867515, cerca de 1% de FA e 3 a 4% de pCA em entrenós maduros (DE CARLI POELKING et al., 2015). Del Rio *et al.* (2015) encontraram razões pCA/FA semelhantes em bagaço (5.2) e palha (3.0), em estudo com variedade comercial de cana-de-açúcar. Entretanto, na análise da composição apenas em lignina (*lignina de madeira moída*, MWL), a proporção sofreu aumento significativo (31.1 em bagaço e 12.8 em palha). Isso é consistente com o fato de pCA ser incorporado

principalmente à lignina, enquanto o FA predomina na hemicelulose (DEL RÍO et al., 2015).

Os elevados níveis de pCA, superiores aos de FA, observados em cana-de-açúcar também são observados em espécies evolutivamente próximas, como sorgo e milho, que possuem altos níveis de pCA éster (3 a 4% em massa). Outras gramíneas C4 como *switchgrass* possuem níveis significativos de pCA, porém muito mais baixos que cana-deaçúcar, sorgo e milho, próximos dos níveis de gramíneas C3 como Brachypodium (1.2-1.5%) (HATFIELD et al., 2009; RANCOUR; MARITA; HATFIELD, 2012). É possível que essas diferenças no conteúdo de *p*-cumarato estejam apenas relacionadas ao porte dessas plantas, ou seja, a lignina de gramíneas maiores e de colmo mais largo apresentam mais *p*-cumarato (HATFIELD; RANCOUR; MARITA, 2017).

Em gramíneas, a composição da parede celular e a recalcitrância variam com o tipo celular e a região do entrenó. Estudos de microscopia UV em colmos de cana-de-açúcar evidenciam que as células de vasos são mais lignificadas, seguidas das fibras, e que ambas contêm quantidades significativas de ácidos hidroxicinâmicos. Nas células de parênquima, a lignificação é mínima, predominando ácidos hidroxicinâmicos (HE; TERASHIMA, 1990; SIQUEIRA et al., 2011). Siqueira et al. (2011) correlacionaram a distribuição topoquímica dos ácidos hidroxicinâmicos e lignina com a digestibilidade. Na região da medula predominaram células de parênquima pouco lignificadas e com menores quantidades de ácidos hidroxicinâmicos quando comparadas com células parenquimáticas do córtex. Os vasos e fibras presentes também apresentaram menos lignina do que os do córtex. Ademais, a região da medula foi mais facilmente degradada por celulases comerciais. Já o córtex apresentou maiores quantidades de vasos e fibras com intensa lignificação e presença de ácidos hidroxicinâmicos, possuindo maior recalcitrância. Entretanto, a remoção dos ácidos hidroxicinâmicos e lignina por tratamento com hipoclorito aumentou significativamente a conversão de celulose por celulases (SIQUEIRA et al., 2011).

Costa *et al.* (2013) avaliaram a composição de lignina, hemicelulose e ácidos hidroxicinâmicos de três híbridos em diferentes frações do entrenó: A fração mais externa, o córtex, a interface córtex-medula e a medula. De um modo geral, a quantidade de lignina e hemicelulose aumentou da medula em direção ao córtex. O conteúdo de ácido *p*-cumárico foi maior do que o ferúlico em todas as amostras, e o total de ácidos hidroxicinâmicos unidos por ligações éster variou de 1,5 a 2% no córtex e medula dos
híbridos. O estudo da digestibilidade com celulases comerciais demonstrou que a recalcitrância foi maior na fração mais externa e no córtex (COSTA et al., 2013).

Outro fator que influencia a recalcitrância da biomassa é o estágio de maturação dos tecidos, uma vez que está diretamente relacionado com a composição da parede celular. Em cana-de-açúcar, a lignificação dos vasos ocorre no início do estágio de maturação, seguida da lignificação das fibras. Os ácidos hidroxicinâmicos são depositados principalmente durante a lignificação (HE; TERASHIMA, 1990).

O estágio de maturação de um entrenó depende de sua posição no colmo. A canade-açúcar cresce iniciando novos entrenós no ápice, que começam a se alongar quando a folha abaixo se expande completamente. Então, o caule é formado por entrenós sucessivamente mais velhos, do topo para a base (LINGLE; THOMSON, 2012). Ademais, dentro de cada unidade fitomérica (composta de uma folha, bainha e entrenó abaixo), a maturação e alongamento acontecem sempre na direção basipetal. Dessa forma, a ponta da folha irá amadurecer primeiro, enquanto as células da base da folha ainda estão imaturas. O amadurecimento ocorrerá então no sentido base da folha, bainha, e atingirá o entrenó abaixo, sendo que a base do entrenó é a última a amadurecer (RAE; MARTINELLI; DORNELAS, 2014).

Dessa forma, a composição da parede celular e a recalcitrância da cana-de-açúcar dependem do tipo celular, do tecido e seu estágio de maturação. Essa complexidade deve ser levada em conta nas análises estruturais e de digestibilidade dos materiais lignocelulósicos.

1.7 FAMÍLIA GÊNICA BAHD ACILTRANSFERASES

Os genes envolvidos na incorporação dos ácidos hidroxicinâmicos na parede celular de gramíneas foram identificados como sendo da família BAHD aciltransferases (MITCHELL et al., 2007, BARTLEY et al., 2014). As enzimas BAHD aciltransferases constituem uma superfamília gênica responsável pela acilação dependente de acil-CoA de metabólitos secundários, resultando na formação de ésteres e amidas (D'AURIA, 2006). O nome desta superfamília gênica surgiu após a caracterização funcional de quatro membros encontrados em distintas eudicotiledôneas: BEAT ou benzil álcool O-acetiltransferase da *Clarkia breweri*; AHCTs ou antocianina Ohidroxicinamoiltransferase de *Petunia*; HCBT ou antranilato N-hidroxicinamoil/ benzoiltransferase de *Dianthus caryophyllus* e DAT ou

deacetilvindoline 4-O-acetiltransferase de *Catharanthus roseus* (ST-PIERRE; LUCA, 2000). Diversas enzimas dessa família já foram caracterizadas em plantas, com papel na modificação de alcalóides ou antocianinas, na síntese de compostos voláteis (relacionados à defesa ou atração de polinizadores), bem como na síntese da parede celular.

D'Auria *et al.* (2006), em estudos filogenéticos com enzimas caracterizadas em diversas espécies de plantas, dividiram essa superfamília gênica em cinco grandes clados (I a V) (D'AURIA, 2006). Mais recentemente, estudos filogênomicos ampliaram essa divisão para oito clados (TUOMINEN; JOHNSON; TSAI, 2011)(Ia, Ib, II, IIIa, IIIb, IV, Va e Vb), baseado nas sequencias BAHD de quatro eudicotiledôneas, *Arabidopsis, Populus* (álamo), *Medicago* (luzerna cortada), *Vitis* (uva) e uma gramínea, *Oryza* (arroz). As análises filogenéticas sugeriram que o clado IV é específico de gramíneas, uma vez que apenas sequências de arroz foram encontradas. Já o clado IIIa, ao contrário seria específico de eudicotiledôneas (Figura 4).

Alguns desses clados podem ser classificados de acordo com o substrato que atuam, enquanto outros não têm atividade bioquímica específica. As enzimas do clado Ia, por exemplo, atuam sobre glicosídeos fenólicos, principalmente antocianinas, enquanto as do Clado IIIa utilizam diversos álcoois como substrato. Já o clado Va pode ser dividido em subgrupos, que aceitam uma gama diferente de substratos. Um deles participa da síntese de ésteres voláteis, enquanto outro é o clado de Mitchell, relacionado à incorporação dos ácidos hidroxicinâmcos. Já o clado Vb consiste de enzimas que usam hidroxicinamoil coA e benzoil coA como doadores de acila (HCT/HQT), que participam da via de síntese da lignina. Já os clados Ib, II, IIIb e IV apresentam poucas enzimas caracterizadas. O clado Ib contém apenas uma sequência caracterizada em Arabidopsis envolvida com a síntese de cutina; o clado II apresenta duas enzimas envolvidas na síntese de graxas (em arabidopsis e milho); IIIb contém uma enzima em petúnia relacionada a compostos voláteis; enquanto no clado IV apenas uma sequência foi caracterizada em cevada, que utiliza aminas como substratos, relacionada a síntese de agamatina (composto de defesa) (BONTPART et al., 2015; D'AURIA, 2006; RANI et al., 2010; TUOMINEN; JOHNSON; TSAI, 2011).

Figura 4 – Filogenia de proteínas BAHD aciltransferases bioquimicamente caracterizadas e putativas identificadas nos genomas de *Populus* (Álamo), *Arabidopsis, Medicago* (Luzerna cortada), *Oryza* (Arroz) e *Vitis* (Uva).



Legenda: A- Filogenia inferida pela máxima verossimilhança, B- Gráfico da quantidade de sequências de cada espécie, por clado. As cores representam as espécies em C. C- Gráfico da quantidade de sequências em cada clado, por espécie. Fonte: TUOMINEN et al., 2011.

A maior parte das enzimas BAHD apresenta dois motivos conservados: HXXXD e DFGWG, sendo o primeiro fundamental para a atividade catalítica. Tuominen *et al.* (2011), em análise de bioinformática para a identificação de motivos conservados nos oito clados, não identificaram o motivo conservado HXXXD, provavelmente, em vista da variação do resíduo X. Entretanto, o motivo DFGWG foi encontrado nos oito clados. Outros motivos conservados também foram identificados, bem como motivos específicos de cada clado (TUOMINEN; JOHNSON; TSAI, 2011).

1.8 ENZIMAS ENVOLVIDAS NA INCORPORAÇÃO DOS ÁCIDOS HIDROXICINÂMICOS NA PAREDE CELULAR

Os genes envolvidos na incorporação dos ácidos hidroxicinâmicos na parede celular começaram a ser caracterizados por Mitchell *et al.* (2007). Nesse trabalho, eles

realizaram uma busca por genes que eram mais abundantemente expressos em gramíneas do que em eudicotiledôneas, e que, portanto, poderiam estar envolvidos na síntese de GAX. Para isso, foi utilizada uma abordagem de bioinformática, que consistiu incialmente na identificação de grupos de genes ortólogos em arroz e Arabidopsis, seguida do mapeamento desses genes em bibliotecas de EST de espécies de cereais (gramíneas) e eudicotiledôneas. Um grupo de 12 *loci* de arroz da família PFAM PF02458 (Figura 5, genes *OsAt1* a *OsAt12*), com apenas um ortólogo em Arabidopsis, foi significativamente mais representado em bibliotecas de cereais do que nas de eudicotiledôneas. Essa superfamília corresponde a coA aciltransferase ou BAHD aciltransferase, como é denominada em plantas. Como tais enzimas catalisam a transferência de grupos acila, dependente de coenzima A, foi sugerido que esse grupo de genes poderia estar envolvido na específica acilação da hemicelulose de gramíneas por hidroxicinamatos (MITCHELL; DUPREE; SHEWRY, 2007).

Análises filogenéticas de proteínas BAHD de diversas espécies de plantas (monocotiledôneas e eudicotiledôneas) demonstraram que os genes identificados por Mitchell pertencem a um clado (Clado de Mitchell) que sofreu expansão gênica em gramíneas relativamente às eudicotiledôneas. Nesse grupo, 20 genes de arroz foram encontrados, nomeados OsAt1 a OsAt20 (Figura 5). Desde a primeira investigação por Mitchel (2007), diversos estudos em arroz, Brachypodium, milho e Setaria associaram genes desse clado à incorporação de pCA à hemicelulose, FA à hemicelulose ou pCA à lignina.

Figura 5 – Filogenia do clado de Mitchell inferida pelo método bayesiano utilizando sequências identificadas em arroz e enzimas BAHD aciltransferases caracterizadas funcionalmente



Fonte: BARTLEY et al., 2013.

Em estudo com 17 linhagens de arroz *activation tagged*, cujas inserções tiveram como alvo 12 dos 20 genes aciltransferase, quatro tiveram alterações no conteúdo de FA e pCA na parede celular, com alteração nos níveis de expressão de *OsAt7*, *OsAt15*, *OsAt5* e *OsAt10*. Em uma das linhagens, houve aumento (~200 vezes) na expressão de *OsAt10*, e os níveis de FA reduziram (~60% menos), enquanto os de pCA aumentaram (300% a mais) tanto na bainha como na lâmina das folhas. O aumento dos níveis de expressão de *At10* (~ 3000 vezes), utilizando o promotor constitutivo de ubiquitina de milho (UBI), resultou na diminuição de FA/pCA em folhas jovens e na palha madura (45% mais pCA e 20% menos FA), sendo que o aumento na quantidade de pCA foi restrito à GAX, não atingindo a lignina. Esses resultados sugerem que o gene *OsAt10* tem potencial função na adição de pCA à GAX. Como esperado pela diminuição de FA, essas linhagens apresentaram aumento nos rendimentos de sacarificação de 20% para hidrólise com celulases, e de 40% para o tratamento biológico com fungos, o que demonstra o potencial do gene *OsAt10* para aplicações biotecnológicas (BARTLEY et al., 2013). Recentemente, a superexpressão de

OsAt10 em *switchgrass (Panicum virgatum L)* resultou no aumento de *p*CA éster (18-28%) e diminuição de FA éster (21-27%) apenas em folhas verdes (jovens). Em folhas senescentes, houve redução de FA (~19%) e *p*CA (~15%). Como resultado a eficiência de sacarificação aumentou 30% (em folhas senescentes) a 40% (folhas verdes) (LI et al., 2018).

Os possíveis genes envolvidos com a incorporação de FA à GAX foram inicialmente identificados por meio do silenciamento simultâneo de quatro genes, usando técnicas de RNA de interferência (RNAi) (PISTON et al., 2010). As plantas transformadas com construções de silenciamento de *OsAt6* a *OsAt10* apresentaram redução significativa dos níveis de transcritos desses genes tanto nos caules (20-60% do controle) como nas folhas (40-60% do controle), resultando em uma diminuição dos níveis de FA na parede celular das folhas em cerca de 20-30%. Esse trabalho sugeriu que pelo menos um desses genes deveria estar envolvido na incorporação de FA à GAX, excluindo *OsAt10* cuja provável função é a incorporação de *p*CA, como supracitado (PISTON et al., 2010).

Mais recentemente, estudos evidenciaram o envolvimento dos genes Atl e At9 na incorporação de FA na hemicelulose. O silenciamento por RNAi do gene BdAt1, ortólogo de OsAt1 em Brachypodium, levou à diminuição no conteúdo de FA de até 35% em folhas e colmos. Ademais, a super-expressão desse gene (utilizando promotor 35S) resultou no aumento dos níveis de FA até 58% e 47% em folhas e colmos, respectivamente (BUANAFINA et al., 2016). Já o silenciamento do gene de Setaria (SvBAHD01), ortólogo de OsAt9, resultou numa redução de 60% da quantidade de FA na GAX no colmo. Essa diminuição provocou um aumento na eficiência de sacarificação em torno de 40%-60%. O silenciamento por RNAi de BdBAHD01, ortólogo de OsAt9 em Brachypodium, também resultou numa diminuição no conteúdo de ferulatos na GAX, porém pequena (10-20% no colmo), provavelmente devido à maior expressão de um gene com redundância funcional (DE SOUZA et al., 2018). O ortólogo de At9 em cana-de-açúcar (SacBAHD01) também foi recentemente identificado com função na incorporação de FA, suportando a caracterização funcional anterior na gramínea modelo Setaria. O silenciamento por RNAi de SacBAHD01 levou à diminuição de até 50% de FA na palha, e resultou em aumentos no rendimento de sacarificação de até 24% (DE SOUZA et al., 2019).

Os genes responsáveis pela incorporação de pCA à lignina já foram caracterizados em arroz, Brachypodium e milho. Em arroz, foram encontradas enzimas com atividade pcumarilCoA monolignol transferase, partindo da hipótese de que os genes que codificam enzimas com tal atividade deveriam pertencer à família BAHD e ter expressão correlacionada com as dos genes da via de síntese de lignina (WITHERS et al., 2012). A sequência LOC_Os01g18744 (*OsAt4*, como em Bartley et al., 2013, e denominada de *OsPMT*) foi a que mais correlacionou com genes tais como 4CL (4-cumaratoCoA ligase) e *COMT* (Cathecol-o-metiltransferase), que participam da síntese dos monolignois. Ensaios enzimáticos com a proteína purificada confirmaram a atividade esperada. Além disso, de todos os monolignois testados (álcool *p*-cumarílico, álcool sinapílico e álcool coniferílico), a enzima *OsPMT* teve maior afinidade pelo álcool sinapílico, confirmando a preferência da ligação de *p*CA a unidades S da lignina, com relação as unidades G (WITHERS et al., 2012).

O ortólogo de OsPMT foi caracterizado em Brachypodium distachion, denominado BdPMT (PETRIK et al., 2014). Mutações deletérias desse gene resultaram na redução da incorporação de pCA na lignina para níveis próximos de zero. Entretanto, não houve alterações na quantidade de lignina ou no crescimento da planta, apenas um ligeiro decréscimo na incorporação de unidades S. O efeito no rendimento de sacarificação foi mínimo, sugerindo que a abolição de pCA na lignina não impacta na recalcitrância da parede celular. Adicionalmente, linhagens transgênicas com o gene BdPMT silenciado por RNAi apresentaram uma redução no conteúdo de pCA especificamente na lignina (não em GAX), demonstrando a especificidade da enzima pela acilição da lignina. Essas linhagens BdPMT RNAi também demonstraram um pequeno decréscimo na proporção S:G, mas nenhuma alteração significativa na quantidade de lignina ou no crescimento da planta, tampouco nos rendimentos de sacarificação, corroborando com os resultados da linhagem mutante com perda de função. Ademais, linhagens transgênicas com super-expressão de BdPMT apresentaram um aumento de pCA (3,4 vezes) especificamente na lignina, um aumento significativo na proporção S:G(~40%), 23% menos lignina e 33% de aumento nos rendimentos de sacarificação. Uma possível explicação para a diminuição da quantidade de lignina é que o excesso na atividade do PMT poderia tornar os monolignois menos disponíveis para a biossíntese de lignina. O aumento na sacarificação provavelmente está relacionado com a redução no conteúdo de lignina. No entanto, essa diminuição afetou o desenvolvimento da planta (PETRIK et al., 2014).

Uma enzima com atividade *PMT* também foi caracterizada em milho, denominada pCAT (*p*-coumaroyl Coa: Hydroxicinnamoyl alcohol transferase), por meio de uma abordagem de proteômica (MARITA et al., 2014). Diferentemente de *OsPMT*, ensaios enzimáticos demonstraram que *pCAT* tem alta afinidade não apenas por ácido *p*-cumárico,

mas também por FA, o que requer estudos adicionais para verificar se em milho existe incorporação de ferulato-monolignol na lignina. Quanto à afinidade por monolignois, pCAT também demonstrou preferência pelo álcool sinapílico. Adicionalmente, a supressão do gene pCAT por RNAi resultou em plantas com menores níveis de pCA tanto na folha como no caule. Nenhum padrão de correlação da quantidade de lignina (aumento ou diminuição) foi verificado nessas linhagens transgênicas. Corroborando com os resultados de Petrik *et al.* (2014) em Brachypodium, as linhagens de milho pCAT:RNAi apresentaram diminuição significativa na proporção S:G (MARITA et al., 2014).

Como a acilação da lignina por pCA é exclusiva de gramíneas, foi estudado o efeito dessa incorporação em plantas que não possuem tal atividade (SMITH et al., 2015). O gene *OsPMT* foi inserido em duas eudicotiledôneas, *Arabidopsis thaliana* e *Populus alba×grandidentata*. De fato, monolignol *p*-cumarato conjugados foram produzidos em níveis que se aproximam dos de monocotiledôneas, demonstrando que monolignois conjugados que não são inerentes ao metabolismo da planta podem ser incorporados na parede celular. No entanto, inesperadamente, *p*CA foi mais frequentemente unido a álcool *p*-cumarílico do que ao álcool sinapílico, mesmo existindo quantidades significativas de unidades S em Arabidopsis e especialmente em poplar. Os motivos para isso ainda precisam ser investigados, podendo estar relacionado a um aumento na disponibilidade de álcool *p*-cumarílico quando *p*CA está disponível para a reação com *PMT* (SMITH et al., 2015).

Ademais, pCA éster foi incorporado na lignina de Arabidopsis ao se expressar o gene *BdPMT* utilizando o promotor constitutivo de ubiquitina de milho, porém os níveis foram baixos (TIMPANO et al., 2015). Para atingir níveis iguais ao de ligninas de colmo maduro de gramíneas, dois genes *BdPMT1* (PETRIK et al., 2014) e *BdPMT2* (ortólogo de *OsAt8*, com função putativa por ser expresso em altos níveis no colmo maduro) foram inseridos em Arabidopsis, sob controle do promotor *AtC4H*, que direciona a expressão para os vasos. Com essa abordagem, a acilação por *p*CA atingiu níveis de 8 a 9% em massa, tanto para as linhagens expressando *BdPMT1* como *BdPMT2*, sem nenhum impacto no desenvolvimento da planta. Esses resultados sugerem que em Brachypodium, duas enzimas da família BAHD apresentam função *p*-cumarato monolignol transferase, existindo redundância funcional. Além disso, demonstra a possibilidade de usar promotores tecido específicos para promover altos níveis de acilação por *p*CA da lignina, o que interfere na sua solubilidade, podendo aumentar a eficiência de pré-tratamentos alcalinos. Para melhor entender como a incorporação de *p*CA na lignina afeta outros fatores estruturais, além da

simples acilação no carbono gama, o gene *BdPMT1* também foi inserido em linhagens de Arabidopsis mutantes *fah1* (deficientes em ferulato 5-hidrolase), que não conseguem formar álcool sinapílico. Quantidades substanciais de *p*CA foram incorporadas à unidades G, confirmando a capacidade das enzimas *PMT* de também induzirem a incorporação de *p*CA-álcool coniferílico. A construção também foi inserida em mutantes *ccr1g*, que possuem menor quantidade de lignina e crescimento prejudicado, e a incorporação de *p*CA ocorreu normalmente, embora em menores níveis. Além do *p*CA, houve a incorporação de FA éster na lignina dessas linhagens, o que provavelmente se deve ao acúmulo de feruloilcoA que existe nos mutantes com perda de função *ccr*. Isso demonstra a capacidade do gene *BdPMT* utilizar também ferulato como substrato, o que corrobora com os resultados de atividade enzimática *in vitro* da enzima *pCAT* (MARITA et al., 2014), mas difere do de *OsPMT* (WITHERS et al., 2012).

Mais recentemente, foi identificado que a enzima OsAt5 ou OsFMT está relacionada à uma atividade que não se esperava ocorrer naturalmente em gramíneas, que é a incorporação de ferulato a monolignois. Técnicas de DFRC (derivatization followed reductive demonstraram embora cleavage) que, em pequenas quantidades, eudicotiledôneas apresentam de fato unidades G-FA na lignina, e grande parte das monocotiledôneas incorpora S-FA. A super-expressão do gene OsAt5 (promotor ubiquitina de milho e linhagens activation tagged) levou a um aumento da quantidade de FAmonolignol liberada pelos ensaios DFRC (até ~7 vezes), sugerindo a participação desse gene na incorporação de FA na lignina de gramíneas, e demonstrando que a lignina também é constituída de conjugados FA-monolignol, e não apenas acilada por ácido pcumárico como se acreditava (KARLEN et al., 2016).

Em suma, diversos trabalhos identificaram genes relacionados à incorporação de ácidos hidroxicinâmicos em gramíneas e verificaram seu impacto na estrutura da parede celular, como sumarizado (Tabela 1). Entretanto, vários aspectos ainda precisam de mais investigações. A maioria desses genes foi estudada em gramíneas modelo. Apenas o gene *At9* foi estudado em cana-de-açúcar e o gene *At4* foi estudado em milho. O gene *At10*, por exemplo, com potencial de reduzir a recalcitrância da biomassa, só foi caracterizado em arroz. A identificação e caracterização nas espécies de interesse econômico é importante, uma vez que as espécies podem apresentar diferenças nos mecanismos de incorporação dos ácidos hidroxicinâmicos, podendo apresentar diferentes níveis de expressão dos genes desse subclado, ou ainda, redundâncias funcionais específicas da espécie. Adicionalmente,

embora os genes *At10* e *At9* sejam fortes candidatos à atividade *p*-cumarilCoA e feruloilCoA transferase, respectivamente, mais evidências são necessárias quanto à atuação dessas enzimas nesses substratos. Ademais, outros genes desse clado também podem ter função relacionada à incorporação dos ácidos hidroxicinâmicos, de modo que ainda não está claro se há redundância funcional. Ainda não se sabe sobre os aspectos relevantes da incorporação, por exemplo, se FA ou *p*CA são adicionados inicialmente a um intermediário no citoplasma, ou diretamente à arabinose, e posteriormente incorporação a hemicelulose, ou ainda se ocorre acilação direta das cadeias de GAX. Também não é claro como ocorre a distribuição de feruloil-coA e *p*-cumaroil-coA entre incorporação à hemicelulose ou lignina. Dessa forma, muitos estudos ainda são necessários para compreender a genética das enzimas envolvidas na incorporação de ácidos hidroxicinâmicos, em diversas espécies de gramíneas.

Espécie	Gene	Expressão	Fenótipo	Efeito	Ref.
-		•	•	digestão	
Oryza sativa	OsAt10	Super-expressão	20% diminuição	Aumento	Bartley et
		(UBI milho)	de FA e 45%	de até	al., 2013
			aumento de <i>p</i> CA	40%	
			na GAX (Palha)		
Oryza sativa	OsAt6 a	Silenciamento	20-30% menos	NA	Piston et
	OsAt10	por RNAi	FA na GAX em		al., 2012
			folhas		
Brachypodium	BdAt1	Silenciamento	35% menos FA	NA	Buanafina
disatachion		por RNAi	em folhas e caule		et al., 2016
		Super-expressão	58-47% de	NA	
		355	aumento de FA		
			nas folhas e no		
			caule		
Brachypodium	BdAt9	Silenciamento	10-20% de	NA	De Souza
distachion		por RNAi	redução de FA		et al., 2018
			no colmo		
Setaria viridis	SvAt9	Silenciamento	60% de	40-60%	De Souza
		por RNAi	diminuição de	de	et al., 2018
			FA no colmo	aumento	
Saccharum	SacAt9	Silenciamento	50% de	Até 24%	De Souza
spp.		por RNAi	diminuição de	de	et al., 2019
			FA na palha	aumento	
Brachypodium	BdPMT	Silenciamento	0,1 X pCA na	Nenhum	Petrik et
distachion		por RNAi	lignina,	efeito	al., 2014
			diminuição S:G		
		Mutante perda	Sem <i>p</i> CA na	Nenhum	
		de função (azıda	lignina	efeito	
		de sódio)		A	
		Super-expressao	3X mais p CA na	Aumento	
		(UBI milho)	lignina, aumento	de 33%	
			de S:G, 25%		
7	T	Silanaiamanta	Diminuição	NT A	Monito at
Zea mays	pCAT		\mathbf{D}	NA	$\frac{1}{2}$
		poi KINAI	pCA ha hghha		al., 2014
			diminuição S·G		
Orwza satiwa	OsA+5	Super-expressão	Aumento de FA	NΔ	Karlen et
Οι για δαπνα	011	(IIRI milho) e	na lignina (~7X)		a = 2016
	$O_{S}FMT$	activation	na iiginna (~7A)		an, 2010
	U 51 M1	taooino			
		iuzzinz			

Tabela 1 – Genes identificados na família BAHD relacionados à incorporação de ácidos hidroxicinâmicos

Fonte: Autoria própria

2 JUSTIFICATIVAS

A cana-de-açúcar (Saccharum spp.) é uma das espécies mais produtivas e eficientes na produção de etanol (40%) comparado a outras gramíneas (BYRT, C. S.; GROF, C. P. L.; FURBANK, R. T., 2011; VAN DER WEIJDE et al., 2013). Entretanto, para alcançar uma melhor eficiência na conversão do material lignocelulósico, aumentando a produção de etanol, biomateriais ou alimentação animal, faz-se necessário conhecer os genes envolvidos na regulação dos ácidos hidroxicinâmicos, reduzindo a recalcitrância da parede celular. Atualmente, membros da superfamília gênica BAHD foram caracterizados em gramíneas demonstrando a alteração do conteúdo dos ácidos hidroxicinâmicos e aumento da sacarificação. Os genes At10 e At4/PMT/pCAT foram identificados como relacionados à adição de pCA à hemicelulose e à lignina, respectivamente (BARTLEY et al., 2013; MARITA et al., 2014; PETRIK et al., 2014), tendo potencial para redução da recalcitrância. No entanto, nenhum destes genes foi caracterizado em cana-de-açúcar. Ademais, embora muitos estudos tenham mostrado a correlação do conteúdo de ácidos hidroxicinâmicos com a recalcitrância da parede celular, ainda faltam informações a respeito do efeito da incorporação dos ácidos ferúlico e cumárico. Detalhes importantes sobre o impacto dos genes At10 e At4 na composição e estrutura dos polímeros da parede celular ainda precisam ser evidenciados. Pouco se sabe sobre a relação entre as vias de incorporação de pCA na hemicelulose e na lignina.

Considerando a importância e o potencial bioenergético da cana-de-açúcar, é necessário identificar os genes responsáveis pela síntese dos ácidos hidroxicinâmicos nessa espécie, bem como compreender melhor o papel desses genes na biossíntese da parede celular de gramíneas. Esse conhecimento poderá contribuir com a futura aplicação biotecnológica desta via metabólica para a redução da recalcitrância da parede celular aumentando as taxas de sacarificação.

3 OBJETIVOS

3.1 OBJETIVOS GERAIS

Considerando que pouco se sabe sobre os genes envolvidos na incorporação dos ácidos hidroxicinâmicos em cana-de-açúcar, o objetivo geral deste trabalho é a identificação dos genes da superfamília gênica BAHD envolvidos com a incorporação dos ácidos hidroxicinâmicos, análise do padrão de expressão desses genes em cana-de-açúcar, e caracterização funcional dos genes *ScAt10 e ScAt4*.

3.2 OBJETIVOS ESPECÍFICOS

Identificação, anotação e análise filogenética dos genes membros da superfamília gênica BAHD (clado Va) em Arabidopsis, Brachypodium, arroz, Setaria, milho, sorgo e em cana-de-açúcar;

Análise de expressão *in silico* da expressão desses genes em diferentes órgãos de cana-de-açúcar e por RT-qPCR em híbridos de cana-de-açúcar com composição contrastante de ácidos hidroxicinâmicos (H89 e H321);

Clonagem das sequências identificadas *ScAt10* e *ScAt4* em vetores Gateway de super-expressão (Ox) usando promotor ubiquitina de milho (UBI) e silenciamento (RNAi), contendo as bordas de T-DNA para transformação em linhagens da planta modelo *Zea mays* (milho);

Caracterização fenotípica e molecular de linhagens transgênicas de milho contendo os vetores de super-expressão (Ox) e silenciamento (RNAi) para *ScAt10* e *ScAt4*.

4 MÉTODOS

4.1 IDENTIFICAÇÃO DOS GENES DA FAMÍLIA BAHD ACILTRANSFERASES EM CANA-DE-AÇÚCAR E GRAMÍNEAS RELACIONADAS

A identificação dos genes da família BAHD foi conduzida através de bioinformática em larga escala, nas gramíneas: Arroz (*Oryza sativa*), brachypodium (*Brachypodium distachyon*), milho (*Zea mays*), sorgo (*Sorghum bicolor*), Setaria (*Setaria italica*) e cana-de-açúcar (*Saccharum spp.*), e na eudicotiledônea modelo Arabidopsis (*Arabidopsis thaliana*).

Para identificação das sequências de gramíneas relacionadas (arroz, Brachypodium, milho e sorgo) e Arabidopsis, foram baixados os bancos genômicos e de proteínas preditas dessas espécies, disponibilizados no site phytozome (https://phytozome.jgi.doe.gov/pz/portal.htmL). Para arroz, Brachypodium e milho, foram utilizadas as versões Oryza sativa 204 v7.0, B distachion 283 e Z mays 181, respectivamente, obtidas no segundo semestre de 2015. Já para sorgo, setaria e Arabidopsis os dados já foram atualizados fazendo uma busca nas versões mais recentes dos genomas S bicolor 313 v3.1, S italica 312 v2.2 e A thaliana 167 TAIR 10, respectivamente. A busca foi feita inicialmente por meio dos softwares BLASTn e BLASTp, utilizando como isca sequências transferases já identificadas na literatura. As proteínas obtidas pelo BLASTp foram utilizadas para nova busca nos referidos bancos usando o software HMMER (Hidden Markov Model). As sequências identificadas foram usadas para construção do perfil HMM criando um arquivo com extensão.hmm usando o algoritmo HMMER build. Esse arquivo foi usado para busca nos bancos de dados usando HMMER search. As sequências obtidas nessa etapa foram anotadas, alinhadas e submetidas a análise filogenética pelo método Neighbor joining (descrito abaixo), para escolha de um representante de cada clado. Esses representantes foram utilizados para nova busca usando o algoritmo JackHMMER, confirmando as sequências já obtidas ou obtendo novas sequencias. Todas as proteínas identificadas tiveram domínio verificado pelo PFAM v30.0 (EL-GEBALI et al., 2019). Apenas as sequências que apresentaram o domínio transferase (PF02458) foram anotadas e submetidas às análises posteriores. Um alinhamento foi feito com as proteínas anotadas usando o software MUSCLE (Multiple Sequence Comparison by Log-Expectation) e analisado no programa BioEdit. Uma inspeção visual foi feita para identificar quais sequências possuíam o domínio HXXXD (presente nas proteínas da família BAHD) conservado.

Para cana-de-açúcar, a identificação foi feita pelo mesmo método, porém usando um banco de sequências expressas (EST), três bancos transcriptômicos e dois bancos genômicos. Foi utilizado o banco de sequências expressas do SUCEST (VETTORE et al., 2003), cujos dados foram disponibilizados pelo Dr. Renato Vicentini. Os dados de RNAseq a partir de entrenó (VICENTINI et al., 2015) e de folha (CARDOSO-SILVA et al., 2014) de diferentes variedades de cana-de-açúcar comerciais também foram cedidos pelo colaborador Dr. Renato Vicentini. Os dados de RNA-seq a partir de raiz foram disponibilizados pelo grupo do Dr. Marcos Silveira Buckeridge (IB/USP). Foram utilizados o genoma parcial e as proteínas preditas do cultivar SP80-3280 (RIAÑO-PACHÓN; MATTIELLO, 2017), disponibilizado pelo colaborador Dr. Diego Riano Pachón e utilizou-se o genoma monoploide de referência, bem como as proteínas preditas, do cultivar R570 (GARSMEUR et al., 2018).

As sequências expressas e os transcritos de cana-de-açúcar foram traduzidos em proteína utilizando os *softwares EXPASY translate* (SWISS INSTITUTE OF BIOINFORMATICS, 2019) e *EMBOSS Transeq* (MADEIRA et al., 2019). A seleção da *ORF (Open Reading Frame)* correta foi feita por meio da verificação do domínio no site PFAM (EL-GEBALI et al., 2019). Apenas as sequências com domínio PFAM *transferase* (PF02458) foram anotadas.

Todas as sequencias de aminoácidos foram alinhadas pelo algoritmo *MUSCLE*. Análises filogenéticas preliminares foram conduzidas no programa MEGA (*Molecular Evolutionary Genetics Analysis*) (TAMURA et al., 2013) pelo método *Neighbor joining*, utilizando modelo de substituição *p-distance*, tratamento de deleções *pairwise deletion* e teste pelo método *bootstrap* com 1000 replicações. Para classificação das sequências entre os clados da família BAHD já nomeados na literatura (D'AURIA, 2006; TUOMINEN; JOHNSON; TSAI, 2011), filogenias foram inferidas pelo método da máxima verossimilhança usando o algoritmo *PhyML3.1* (GUINDON et al., 2010) no *software Seaview* (GOUY; GUINDON; GASCUEL, 2010). Inicialmente o algoritmo *model test* do programa MEGA foi utilizado para escolha do melhor modelo. O teste foi executado pelo método aLRT.

4.2 ANÁLISE *IN SILICO* DO PADRÃO DE EXPRESSÃO DAS SEQUÊNCIAS IDENTIFICADAS EM CANA-DE-AÇÚCAR

Para análise do padrão de expressão *in silico* das sequências BAHD identificadas em cana-de-açúcar, foram utilizados os dados de expressão do SUCEST, do RNA-seq de folha e de entrenó, cedidos pelo colaborador Dr. Renato Vicentini, e de raiz cedidos pelo Dr. Marcos Buckeridge. Os dados do SUCEST são provenientes de diferentes bibliotecas (VETTORE, 2003), em que cada uma corresponde a um tecido ou condição distinta, de diferentes variedades comerciais de cana-de-açúcar. Já o RNA-seq gerado a partir de amostras de folha utilizou variedades diferentes de cana-de-açúcar, IAC-SP96-3046, SP81-3250, SP80-3280, RB83-5486, RB92-5345, IAC-SP95-3018, que contrastam em algumas características como a capacidade de perfilhamento e resistência a fungos, sendo que todas possuem alto conteúdo de sacarose (CARDOSO-SILVA et al., 2014). O RNA-seq realizado a partir de entrenós (Quinto entrenó) utilizou dois genótipos com conteúdo contrastante de lignina: IAC-SP04627 (maior teor de lignina) e IAC-SP04065 (menor teor de lignina) (VICENTINI et al., 2015). Os dados do RNA-seq de raiz foram obtidos usando diferentes regiões desse tecido (De S1, com ausência total de aerênquima, a S4, onde ele já está formado) (PIOVEZANI, 2017).

A análise de expressão foi feita com a quantidade de sequências expressas de cada biblioteca do SUCEST (número de *reads* dentro de cada sequência/tamanho da biblioteca) e a abundância dos transcritos em cada biblioteca de RNA-seq de raiz (FPKM, na escala log2), de RNA-seq de folha (transcritos por milhão, TPM) e de RNA-seq de entrenó (transcritos por milhão, TPM). Para facilitar a identificação das sequências diferencialmente expressas, os dados dos bancos de folha e entrenó também foram normalizados para escala log2. Esses dados foram utilizados para construção de *heatmaps* no *software* R (R CORE TEAM, 2013), utilizando agrupamento (*clusterização*) hierárquico pelo método euclidiano e distância completa (*complete*).

4.3 COLETA DE AMOSTRAS DOS HÍBRIDOS DE CANA-DE-AÇÚCAR COM CONTEÚDO CONTRASTANTE DOS COMPONENTES DA PAREDE CELULAR

O plantio de 11 híbridos e três cultivares de referência de cana-de-açúcar (MASARIN et al., 2011) foi realizado em outubro de 2013 em campo experimental na

localidade de Lorena-SP. Esses híbridos são provenientes do programa de melhoramento da RIDESA (Rede Interuniversitária para Desenvolvimento do Setor Sucroenergético) associado com a Universidade Federal de Viçosa, com o objetivo de obter plantas com composição ou conteúdo de lignina modificados, e foram cedidos para o Departamento de Biotecnologia (EEL-USP) para estudo. Dois dos híbridos plantados, denominados H89 e H321, foram selecionados para este trabalho por possuírem conteúdo contrastante de ácidos hidroxicinâmicos na parede celular (COSTA et al., 2013; MASARIN et al., 2011). Amostras biológicas de entrenós em fases distintas do desenvolvimento, separados nas regiões medula e córtex, foram coletadas, congeladas imediatamente em nitrogênio líquido e armazenadas a -80°C em *Ultrafreezer* modelo U410 (*New Brunswick Eppendorf*).

4.4 EXTRAÇÃO DE RNA DOS TECIDOS DE CANA-DE-AÇÚCAR, TRATAMENTO COM DNAse E SÍNTESE DE cDNA

Os tecidos do córtex e medula em diferentes fases de desenvolvimento (entrenós 1, 5 e 9, numerados a partir do topo do colmo), bem como os de folha do híbrido H89 (cedidos pela Dra. Tatiane da Franca, EEL-USP), foram retirados do *ultrafreezer* (-80°C) e mergulhados imediatamente em nitrogênio líquido. As amostras foram maceradas em moinho criogênico (*2010 Geno/Grinder*[®], *Spex SamplePrep*[®]). O material macerado foi separado em microtubos - mantendo sempre as espátulas e os microtubos em contato com o nitrogênio líquido - e armazenado em *ultrafreezer*.

A extração de RNA do material macerado foi realizada usando o kit *Concert* TM *Plant RNA Reagent (Invitrogen*TM), utilizando o protocolo do fabricante com otimizações. Adicionou-se 500 μ L do reagente *Concert* frio (4°C) a 300 mg de material (para extração de entrenó e raiz) ou 150 mg (para a extração de folha). Misturou-se em agitador até a solução se tornar homogênea. Os microtubos foram incubados, na posição horizontal, à temperatura ambiente por 5 minutos. Centrifugou-se a solução a 12000g por 5 minutos, à temperatura ambiente. O sobrenadante foi separado em um novo microtubo de 2,0 mL, e o precipitado foi descartado. Foram adicionados 200 μ L de NaCl 5M e agitaram-se os tubos. Adicionou-se 300 μ L de clorofórmio e misturou-se por inversão. Centrifugou-se a 1200g por 10 minutos, a 4°C, para separar as fases aquosa e orgânica. A fase aquosa, na parte superior, foi separada em novo microtubo de 2,0 mL, enquanto a fase orgânica (inferior) foi descartada. Adicionou-se igual volume de isopropanol gelado (4°C) à fase aquosa,

agitou-se por inversão, e os tubos foram incubados à temperatura ambiente por 10 minutos. As amostras foram centrifugadas a 12000g por 10 minutos, a 4°C. Descartou-se o sobrenadante. O precipitado foi lavado adicionando-se 1 mL de etanol 75% e centrifugando a 12000g, a 4°C por 1 minuto. O procedimento de lavagem foi repetido mais duas vezes. O precipitado de RNA foi suspendido em 30 µL de H₂O *Milli-Q* estéril. A quantidade do RNA extraído e a qualidade, indicada pelas relações de absorbância 260nm/280nm e 230nm/280nm, foram determinadas em espectrofotômetro *Biodrop duo (Biochrom)*. A integridade do RNA foi avaliada por meio de eletroforese em gel de agarose 0,8% em TAE 0,5X, imerso em cuba contendo TAE 0,5X, previamente mergulhada em solução detergente por uma hora. Em seguida, 1 µg de cada uma das amostras de RNA foi tratada com *RQ1 RNAse-free DNAse*[®] (*Promega Co*), seguindo as recomendações do fabricante, apenas com a modificação de que as reações foram incubadas à 37°C por 45 minutos. A síntese da primeira fita de cDNA foi conduzida usando a enzima *SuperScript III Reverse Transcriptase*[®] (*Invitrogen*TM), utilizando 500 ng da solução de RNA tratada com DNAse e o iniciador *oligo*dT, seguindo o manual do fabricante.

4.5 ANÁLISE DO PADRÃO DE EXPRESSÃO DOS GENES SELECIONADOS EM DIVERSOS ÓRGÃOS E TECIDOS EM CANA-DE-AÇÚCAR POR RT-qPCR

4.5.1 Desenho de iniciadores específicos para os transcritos de cana-de-açúcar identificados na família BAHD

Para análise da expressão gênica das sequências BAHD identificadas no clado relacionado com a incorporação dos ácidos hidroxicinâmicos na parede celular (grupo A do clado de *Mitchell*), foram desenhados iniciadores utilizando as sequências identificadas nos bancos transcriptômicos de cana-de-açúcar. Foram utilizadas para o desenho as regiões entre o último éxon e a região 3'UTR, identificadas a partir de alinhamentos entre as sequências de cana-de-açúcar e o ortólogo em sorgo. Buscaram-se regiões conservadas entre os transcriptomas analisados, mas sempre que possível, o transcrito de entrenó foi priorizado como alvo, uma vez que a expressão seria analisada principalmente nesse tecido. Foi utilizado o *software primer 3 plus*, obedecendo aos parâmetros: Tamanho do fragmento entre 80 pb e 150 pb, tamanho dos iniciadores entre 20 pb e 22 pb, temperatura de anelamento (Tm) entre 57°C e 63°C (Tabela 2). A especificidade foi checada por meio de alinhamento dos oligonucleotídeos com outros transcritos homólogos da família BAHD.

Ademais, foi feita uma busca na literatura por iniciadores para genes constitutivos de cana-de-açúcar com expressão estável em diferentes tecidos, possibilitando a normalização da expressão gênica. Foram selecionados os iniciadores GAPDH (BOTTCHER et al., 2013; CESARINO et al., 2013; VICENTINI et al., 2015) (Tabela 2). A síntese de todos os oligonucleotídeos foi encomendada para a empresa *EXXTEND*.

4.5.2 Reação RT-qPCR e análise da curva de melting

Inicialmente, foram otimizadas as concentrações de iniciadores e de cDNA, usando um pool de amostras de cDNA dos entrenós 1,5 e 9, e folha jovem do híbrido H89, provenientes de duas plantas diferentes. Para cada reação, foram combinados 2,5 μ L de amostra (diluições 1:25 ou 1:50), os iniciadores (concentrações 200 μ M ou 400 μ M), 5 μ L de master mix *SYBR*TM *GREEN PCR master mix* (*Applied Biosystems*TM, *ThermoFisher Scientific*TM) e água *Milli-Q*, totalizando 20 μ L. O controle negativo foi feito misturando-se apenas o master mix, iniciadores e água. As reações de RT-qPCR foram executadas em triplicata técnica e duplicata biológica em placas de 96 poços no 7500 Fast Real-time PCR system (Applied Biosysitems, ThermoFisher ScientificTM), utilizando o programa 95°C(20s)/95°C(3s)/ 40 ciclos 60°C (30s) e curva de *melting* 95°C (20s)/60°C (1min)/95°C (15s)/60°C (15s). Foram consideradas ótimas as concentrações de iniciadores e cDNA mínimas capazes de amplificar o produto (Tabela 2). A especificidade da reação foi avaliada para cada par de iniciadores por meio da análise da curva de *melting* (Anexo C), buscando-se picos únicos e em caso de presença de picos no controle negativo, que estes fossem em temperaturas distintas.

Os resultados de todas as reações foram analisados no software RQ *Relative Quantification, ThermoFisherCloud (ThermoFisher Scientific*TM). As eficiências médias de cada par de iniciadores em todas as reações foram calculadas usando o software *LinRegPCR* (RAMAKERS et al., 2003; RUIJTER et al., 2009). As eficiências obtidas por esse método foram incorporadas na análise pelo software RQ *Relative Quantification*, o qual calcula o valor de Ct equivalente, que corresponde a uma estimativa do Ct caso a eficiência fosse 100% (PFAFFL, 2001). Os valores de Δ Ct (Ct gene de interesse – Ct GAPDH), calculados a partir dos Ct corrigidos pela eficiência, foram utilizados para a análise de expressão (PFAFFL, 2001).

	Diluição das	Concentração dos	Eficiência média dos
Nome	amostras	iniciadores	iniciadores
ScAt1	1:50	400	1.88
ScAt2	1:50	200	1.85
ScAt4*	1:50	400	1.92
ScAt5	1:25	400	1.83
ScAt6*	1:50	400	1.92
ScAt7*	1:50	400	1.90
ScAt8	1:50	200	1.79
ScAt9*	1:50	400	1.92
ScAt10	1:50	400	1.86
ScGAPDH	1:25/1:50	200	1.91

Tabela 2– Padronização dos iniciadores dos unigenes de cana-de-açúcar da família BAHD, grupo A do clado de *Mitchell*, para uso em RT-qPCR

Nota: Nome dos iniciadores de acordo com a nomenclatura dos genes BAHD do clado de Mitchell de Bartley *et al.* 2013.

*Iniciadores que foram redesenhados para atingir a padronização, pelo fato de o primeiro iniciador encomendado não passar no teste da curva de *Melting*.

Fonte: Autoria própria.

4.6 CLONAGEM DE GENES DE CANA-DE-AÇÚCAR SELECIONADOS COM POTENCIAL BIOTECNOLÓGICO

4.6.1 Desenho de iniciadores para amplificação do cDNA

Após triagem e identificação dos genes *ScAt10* e *ScAt4* nos bancos genômicos e transcriptômicos, foram desenhados iniciadores, usando o *software primer 3 plus* (*http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi*), para amplificação do cDNA das sequências completas identificadas. Também foram desenhados iniciadores em regiões dessas sequências para a amplificação de fragmentos que sejam capazes de gerar pequenos RNAs de interferência (*siRNA*) para silenciamento desses genes. Para a seleção dessas regiões, foi utilizado o software *pssRNAit* (*http://plantgrn.noble.org/pssRNAit/*). Este programa busca candidatos à *siRNA* de 21 nucleotídeos em uma sequência alvo, contra um banco de dados de *RNAm* da espécie selecionada, calculando os *siRNA* mais específicos, eficientes e não-tóxicos. A busca foi feita utilizando as sequências *ScAt10* e *ScAt4* contra o banco de milho (*Zea mays*), uma vez que as construções deverão ser capazes de silenciar os genes desta espécie modelo, escolhida para transformação e análise fenotípica de linhagens transgênicas. Das regiões com candidatos a *siRNA* identificadas

pelo programa, foram selecionadas as que eram específicas das sequências *ScAt10* e *ScAt4*, identificadas por meio do alinhamento com as sequências homólogas, conferindo mais especificidade ao silenciamento. Em seguida, essas regiões foram novamente utilizadas para busca no programa, contra os bancos de *Zea mays* e cana-de-açúcar, selecionando a que possuía maior quantidade de siRNA comuns para as duas espécies, o que permite futuro uso também para silenciamento de linhagens de cana-de-açúcar. Ademais, seguindo o manual da clonagem *Gateway* (*Invitrogen*TM), foram desenhados iniciadores consistindo no desenho inicial (para amplificação dos fragmentos de *RNAi*) mais as pontas recombinantes *attb*, necessárias para a reação de recombinação BP (Tabela 3). A síntese dos oligonucleotídeos foi encomendada para a empresa *EXXTEND*.

Nome Alvo/Uso			Tm
OxAt4 5'end	Amplificação do cDNA de ScAt4	1299	55.1
OxAt4 3'end			55.6
attB1+At4	Inserção das extremidades <i>attB</i> no cDNA de <i>ScAt4</i>	1360	67.8
attB2+At4			67.9
OxAt10 5'end	Amplificação do cDNA de ScAt10	1362	61.4
OxAt10 3'end			62.0
attB1+At10	Inserção das extremidades attB no cDNA de ScAt10	1423	67.8
attB2+At10			68.2
At10RNAi F	Amplificação do fragmento de <i>ScAt10</i> para silenciamento por RNAi	163	61.6
At10RNAi R	-		61.5
attB1+ At10RNAi	Inserção das extremidades <i>attB</i> no fragmento de <i>At10RNAi</i>	224	69.7
attB2+At10RNAi			69.7
At4RNAi F	Amplificação do fragmento de <i>ScAt4</i> para silenciamento por RNAi	162	56.8
At4RNAi R			56.9
attB1+At4RNAi	Inserção das extremidades <i>attB</i> no fragmento de <i>At4RNAi</i>	223	69.0
attB2+At4RNAi			70.4

Tabela 3– Sequência de iniciadores para clonagem de *Ox:ScAt10*, *Ox:ScAt4*, *At10:RNAi*, *At4:RNAi* e temperatura de anelamento predita pelo *software primer3plus*

Fonte: Autoria própria.

4.6.2 Síntese do cDNA e amplificação por reação de PCR para as construções de silenciamento

Para obtenção dos produtos de PCR para a clonagem gênica das construções de silenciamento, inicialmente o cDNA foi sintetizado a partir do RNA extraído dos híbridos de cana-de-açúcar. Para diminuir a chance de obter produtos inespecíficos durante a reação de PCR, bem como aumentar a concentração inicial de cDNA especificamente de *ScAt10* e *ScAt4*, o cDNA foi sintetizado a partir apenas do *RNAm* dessas sequências, utilizando iniciadores específicos dos genes, desenhados na região 3'UTR (Iniciadores reversos, Tabela 2).

O RNA foi extraído de tecidos que expressam esses genes (Folha para *ScAt10*, e entrenó para *ScAt4*) como descrito em 4.4. Em seguida, 1 µg do RNA foi tratado com DNAse e a síntese cDNA foi conduzida como descrito no item 4.4, apenas com a substituição do *oligo*dT por 1µL do iniciador reverso Sc*At10*(100µM) (Tabela 4), na reação com amostra de RNA proveniente de folha, e 1µL do iniciador reverso *ScAt4* (100µM) (Tabela 4) para a reação com RNA proveniente do entrenó.

As reações de PCR foram conduzidas em uma solução de 25 μ L contendo 1 μ L do cDNA específico, utilizando os iniciadores para clonagem (*At10RNAi* e *At4RNAi*, Tabela 3) e as enzimas de alta fidelidade *Phusion*® *High fidelity DNA polymerase* (*New England Biolabs*TM) ou *Q5*TM *High fidelity DNA polymerase* (*New England Biolabs*TM), seguindo as recomendações do fabricante. Para as reações com a polimerase *Q5*, é recomendado o cálculo da temperatura de anelamento (Tm) pelo programa *Neb Tm calculator* (http://tmcalculator.neb.com/#!/), uma vez que essa enzima de um modo geral requer Tm mais elevada. As temperaturas preditas pelos *softwares primer3plus* e *Neb Tm calculator*, bem como o programa de PCR utilizado foram detalhados na Tabela 4. Nos casos em que mais de uma temperatura de anelamento foi testada, a que proporcionou melhor resultado foi marcada em negrito. Os tamanhos dos produtos de PCR foram verificados por eletroforese em gel de agarose e as imagens capturadas em fotodocumentador transiluminador *LPIX touch (Loccus* biotecnologia).

Os produtos de PCR foram purificados usando kit *GeneJET PCR Purification Kit* (*Thermofischer*TM). No caso de visualização de bandas adicionais na eletroforese, menores que o tamanho esperado (ex. dímeros de iniciadores em excesso), os produtos foram sacados do gel de agarose e então purificados.

Nome	Enzima	Tm primer 3 plus	Tm Neb calculator	Programa de PCR (Tm ótima em negrito)
At4RNAi	Phusion	56.8	67	98°C (30s), 35 ciclos 98°C (10s)/ 57 °C(30s)/72°C (7s)
At4KNA1		56.9	63	72° C(10 min)
At10RNAi	Q5	61.6	72	98°C (30s), 35 ciclos 98°C (10s)/ 68 °C(30s)/72°C (7s)
At10RNAi		61.5	72	72° C(2 min)

Tabela 4– Condições de amplificação dos produtos para clonagem por PCR, comparando as temperaturas de anelamento utilizadas com a temperatura predita pelos *softwares primer 3 plus e Neb calculator*.

Fonte: Autoria própria.

Os produtos de PCR obtidos para cada uma das sequências (*At10RNAi*, *At4RNAi*) foram submetidos à uma nova reação de PCR, para inserção das pontas recombinantes *attB1* e *attB2* necessárias para o sistema de clonagem *Gateway*. A nova reação de PCR foi conduzida com a enzima $Q5^{TM}$ (*New England Biolabs*TM), utilizando 1 µL do produto de PCR purificado, seguindo as recomendações do fabricante, nas condições e ciclagem descritas na Tabela 5. Os produtos de PCR foram purificados como descrito acima.

Tabela 5– Condições de amplificação dos produtos attB PCR, comparando as temperaturas de anelamento utilizadas com a temperatura predita pelos *softwares primer 3 plus e Neb calculator*.

Nome	Enzima	Tm primer 3 plus	Tm Neb calculator	Programa de PCR (Tm ótima em negrito)
At4RNAi	Dhusion	69	72	98°C (30s), 35 ciclos 98°C (10s)/ gradiente 56 °C, 60°C,
At4RNAi	Phusion	70.4	72	65°C (30s)/72°C (7s), 72° C(10 min)
At10RNAi		69.7	72	98°C (30s), 15 ciclos 98°C (10s)/ 61 °C(30s)/72°C (7s),
At10RNAi	Q5	69.7	72	20 ciclos 98°C (10s)/ 61°C(30s)/72°C (7s), 72° C(2 min)

Fonte: Autoria própria.

4.6.3 Síntese das sequências codificantes para as construções de super-expressão

Para a construção Ox:ScAt10, houve dificuldade na amplificação do cDNA completo do gene por meio de reações de PCR, dado que muitos produtos inespecíficos também eram amplificados. Dessa forma, a síntese foi encomendada para a empresa *Fastbio*, que sintetizou a sequência codificante completa fornecida juntamente com as pontas *attB* adjacentes (1424 pb) e inseriu em vetor puc57 (totalizando 4134pb). A sequência *ScAt10* com as pontas *attB* para recombinação foi amplificada a partir do vetor sintetizado usando os iniciadores *attBScAt10* (Tabela 3), uma vez que a reação BP do sistema de clonagem *gateway* (*Invitrogen* TM) exige produtos linearizados (segundo o manual do fabricante). A reação de PCR foi conduzida em 25 µL de solução contendo 1 µL do vetor sintetizado diluído (1:4), com a enzima $Q5^{TM}$ (*New England Biolabs*TM) seguindo as recomendações do fabricante. A temperatura de anelamento foi de 68°C, e o programa de PCR foi 98°C (30 seg), 35 ciclos de 98°C 10 seg/ 68°C 30 seg/ 72°C 50 seg, 72°C 10 min. A banda obtida de tamanho esperado foi sacada e purificada usando o kit *GeneJET PCR Purification Kit (Thermofischer*TM).

Para a construção Ox:ScAt4, foi possível a amplificação do cDNA completo por meio de reações de PCR utilizando os iniciadores para clonagem Ox:At4 5 'end (Tabela 3), a enzima $Q5^{\text{TM}}$ (New England BiolabsTM) seguindo o protocolo do fabricante, temperatura de anelamento 66°C, e programa de PCR 98°C (30 seg), 35 ciclos de 98°C 10 seg/ 66°C 30 seg/ 72°C 50 seg, 72°C 10 min . Entretanto, a segunda amplificação para inserção das pontas recombinantes *attb* gerou diversas bandas inespecíficas, de forma que não foi possível a obtenção do produto de PCR *attb* para a clonagem *Gateway*®(*Invitrogen*TM). Assim, a síntese foi encomendada para a empresa *Fastbio*. De forma similar à síntese de *ScAt10*, a sequência codificante de *ScAt4* com as pontas recombinantes *attb* foi sintetizada (1351pb) e inserida em vetor puc57 (totalizando 4079 pb). O vetor foi enviado pela empresa para o *Dairy Forage Research Center* (DFRC, USDA-ARS), Madison, WI, EUA, onde as posteriores etapas de clonagem foram realizadas. A sequência linear para a reação BP foi obtida por digestão do vetor *puc57:OxScAt4* com a enzima de restrição HindIII (*New England Biolabs*TM), seguindo o protocolo do fabricante para um volume de reação de 25µL, seguida da inativação da enzima por calor (80°C, 20 minutos).

4.6.4 Preparo de células competentes de Escherichia coli

Para a clonagem por eletroporação, linhagens One Shot[®] TOP10 Chemically Competent E. coli (Invitrogen[™]) e One Shot[®] ccdB Survival[™] 2 T1R Competent Cells (InvitrogenTM) foram preparadas para se tornarem eletrocompetentes. Inicialmente, as bactérias foram inoculadas em 3 mL de meio LB líquido contendo 50 µg/mL de antibiótico estreptomicina. Incubou-se sob agitação (225 rpm) a 37°C por cerca de 16 horas. Inoculouse 500 µl do pré-inóculo em 200 mL de LB líquido sem antibiótico em um Erlenmeyer de 500 mL. Incubou-se novamente sob agitação (225 rpm) a 37°C. Após aproximadamente 3 horas de incubação, a OD 600 nm começou a ser medida em espectofotômetro Biodrop duo (Biochrom) a cada quinze minutos, até que atingisse um valor entre 0,5-0,6. Quando a cultura atingiu a OD correta, as células foram incubadas no gelo por 30 minutos. O meio foi distribuído em tubos de centrífuga e centrifugou-se por 15 minutos, a 5000 g a 4°C. Em seguida, removeu-se o sobrenadante e as células foram suspendidas em 25 mL de água Milli-Q estéril, apenas agitando o tubo levemente no gelo. Centrifugou-se novamente por 15 minutos a 5000 g à 4°C. Descartou-se novamente o sobrenadante e as células foram suspensas em 10 mL de água Milli-Q estéril. Os volumes de todos os tubos foram misturados, obtendo dois tubos com 30 mL cada. Centrifugou-se novamente por 15 minutos a 5000 g a 4°C. O sobrenadante foi descartado e as células foram suspendidas em 1,9 mL de Glicerol 10% estéril. Novamente, centrifugou-se por 15' a 5000 g a 4°C. Todo o sobrenadante foi removido, retirando o excesso com auxílio de uma pipeta. O precipitado de cada um dos tubos foi suspenso em 784 µL de glicerol 10% estéril. Foram feitas alíquotas de 90 µL em tubos de 1,5 mL, que foram armazenadas em *ultrafreezer* (-80°C).

Para a clonagem por choque térmico, foram utilizadas linhagens de *E. coli* DH5α cedidas pelo colaborador Dr. David Rancour (*Dairy Forage Research Center*, USDA), preparadas para se tornarem termocompetentes utilizando o protocolo descrito (INOUE; NOJIMA; OKAYAMA, 1990).

4.6.5 Reação de recombinação Gateway e inserção em Escherichia coli

Após a obtenção dos produtos de PCR com as pontas recombinantes attB1 e attB2 ou do vetor linearizado, as sequências foram inseridas no vetor pDONR221 (*Invitrogen*TM) por meio da reação *BP clonase* (*Invitrogen*TM), seguindo as instruções do fabricante,

apenas com o detalhe de que a reação foi incubada à 25°C por 4 horas antes de ser terminada pela adição de proteinase K.

Para as construções ScAt4:RNAi, ScAt10:RNAi e Ox:ScAt10 os vetores de entrada foram inseridos nas linhagens E coli TOP10 eletrocompetentes, por eletroporação, usando o Gene Pulser XcellTM electroporation System (BIO RAD). 40µL de célula competente foram misturados com 1 µL do produto da reação BP e incubados por 1 minuto. Em seguida, a solução foi transferida para cubetas de 0.1mm ou 0.2 mm (BIO RAD) e submetidas à eletroporação, seguindo o protocolo do fabricante do eletroporador para transformação em E coli. As células transformadas foram inoculadas em meio LB sólido contendo 50 µg/mL de antibiótico Kanamicina. As placas foram incubadas a 37°C por cerca de 16 horas. Em seguida, as colônias foram individualmente inoculadas em meio LB líquido contendo também 50 µg/mL de antibiótico Kanamicina, e incubadas à 37°C por cerca de 16 horas sob agitação de 200 rpm. Os vetores de entrada foram extraídos das colônias selecionadas utilizando o kit GeneJET Plasmid Miniprep Kit (ThermofischerTM) seguindo as instruções do fabricante. Todos os vetores de entrada foram confirmados por reação com enzima de restrição KpnI, PvuI e PvuII (New England BiolabsTM) e reação de PCR com os iniciadores específicos das sequências (Tabela 3). Os vetores foram enviados para sequenciamento pela empresa *Eurofins*, confirmando a inserção das sequências.

Os vetores de entrada foram recombinados com os vetores de destino de superexpressão *pzp221bOX* e de silenciamento *pzp221bRNAi* (MARITA et al., 2014; RANCOUR et al., 2015) cedidos pelos colaboradores Dr. David Rancour e Dr. Ronald Hatfield (U.S Dairy Forage Research Center, USDA) para uso exclusivo no presente trabalho. Tais vetores possuem o sistema de recombinação Gateway e bordas T-DNA para posterior inserção das sequências em linhagens de milho por Agrobacterium tumefaciens. No caso do vetor de RNAi, o sistema de recombinação faz com que a sequência seja inserida de forma repetida e invertida, para gerar o RNA dupla fita que desencadeia o silenciamento dos genes na planta. Inicialmente, para obter maior quantidade dos vetores de destino, estes foram propagados nas células eletrocompetentes One Shot® ccdB Survival[™] 2 T1R Competent Cells, resistentes ao gene ccdB, em placas de petri contendo meio LB sólido contendo o antibiótico Espectinomicina na concentração de 25 µg/mL, como descrito acima. Os plasmídeos foram purificados das colônias selecionadas pelo GeneJET Plasmid Miniprep Kit (ThermofischerTM) e confirmados por reações com enzimas de restrição KpnI e BamHI (New England BiolabsTM), seguindo o protocolo do fabricante para 25μ L de reação. Em seguida, a reação LR clonase (InvitrogenTM) foi conduzida entre os vetores de entrada e os de destino, seguindo o manual do fabricante, com a única observação de que a reação foi incubada à 25°C por 4 horas antes de ser terminada pela adição de proteinase K. Os vetores finais foram inseridos em bactérias *E.coli* TOP 10 por eletroporação, as colônias foram selecionadas em meio contendo Espectinomicina 25µg/mL e os plasmídeos foram purificados como descrito acima.

Todos os vetores foram confirmados por digestão com enzima de restrição *KpnI* e *BamHI* (*New England Biolabs*TM), seguindo o protocolo do fabricante e por sequenciamento pela empresa *Eurofins*.

Para a construção Ox:ScAt4, após a reação BP, os vetores de entrada foram inseridos nas células de E.coli DH5a por choque térmico. Em um microtubo estéril, em contato com gelo, 25 µL de célula competente foram misturados com 1 µL da reação BP. Incubou-se o tubo no gelo por 30 minutos. O tubo foi então incubado em banho de água quente pré-aquecido a 42°C, durante exatamente 1 minuto. Transferiu-se o microtubo novamente para o gelo por 2 minutos. Em seguida, adicionaram-se 125 µL de meio SOB (Super Optimal Broth, Sigma AldrichTM) e incubou-se o tubo à 37°C sobe agitação (200 rpm) por 1 hora. As células foram inoculadas em meio LB sólido contendo 50 µg/mL de antibiótico Kanamicina. A placa de petri foi incubada à 37°C por cerca de 16 horas. As colônias foram individualmente inoculadas em meio LB líquido, e novamente propagadas a 37°C sobe agitação por 16 horas. O vetor recombinante foi purificado utilizando o Kit GeneJET Plasmid Miniprep Kit (ThermofischerTM) e confirmado por dupla digestão com as enzimas KpnI e PvuII (New England BiolabsTM), seguindo o protocolo do fabricante para reação de 10 µL. Os vetores de entrada e de destino foram recombinados por meio da reação LR clonase como descrito acima. O vetor final foi transformado em células DH5a, as colônias foram selecionadas em meio contendo Espectinomicina (25 µg/mL) e os plasmídeos foram purificados como descrito para o vetor de entrada. O vetor pzp221b:Ox:ScAt4 foi confirmado por reação com enzimas de restrição PvuII e BamHI e por sequenciamento pelo Centro de Biotecnologia UW Madison.

4.7 TRANSFORMAÇÃO DAS LINHAGENS DE MILHO, CULTIVO DAS PLANTAS TRANSGÊNICAS E GENOTIPAGEM POR PCR

4.7.1 Transformação das linhagens de milho pela Plant Transformation Facility, Iowa State University

Todos os vetores confirmados *pzp221b:ScAt10:RNAi*, *pzp221b:ScAt4:RNAi*, *pzp221b:Ox:ScAt10* e *pzp221b:Ox:ScAt4* foram enviados para a *Plant Transformation Facility* (IOWA STATE UNIVERSITY, 2019), para transformação de linhagens de milho (genótipo *HiII*) mediada por *Agrobacterium tumefaciens*. Esse procedimento foi financiado pelo projeto FAPESP temático (2014/06923-6). Na *facility*, os calos embriogênicos transformados foram regenerados e selecionados em placa de petri com meio de regeneração contendo o herbicida *bialaphos* (glufosinato de amônio). Cerca de 10 eventos independentes foram obtidos para cada uma das quatro construções.

4.7.2 Cultivo das plântulas T₀, cruzamento e geração de sementes T₁

As plântulas em placas de petri contendo o cassete de silenciamento de *ScAt10 e ScAt4* (denominadas de eventos A872 e A873, respectivamente), bem como superexpressão de *ScAt10 e ScAt4* (denominadas de eventos A871 e A885) foram enviadas pela *facility* (Figura 6, A) para o *U.S Dairy Forage Research Center*, *USDA*. Após o recebimento, elas foram transferidas para pequenos vasos (Figura 6, B) contendo uma mistura de 20 kg do substrato *PRO-MIX BX mycorrhizae general purpose growing medium (PROMIX[®])* para cada 60 L do solo *Happy Frog Potting Soil* (*Smart Naturals*TM). Nessa etapa, o regime de iluminação foi de 24 horas de luz. Após um mês, quando as plântulas atingiram um tamanho adequado, elas foram transferidas para vasos de 11,5 L (Figura 6, C e D), contendo a mesma mistura de substrato. O regime de iluminação foi de 16 horas de luz. As plantas foram fertilizadas mensalmente com adubo NPK 18-2-18 (*Water soluble fertilizer 18-2-18 boosted base FeED, Jack's professional*[®]) e regadas de acordo com a necessidade.



Figura 6 – Linhagens transgênicas de milho em diferentes estágios de crescimento

Legenda: A - Plântulas recebidas em placa de petri. B - Plântulas após serem transferidas para vasos de cerca de 5cm, em câmara de crescimento (24 horas de luz). C - Plântulas crescidas por um período de um mês, cultivadas em casa de vegetação com fotoperíodo de (16 horas de luz, 8 de escuro), em vasos de 11,5 L. D - Plantas adultas em fase reprodutiva. Fonte: Autoria própria.

Quando as plantas T_0 atingiram a fase reprodutiva (Após cerca de 3 meses), foi feito o cruzamento entre as plantas de cada construção, obtendo a geração T_1 . Foram utilizados como doadores de pólen milhos selvagens genótipo B73, crescidos na mesma condição descrita para os transgênicos. Cortaram-se os pendões dos milhos B73 e foi feita a polinização manual dos estigmas dos cultivares transgênicos. Os estigmas foram cobertos com uma embalagem de papel para prevenir polinização indesejada pelo vento. Para as construções A871 (*Ox:ScAt10*) e alguns eventos A885 (*Ox:ScAt4*), os milhos transgênicos amadureceram um pouco antes dos B73, de forma que atingiram a fase reprodutiva antes dos selvagens iniciarem a produção de pólen. Nestes casos, eventos independentes da mesma construção foram cruzados entre si, ou seja, utilizando uma planta de um evento como doadora de pólen e outra planta de outro evento como receptora.

4.7.3 Extração de DNA e genotipagem por PCR

Após cerca de um mês de crescimento das plantas, parte da folha foi coletada para extração de DNA e genotipagem por PCR, a fim de verificar a presença do inserto nos transgênicos. Foram coletadas folhas tanto dos eventos transgênicos como selvagem (B73, controle).

O DNA foi extraído utilizando o método CTAB. As folhas foram coletadas em microtubos de 1.5mL e armazenados no gelo. Foram adicionados 300 μ L de uma solução de extração contendo 2% m/v de brometo de cetiltrimetilamônio (CTAB), 0.1 M Tris-HCl pH 8.0, 1%(v/v) 2-mercaptoetanol e 1% (m/v) bissulfito de sódio. Os tecidos foliares foram macerados com um pilão de tamanho adequado para o microtubo. Adicionaram-se mais 300 μ L da solução de extração e os tubos foram incubados a 55°C por 30 minutos. As amostras foram centrifugadas rapidamente, e os sobrenadantes foram recolhidos em microtubos de 1.5mL. Em cada tubo, adicionou-se 0.8mL de uma solução 24:1 (v/v) clorofórmio/octanol e misturou-se por inversão. Centrifugaram-se os tubos a 14000 g por 3 minutos. Para cada amostra, 600 μ L da fase superior foram transferidos para outro microtubo de 1.5mL. Adicionaram-se 600 μ L de isopropanol e misturou-se por inversão. Os tubos foram incubados à temperatura ambiente por 15 minutos. As amostras foram centrifugadas a 14000g por 3 minutos. Os sobrenadantes foram removidos, os *pellets* foram lavados três vezes com 1 mL de etanol 70% e deixados secar ao ar. Os *pellets* foram então suspendidos em 50 μ L de água *Milli-Q* estéril.

Para a genotipagem por PCR, iniciadores específicos foram desenhados para cada construção (Tabela 6). Para confirmar as construções de super-expressão, foram desenhados iniciadores específicos para cana-de-açúcar, na região dos dois éxons da sequência codificante clonada, garantindo que os iniciadores não amplificariam os genes endógenos de milho *ZmAt10* e *ZmAt4* (Figura 7). A diferenciação entre as sequências codificantes de cana-de-açúcar e as genômicas endógenas de milho foi garantida tanto pela diferença na composição de bases como pelo tamanho, já que as sequências genômicas incluiriam o íntron. Para as construções de silenciamento, utilizou-se um trecho do cassete inserido como alvo. Foram utilizados os iniciadores reversos para os fragmentos de *RNAi ScAt10* ou *ScAt4* (Tabela 6), como iniciadores reversos para a sequência anti-senso. Como iniciador *forward*, foi desenhado um oligonucleotídeo possuindo como alvo o trecho do vetor *pzp221bRNAi* após a região NOS terminal, localizada após o fragmento no sentido

anti-senso (Tabela 6, Figura 8). Assim, garantiu-se especificidade do conjunto de iniciadores, de forma a não amplificar genes endógenos de milho.

Tabela 6 – Iniciadores específicos para genotipagem por PCR das linhagens A871 (*Ox:ScAt10*), A872 (*ScAt10:RNAi*), A873 (*ScAt4:RNAi*) e A885(*Ox:ScAt4*)

Nome	Sequência	Alvo	Produto	Tm(°C)	GC
Or:ScAtlOgan F	TGACCACCAAGCAGAA	A871	293	59.9	55
Ox.Schilogen I	GTCC				
Ox:ScAtlOgen R	CGTGACCTGTATCATGG	CGTGACCTGTATCATGG A871		59.9	55
Ox. Serii Togen K	CGA				
Ox:ScAt4gen F	GAGGAGGAGGATGGGG	A885	420	62.3	65
Ox.Selfi Igen I	AGAC				
Ox:ScAt4gen R	CAGGTCCGTCGTGAAG	A885		59.7	55
embern rgen n	TACA				
ScAt10:RNAi F	CTCCGACGACGTCGCG	A872	574	61.6	65
	AAGA				
ScAt4·RNAi F	GTCACCAAACCCCGCA	A873	574	61.1	56,8
	TC				
RH3	GAGCGGATAACAATTT	A872/		54.3	
	CACACAG	A873			

Fonte: Autoria própria

Figura 7– Alinhamento de regiões das sequências codificantes clonadas de cana-de-açúcar (*ScAt10* e *ScAt4*) com as sequências genômicas dos ortólogos em milho (*ZmAt10* e *ZmAt4*), mostrando a posição e especificidade dos iniciadores para genotipagem das linhagens de super-expressão



Legenda: Alinhamento pelo algoritmo *Clustal W*, no software *MegAlignPro*[®]. A – Região de anelamento do iniciador Ox:ScAt10gen F. B- Região de anelamento do iniciador Ox:ScAt10gen R. C- Região de anelamento do iniciador Ox:ScAt4gen F. D- Região de anelamento do inciador *Ox:ScAt4gen R*. Fonte: Autoria própria.



Figura 8- Região de anelamento dos iniciadores para genotipagem das linhagens de silenciamento

Legenda: A - Região proveniente do vetor *pzp221b:RNAi:At10*. B - Região proveniente do vetor *pzp221b:RNAi:At4*. Construção de silenciamento entre as bordas T-DNA. Mapa completo do vetor em anexo. Fonte: Autoria própria.

Todos os iniciadores foram desenhados usando o *software Primer3Plus* e tiveram a síntese encomendada para a empresa IDT.

As reações de PCR foram realizadas utilizando o mix *EconoTaq*[®] *PLUS GREEN* 2X Master Mix (Lucigen[®]), seguindo as instruções do fabricante para 10 µL de reação. Os ciclos da reação de PCR foram: Para A871 94°C 2 min, (94°C 30s, 60°C 30s, 72°C 30s)30X, 72°C 10 min; para A872, A873 e A885 94°C 2 min, (94°C 30s, 60°C 30s, 72°C 40s)30X, 72°C 10 min. Os produtos da reação de PCR foram visualizados por eletroforese em gel de agarose 3% em tampão TBE 1 X, corado com brometo de etídeo. As imagens foram capturadas usando o transiluminador.

4.8 ANÁLISE DO PADRÃO DE EXPRESSÃO DOS GENES *AT10* E *AT4* NAS LINHAGENS DE MILHO TRANSGÊNICAS POR RT-qPCR

4.8.1 Extração de RNA dos tecidos foliares, tratamento com DNAse e síntese de cDNA

Para verificar se de fato ocorreu o efeito de super-expressão ou silenciamento nas linhagens de milho transgênicos, foi feita uma análise da expressão gênica por RT-qPCR. A fim de realizar essa análise em tecidos aproximadamente no mesmo estágio de maturação, uma vez que os níveis de expressão podem mudar ao longo do desenvolvimento, foram coletadas para todos os eventos transgênicos e milho selvagem B73 a primeira folha logo após o desenvolvimento da orelha (órgão reprodutor feminino). As folhas foram colhidas, imediatamente submersas em nitrogênio líquido e armazenadas em ultrafreezer -80°C até o seu uso.

As amostras foram maceradas em moinho criogênico e armazenadas em microtubos de 1.5 mL. Foi feita a extração do RNA dos tecidos macerados usando o kit *Spectrum*TM *Plant Total RNA kit (Sigma-Aldrich*[®]), seguindo as instruções do protocolo A do fabricante. A quantidade do RNA extraído e a qualidade, indicada pelas relações de absorbância 260nm/280nm e 230nm/280nm, foram determinadas em espectrofotômetro. A integridade do RNA foi avaliada por meio de eletroforese em gel de agarose 0,8% em TBE 1 X, imerso em cuba contendo TBE 1X, previamente mergulhada em solução detergente por uma hora. Em seguida, 1 µg de cada uma das amostras de RNA foi tratada com *RQ1 RNAse-free DNAse*[®] (*Promega Co*), seguindo as recomendações do fabricante, apenas com a modificação de que as reações foram incubadas à 37°C por 45 minutos. A síntese da primeira fita de cDNA foi conduzida usando a enzima *GoScript*TM *Reverse Transcription System*[®] (*Promega Co*), utilizando 400 ng da solução de RNA tratada com DNAse e o iniciador *oligo*dT, seguindo o manual do fabricante.

4.8.2 Desenho de iniciadores específicos para as sequências ScAt10, ScAt4, ZmAt10 e ZmAt4

Foram desenhados iniciadores adequados para as reações de PCR em tempo real específicos para as sequências de cana, que não amplificassem os endógenos de milho, a fim de verificar os níveis de expressão dos genes *ScAt10* e *ScAt4* nos eventos A871 e A885, respectivamente (Tabela 7). Para isso, foi utilizada a região entre o último éxon e o trecho entre a extremidade recombinada da clonagem gateway (*attb*) e o NOS terminal (Figura 9). Ademais, com o objetivo de verificar os níveis de expressão dos genes endógenos de milho *ZmAt10* e *ZmAt4*, tanto nas construções de super-expressão (A871 e A885, respectivamente) como nas de silenciamento (A872 e A873), foram desenhados iniciadores específicos nas regiões do último éxon e 3'UTR (Tabela 7). Todos os iniciadores foram desenhados usando o software Primer3Plus, respeitando os parâmetros: Tamanho do fragmento entre 80 pb e 150 pb, tamanho dos iniciadores entre 20 pb e 22 pb, temperatura de anelamento (Tm) entre 57°C e 63°C. A síntese dos oligonucleotídeos foi encomendada para a empresa EXXTEND.

Foi feita uma busca na literatura por iniciadores para genes constitutivos estáveis em tecidos de milho, que pudessem ser utilizados para a normalização dos níveis de expressão. Foram selecionados os genes *ZmLEU*, *ZmCUL* e *ZmMEP* (MANOLI et al., 2012) (Tabela 7).

Tabela 7 – Iniciadores específicos para análise do nível de expressão por RT-qPCR das linhagens A871 (*Ox:ScAt10*), A872 (*ScAt10:RNAi*), A873 (*ScAt4:RNAi*) e A885(*Ox:ScAt4*)

Nome	Sequência	Uso	Produto	Tm(°C)	GC
RTq ScAt4E3 F	CCGAGTTCCAGGAGCA	A885	118	60.2	60
	GATG				
RTq attp-rec	TCGAGCTCGGATCCACT			59.8	55
	AGT				
RTq ScAt10 ex3'.2	AGCTGCGTCACCAAGG	A871	128	60.1	61.1
	AG			(1.0	(2, 2)
RIq rec.2	CIAGIAACGGCCGCCA			61.8	63.2
DT 7m A+A 2 E		1005/1	05	60 7	61.1
<i>ΚΙ ΖΜΑΙ</i> 4.2 Γ	ACCATOTOCOTCCAOO	873	95	00.7	01.1
RT 7mAt4 ? R	GATGGGATGATCAGGA	075		59.8	50
RI Znatti.2 R	AGGA			57.0	50
RT ZmAt10.2 F	AGCTGCGTCACCAAGG	A871/A	96	60.1	61.1
	AG	872			
RT ZmAt10.2 R	CAGGGTTCCATCATCTC			59.8	47.6
	CATA				
RTq_LEU_F	CAGGGAAGGTTGCCTC	A885/A	110	60.2	55
	AGTA	871/A88			
		73		(0)	4.5
RIQ_LEU_R				60	45
	IIG				
RTa MEP E	CCATCTGTCTGGGTCAG	4871	120	50.0	55
KIY_MLI_I	GAT	A071	129	59.9	55
RTa MEP R	TTTGATGCTCCAGGCTT			60.2	50
	ACC			00.2	20
RTq_CUL_F	TACCTTGCCTGATTGGT	A885/A	125	60.4	
	GGT	873			
RTq_CUL_R	ACGTGCTGCAAATCAT			60.1	
	GGTA				

Fonte: Autoria própria.

Figura 9– Posição dos iniciadores para análise dos níveis de expressão de *ScAt10* e *ScAt4* por RT-qPCR



Legenda: A - Região proveniente do vetor *pzp221b:Ox:ScAt10*. B - Região proveniente do vetor *pzp221b:RNAi:ScAt4*. Construção de silenciamento entre as bordas T-DNA. Mapa completo do vetor em anexo. Fonte: Autoria própria.

4.8.3 Reação RT-qPCR e análise da curva de melting

Inicialmente, foram otimizadas as concentrações de iniciadores e de cDNA, usando amostras de cDNA da folha selvagem (B73), como descrito na seção 4.5.2. Foram consideradas ótimas as concentrações de iniciadores e cDNA mínimas capazes de amplificar o produto (Tabela 8). A especificidade da reação foi avaliada para cada par de iniciadores por meio da análise da *curva de melting* (Anexo D). As eficiências dos iniciadores foram calculadas como descrito na seção 4.5.2 (Tabela 8). A expressão relativa foi calculada pelo método do $\Delta\Delta$ Ct com a incorporação das eficiências e normalização pela média geométrica de múltiplos genes de referência (PFAFFL, 2001; VANDESOMPELE et al., 2002), usando o software *Relative Quantification* (*Thermo Fisher Cloud*TM).

Nome	Diluição (RNAi)	Diluição (Ox)	Concentração dos iniciadores	Eficiência
ScOxAt4	-	1:50	200	1.87
ScOxAt10*	-	1:50	200	1.68
RTZmAt4*	1:25	1:50	200	1.68
RTZmAt10*	1:25	1:25	400	1.90
RTqZmLEU	1.25	1:25/1:50	200	1.90
RTq ZmMEP	-	1:50	200	1.74
RTq ZmCUL	1:25	-	200	1.80

Tabela 8 – Padronização dos iniciadores para análise de expressão dos genes ScAt10, ScAt4, ZmAt10 e ZmAt4 nas linhagens transgênicas de milho por RT-qPCR

Nota: *Iniciadores que foram redesenhados para atingir a padronização, pelo fato de o primeiro iniciador encomendado não passar no teste da curva de *Melting*. Fonte: Autoria própria
4.9 ANÁLISE FENOTÍPICA DA COMPOSIÇÃO QUÍMICA DA PAREDE CELULAR DAS LINHAGENS DE MILHO TRANSGÊNICAS E SELVAGEM B73

4.9.1 Isolamento da parede celular do colmo das plantas de milho

Após as plantas atingirem o estágio de senescência, os colmos foram colhidos e macerados em moinho tipo ciclone UDY (Udy Corp, Fort Collins, CO, USA), com uma peneira de 1mm. O material macerado foi deixado secar por 24 horas em uma estufa a 55°C. Tubos para centrífuga de 50 mL (*Nalgene Oak Ridge*) também foram secos a 55°C e pesados. O material macerado seco foi então pesado diretamente nessas garrafas. As amostras foram incubadas em solução tampão 50mM Tris-acetato (pH 6) de um dia para o outro à 4°C, e depois por 2h sob agitação em temperatura ambiente. O material foi centrifugado à 6500 rpm por 20 min. Os sobrenadantes foram cuidadosamente separados e descartados. Os pellets foram extraídos mais duas vezes com o mesmo tampão. As amostras foram então extraídas com etanol 80% usando o mesmo procedimento, por 3 vezes. Os pellets foram extraídos mais 3 vezes com acetona e uma vez com uma mistura clorofórmio: metanol (2:1 v/v). As amostras foram então lavadas com acetona para remover a mistura de clorofórmio e etanol. Os resíduos finais de parede celular foram secos ao ar deixando-se os tubos abertos numa capela. O material foi armazenado na estufa a 55°C até o uso posterior nos experimentos.

4.9.2 Determinação do conteúdo total de ácidos hidroxicinâmicos éster da parede celular

A quantidade total de ácidos ferúlico e *p*-cumárico unidos por ligação éster aos componentes da parede celular foi determinada por meio de tratamento alcalino brando e quantificação por cromatografia gasosa com detecção por ionização em chama (GRABBER et al., 1995).

Aproximadamente 100mg do material da parede celular seco a 55°C foram pesados em frascos *Nalgene* de 2.5 mL com tampa revestida de teflon, em duplicata. Foram adicionados 100µL de uma solução 1mg/L de ácido 2-hidroxicinâmico em hidróxido de sódio (NaOH) como padrão interno. As paredes celulares foram hidrolisadas com 2.4 mL de NaOH 2M, previamente desgaseificado com $N_{2(g)}$, à temperatura ambiente por cerca de 20 horas. As amostras foram acidificadas para pH menor que 2 com ácido

clorídrico HCl 12.1 M. Em cada frasco, foram adicionados 2 mL de éter dietílico para extração dos compostos fenólicos. Agitaram-se e centrifugaram-se os frascos à 1000 rpm por 5 minutos. A fase orgânica (topo) foi transferida para frascos de vidro para cromatografia de 3 mL, usando uma pipeta Pasteur. O processo de extração foi repetido um total de 3 vezes. O solvente foi evaporado usando um concentrador de amostras. Os ácidos ferúlico e p-cumárico foram identificados e quantificados como derivativos de Trimetilsilano. Para isso, foram adicionados aos frascos 40µL de TMSI (N-trimetilsililimidazol) e 10µL de piridina e deixou-se reagir por 30 minutos à 60°C. As amostras foram então injetadas em um aparelho de cromatografia gasosa com deteccão por ionização em chama GLC-FID (Agilent technologies 7890 GC system), coluna ZB-1 (Phenomenex, Zebron 100% dimethylpolysiloxane; 30m X 0.25 mm, filme 0.25µm). As condições foram: Injetor 315°C, detector 300°C, e um programa de temperatura de 200°C por 1 min, aumento de 4°C min⁻¹ até 248°C, fixo por 2 minutos, seguido de aumento de 30°C min⁻¹ até 300°C, mantendo essa temperatura por 20 minutos. Todos os programas de temperatura foram corridos a uma pressão constante de 20 psi e taxa de separação de 30:1. Soluções padrão de FA e pCA foram utilizadas para identificação dos compostos pelo tempo de retenção e quantificação pelo método do padrão interno (Figura 10). O fator de resposta médio (FR) para cada um dos compostos foi calculado por meio da equação (1), em que Ap e Cp são a área do pico no cromatograma e concentração padrão de um analito (pCA ou FA), enquanto Api e Cpi são a área do pico no cromatograma e concentração do padrão interno (ácido 2-hidroxicinâmico), respectivamente.

Figura 10– Cromatograma das soluções padrão contendo padrão interno ácido 2hidroxicinâmico e os padrões ácido *p*-cumárico e ácido ferúlico na concentração de 1mg/ml



Legenda; 1- Ácido 2-hidroxicinâmico, tempo de retenção 4.7 min. 2- Ácido p-cumárico, tempo de retenção 6.0 min. 3 – Ácido Ferúlico, tempo de retenção 7.9 min. Os tempos de retenção foram anteriormente determinados injetando-se cada padrão separadamente. Fonte: Autoria própria.

$$FR = \frac{Api}{Cpi} \times \frac{Cp}{Ap} \tag{1}$$

Após a obtenção do fator de resposta, reescrevendo a equação (1), a concentração do analito (pCA ou FA) na amostra foi calculada pela equação (2), em que Cx é a concentração desconhecida do analito, FR é o fator de resposta para este analito, Ax é a respectiva área do pico deste analito e Api é área do padrão interno no cromatograma.

$$Cx = FR \times Cpi \times \frac{Api}{Ax}$$
(2)

4.9.3 Determinação do conteúdo de ácidos hidroxicinâmicos unidos especificamente à hemicelulose GAX

Com o objetivo de quantificar apenas os ácidos hidroxicinâmicos provenientes da hemicelulose, e não da lignina, foi feita a acidólise branda da parede celular, seguindo o protocolo descrito por Lapierre *et al.* (2018). Com esse procedimento, são clivadas as ligações glicosídicas da xilana, mas a maior parte das ligações ésteres que unem os ácidos hidroxicinâmicos à arabinose são preservadas (LAPIERRE et al., 2018). Assim, são liberados os dímeros Ara-FA (Arabinose-ferulato) e Ara-*p*CA (Arabinose-*p*-cumarato), os quais são quantificados.

O reagente de acidólise foi preparado fresco misturando-se dioxano, metanol e 2M HCL(aq) (60:30:10, v/v/v). Cerca de 10 mg do material extraído da parede celular foi adicionado com 1 mL do reagente de acidólise em um frasco de 4mL equipado com uma tampa revestida de Teflon. Os frascos foram incubados por 3 horas a 80°C sob agitação a 500 rpm. A mistura reacional resfriada foi diluída com cerca de 2 mL de água, e 0.1 mL de padrão interno (ácido 2- hidroxicinâmico, oCA), solução 0.5 mg/mL em metanol. Antes da quantificação por cromatografia líquida de alta pressão (HPLC), os fenólicos foram extraídos da mistura diluída por extração em fase sólida em um cartucho Sep-Pak tC18 Plus Short cartridge (Waters, Guyancourt, France). Antes do uso, a coluna foi précondicionada pela lavagem (2 × 3 mL MeOH, depois 2 × 3 mL 0.1% aq HCOOH). Então, uma alíquota da mistura reacional diluída (cerca de 1 mL) foi centrifugada e cerca de 0.5 mL do sobrenadante foi passado pelo cartucho. O cartucho foi então lavado com cerca de 1 mL 0.1% HCOOH(aq) e eluído com 1mL MeOH. A solução eluída foi guardada refrigerada a -20C por 30 minutos antes de ser ultrafiltrada em um filtro de disco de poliéster Chromafil Xtra 0.45 µm (Macherey-Nagel, Hoerdt, France). Cerca de 1 µL de cada amostra foi injetado em um Shimadzu Nexara X2 HPLC equipado com uma coluna C18 Phenomenex Kinetex (2.6 µm X 2.1 mm X 150 mm, P/N: 00F-4462-AN). A fase móvel foi um gradiente binário do solvente A: Água mais 0.1% de ácido fórmico e solvente B: Acetonitrila + 0.1% de ácido fórmico. O detector foi um arranjo de foto-diodo $\lambda = 250-600$ nm, a quantificação foi feita a $\lambda = 270$, e uma curva de calibração de 5 pontos foi determinada usando autênticos MeAra-pCA, MeAra-FA, e oCA (54 µg/mL, padrão interno).

4.9.4 Determinação do conteúdo de ácidos hidroxicinâmicos unidos especificamente à lignina

Com o objetivo de quantificar apenas os ácidos hidroxicinâmicos unidos à lignina, e não à hemicelulose, foi utilizado o método DFRC (*Derivatization Followed by Reductive Cleavage*) modificado (REGNER et al., 2018). Este método cliva seletivamente ligações β -éter da lignina enquanto retém as ligações éster, fornecendo evidências da incorporação de *p*CA-monolignol à lignina (Figura 11).

Figura 11 – Mecanismo das etapas do processo DFRC (*Derivatization Followed by Reductive Cleavage*)



Legenda: A acetilação das unidades hidroxila e bromação das posições benzílicas por brometo de acetila (AcBr)/Ácido acético (AcOH) (derivatização) precede a clivagem redutora de 2 elétrons de ligações β -éteres e debromação de unidades derivadas de cinamato via zinco/ ácido acético. Na etapa final a mistura de produtos é acetilada usando uma mistura de anidrido acético (Ac₂O) / piridina. Fonte: Regner et al., 2018.

Amostras da parede celular isolada (50 mg) foram pesadas em um frasco de 8 mL com tampa revestida de teflon (PTFE) (*Chemglass* GC-4912-02), em triplicata, e incubadas em uma solução de brometo de acetila/ ácido acético (1/4 v/v, 3 mL) à 50 °C por 2.5 horas, sob agitação. O solvente foi removido em um evaporador *SpeedVac* (*Thermo Scientific*TM *SPD131DDA*, 50 °C, 35 min, 1.0 torr, 35 torr/min). O filme resultante foi suspendido em etanol absoluto (0.5 mL), e em seguida o etanol foi removido no evaporador *SpeedVac* (50 °C, 15 min, 6.0 torr, 35 torr/min). O resíduo foi suspendido em uma mistura 1,4 dioxano: acido acético: água (5/4/1 v/v, 5 mL) e 150 mg de zinco em nano-pó foram adicionados. Os tubos foram vedados e sonicados por 1 hora a temperatura ambiente. As reações foram calibradas com os padrões internos isotopicamente marcados (H-*d*₈, G-*d*₈, S-*d*₈, G-DD*p*CA-*d*₁₀, S-DD*p*CA-*d*₁₀, G-DDFA-*d*₁₀, e S-DDFA-*d*₁₀) e transferidas quantitativamente, usando diclorometano (DCM, 2x2mL), para um funil

separatório carregado com cloreto de amônio saturado (10mL). As fases orgânicas foram extraídas três vezes com 10 mL de DCM, combinadas e os resquícios de água removidos com sulfato de sódio anidro. As soluções foram filtradas por meio de papel de filtro qualitativo e o solvente removido em um evaporador rotatório (banho quente a menos de 50°C). Os grupos hidroxila livres foram então acetilados por 30 minutos em uma solução de piridina e anidrido acético (1/1 v/v, 5mL). O solvente foi removido por evaporação rotatória resultando em um filme bruto oleoso. Para remover a maior parte dos produtos derivados de polissacarídeos presentes neste filme, o produto bruto da reação DFRC foi transferido quantitativamente com DCM (3x 1.0mL) para um cartucho para extração em fase sólida (Supelco Supelclean LC-Si SPE tube, 3 mL, P/N: 505048). Os produtos foram eluídos com uma mistura hexano: etil acetato (1:1 v/v, 8mL). Os solventes combinados foram removidos num evaporador rotatório e transferidos com cerca de 1 mL de DCM para frascos para cromatografia gasosa. As amostras foram analisadas em um aparelho de cromatografia gasosa acoplado a um espectrômetro de massa (GC-MS/MS) triploquadrupolo (Shimadzu GCMS-TQ8030) operando em modo de monitoramento de múltiplas reações (MRM), usando padrões sintéticos para calibração e identificação. Curvas de calibração foram determinadas pelo método do padrão interno como descrito no item 4.9.3, ou seja, o fator de resposta foi obtido a partir da razão entre as áreas dos picos dos produtos sintéticos trans e os correspondentes padrões internos marcados isotopicamente multiplicados pela razão de suas concentrações.

4.9.5 Determinação do conteúdo de açúcares neutros totais da parede celular

O total de açúcares neutros da parede celular foram determinados seguindo o procedimento de hidrólise de Saeman (SAEMAN; MOORE; MILLETT, 1963), como modificado por (HATFIELD et al., 2009). As amostras de parede celular isoladas foram pesadas (100mg) em tubos de ensaio, em duplicata. Em paralelo, também foram analisados em duplicata 10g de uma amostra padrão de açúcares neutros, contendo 5% de ramnose, 5% fucose, 20% arabinose, 20% xilose, 10% manose, 10% galactose, 30% glicose (m/m). As amostras e o padrões foram incubados em 1.5 mL de ácido sulfúrico 12M gelado, por 2 horas à temperatura ambiente, e então diluídos com 10.5 mL de H₂O *milli-Q* e hidrolisados por 3 horas à 100°C. Após esfriarem, as amostras foram centrifugadas (1000 rpm, 10 min) para precipitação dos sólidos. Foram adicionados 200µL de uma solução 50mg/mL de mio-inositol como padrão interno, e 1.5 mL do hidrolisado foram transferidos para tubos

de centrífuga de 50mL contendo 8.5 mL de água. As amostras foram neutralizadas com carbonato de bário e centrifugadas (4000 rpm, 10 min). Os sobrenadantes foram filtrados com filtros de seringa de fibra de vidro (GF 1.0 μ m *syringe filter*, diâmetro de 25 mm, hidrofílica, *TISCH scientific*TM) para frascos de 4mL e evaporados em um concentrador de amostras (N-vap).

Os açúcares foram então convertidos em seus derivados de alditol para quantificação por GC (BLAKENEY et al., 1983). Os filmes formados contendo os resíduos de açúcar foram suspendidos em 100 µL de hidróxido de amônia (NH4OH) 1 M. Foram adicionados 500 µL de uma solução de Borohidreto de sódio (NaBH₄) 25 mg/mL em Dimetilsulfóxido (DMSO). Os frascos tampados foram incubados à 45°C por 90 minutos (Banho quente), agitando-os a cada 30 minutos. O excesso de NaBH4 foi destruído com 100 µL de ácido acético glacial. Os grupos álcoois obtidos após essa etapa foram então acetilados. Adicionaram-se 100 µL de 1-metil-imidazol aos frascos e 1 mL de anidrido acético, e agitou-se vigorosamente. Os frascos foram incubados à temperatura ambiente por 20 minutos. Foram adicionados 2.5 mL de H₂O milli-Q e a fase orgânica (inferior) foi extraída duas vezes com 1 mL de DCM. Os extratos foram combinados em um frasco de 8 mL e o solvente foi removido em um concentrador de amostras (N-vap). Os derivados de açúcar foram transferidos com cerca de 100 µL de DCM para um frasco para GC e quantificados por cromatografia gasosa com detecção por ionização em chama (GLC-FID, Shimadzu GC-2010) usando uma coluna 07-225 methylpolysiloxane (30m X 0.25mm with 0.25µm filme, Qadrex corporation). As condições foram injetor 220°C, detector 240°C, e um programa de temperatura de 215°C por 2 minutos, aumento de 4°C até 230°C mantidos por 11.25 min, corrida a velocidade linear constante de 33.4cms⁻¹ e razão de separação 25:1.

4.9.6 Análise estatística

Os experimentos de RT-qPCR e acidólise branda foram executados em triplicata, e a quantificação de fenólicos ésteres totais, açúcares neutros da parede celular e DFRC em duplicata. Todas as análises estatísticas foram executadas no *software* R (*https://www.R-project.org/*). As médias foram analisadas e as diferenças estatísticas foram determinadas usando análise da variância *One Way ANOVA* seguida de testes *post-hoc* de comparação múltipla (Tukey ou Dunnet, α =0.05).

5 RESULTADOS E DISCUSSÃO

5.1 IDENTIFICAÇÃO, ANOTAÇÃO E ANÁLISE FILOGENÉTICA DOS GENES DE BIOSSÍNTESE DOS ÁCIDOS HIDROXICINÂMICOS

As proteínas pertencentes à superfamília das BAHD aciltransferases, contendo o domínio PFAM PF02458 (transferase), foram anotadas em espécies de gramíneas e na eudicotiledônea Arabidopsis. Foram identificadas 63 sequencias de proteínas preditas do genoma em Arabidopsis e 105 (milho), 132 (arroz), 96 (sorgo), 95 (Brachypodium) e 130 (Setaria), respectivamente. As proteínas BAHD codificadas pelos genomas de Arabidopsis, sorgo, arroz e Brachypodium já foram identificadas em trabalhos anteriores (D'AURIA et al. ,2005; YU et al., 2009; TUOMINEN et al., 2011; BARTLEY et al., 2013), enquanto as de milho e Setaria foram identificadas pela primeira vez nesse trabalho. Fazendo-se uma comparação com a quantidade de proteínas identificadas na literatura, observou-se que no presente trabalho foi encontrado número similar ou um pouco maior, o que demonstra a eficácia do método de busca adotado (Tabela 9). As diferenças obtidas estão relacionadas principalmente ao método de busca e a versão do genoma utilizado. Em seus trabalhos, tanto Bartley et al. (2013) como Tuominen et al. (2011) não consideraram as sequências que não possuíam o motivo conservado HXXXD, considerado essencial para a função das enzimas BAHD. Tuominen et al. (2011) excluíram ainda sequências muito curtas, considerando-as como provenientes de possíveis pseudogenes. No presente trabalho, a quantidade de sequências que visualmente não apresentaram o motivo HXXXD conservado foi representada entre parênteses na tabela 6. Entretanto, todas foram consideradas nas análises posteriores, uma vez que se trata de um motivo com alta variabilidade (3 de 5 aminoácidos) e enzimas BAHD funcionais já foram caracterizadas com variações no primeiro H (TUOMINEN et al., 2011).

Doforôncio	Espécie						
Kelelencia	Arroz Brachypodium		Sorgo	Milho	Setaria	Arabidopsis	
(D'Auria et al., 2005)	-	-	-	-	-	64	
(Yu et al., 2009)	-	-	-	-	-	62	
(Tuominen et al., 2011)	84	-	-		-	55	
(Bartley et al., 2013)	117	78	85	-	-	64	
Presente trabalho	132 (115)	95(85)	96(84)	105	130	63	

Tabela 9– Quantidade total de proteínas BAHD aciltransferases identificadas nas diferentes espécies vegetais e comparação com os dados da literatura

Nota: Os números entre parênteses indicam a quantidade de proteínas possuindo o domínio HXXXD conservado.

Fonte: Autoria própria.

Este trabalho também é o primeiro a identificar o conjunto de sequências pertencentes à família BAHD aciltransferases em cana-de-açúcar. A quantidade de sequências identificadas em cada um dos bancos genômicos, de sequências expressas e transcriptômicos foi representada na tabela 10.

Tabela 10– Número de sequências identificadas em cada banco de dados de cana de açúcar: Genoma (DNAg), SUCEST, RNA-seq a partir de plântula, entrenó, folha e raiz.

Referência	Banco de dados	#sequências
(Riaño-Pachón e Matiello, 2017)	DNAg parcial SP80-3280	183
(Garsmeur et al., 2018)	DNAg monoploide R570	60
(Vettore et al., 2003)	SUCEST	75
Dra. Adriana Hemerly	Plântula	74
(Vicentini et al., 2015)	Entrenó	54
(Cardoso-Silva, 2014)	Folha	58
Dr. Marcos Buckeridge	Raiz	23

Fonte: Autoria própria

O banco de proteínas preditas do genoma parcial de cana-de-açúcar apresentou 183 sequências BAHD, maior do que o presente nas demais gramíneas estudadas neste trabalho (Tabela 10). Esse valor é bem alto, sobretudo considerando que ainda não representa o do genoma completo. Isso é um provável reflexo do grande tamanho e poliploidia do genoma de cana-de-açúcar. Algumas sequências podem representar apenas alelos diferentes do mesmo gene. Já o valor de 60 sequências identificadas para o banco de proteínas preditas do cultivar R570 refere-se apenas ao genoma monoplóide, e foi significativamente menor comparado com as demais gramíneas analisadas. No entanto, não é possível afirmar que a espécie cana-de-açúcar possuí menos sequências BAHD do que as

demais gramíneas, já que esse genoma é monoplóide e parcial, representando apenas uma região enriquecida de genes (GARSMEUR et al., 2018). O número de sequências identificadas no genoma R570 também foi menor do que o de sequências expressas e transcritos de cana-de-açúcar (SUCEST e plântula), corroborando com o fato de que este genoma recentemente publicado ainda não está completo.

No banco do SUCEST, foram identificadas 75 sequências expressas, que são provenientes de um conjunto de bibliotecas de diversos tecidos (flor, folha, entrenó). No RNA-seq a partir de plântula, foi encontrado um número similar (78), indicando que os genes BAHD são bastante expressos nesta fase inicial do desenvolvimento da planta. Já as sequências identificadas nos bancos de RNA-seq a partir de entrenó (54), folha (58) e raiz (23), representam apenas os transcritos específicos destes tecidos. Esses valores sugerem que existem menos sequências BAHD expressas em raiz do que em folha e entrenó.

Uma análise comparativa da evolução da família BAHD nas diferentes espécies foi feita inferindo-se uma filogenia pelo método da máxima verossimilhança. As sequências identificadas foram divididas entre os clados previamente classificados por Tuominen *et al.* (2011) como Ia, Ib, II, IIIa, IIIb, IV, Va e Vb. A quantidade de proteínas em cada um dos clados provenientes do genoma de cada espécie analisada, bem como a de sequências de cada um dos bancos de cana-de-açúcar foi representada na Figura 12.

Corroborando com os dados de Tuominen *et al.* (2011), no clado IV somente foram identificadas sequências de gramíneas, enquanto no clado IIIa foram identificadas apenas sequências da eudicotiledônea *Arabidopsis thaliana*. O fato do clado IV conter apenas sequências de gramíneas pode ser relevante, uma vez que podem estar relacionadas a características específicas deste grupo de angiospermas. Outro aspecto é que este clado sofreu retração em sorgo relativo às demais gramíneas. Ademais, verificou-se que o clado Va, que abrange o subclado relacionado com a incorporação dos ácidos hidroxicinâmicos (Clado de Mitchell), contém grande quantidade de sequências em todas as gramíneas analisadas (54 em arroz, 34 em sorgo, 39 em milho, 30 em Brachypodium, 39 em Setaria e 15 no genoma monoplóide de cana-de-açúcar). Por essa comparação, nota-se também que este clado sofreu expansão gênica em arroz. Já o clado Vb sofreu expansão em sorgo e milho, espécies próximas evolutivamente, uma vez que foram encontradas 22 sequências em sorgo, 20 em milho e 10 no genoma monoplóide de cana-de-açúcar, enquanto apenas 10, 13 e 15 foram identificadas em arroz, Brachypodium e Setaria, respectivamente. No caso de cana-de-açúcar, o sequenciamento completo do genoma ainda é necessário para dar suporte à comparação com as outras espécies de gramíneas, uma vez que podem haver genes não incluídos na análise.

Comparando-se a quantidade de sequências identificadas no genoma e transcriptomas de cana-de-açúcar, nota-se que de uma forma geral a distribuição entre os clados é similar, com maior quantidade de sequências nos clados Va e Vb. Essa informação é relevante uma vez que o clado Va contém o grupo relacionado à incorporação dos ácidos hidroxicinâmicos. Considerando que parte das enzimas desse grupo, bem como as do clado Vb estão relacionadas com a síntese da parede celular, esperava-se que de fato os genes do clado Va e Vb fossem expressos em todos os tecidos analisados (folha, entrenó e raiz), como foi encontrado. Destacam-se os padrões de distribuição entre os clados nos transcritos de raiz e entrenó. No banco de RNA-seq proveniente de raiz apenas um transcrito pertence ao clado Ia enquanto os demais pertencem ao clado Va e Vb. Outro aspecto interessante é que no banco de RNA-seq de entrenó, foram identificados mais transcritos no clado Vb (21), que contém enzimas relacionadas com a síntese de lignina, que provavelmente são expressas no colmo em lignificação.

Figura 12 – Distribuição entre os clados filogenéticos das sequências BAHD identificadas no genoma das espécies e nos bancos genômicos, de sequências expressas e transcritos de cana-de-açúcar.



Legenda: Quantidade de sequências em cada um dos clados da família BAHD de acordo com a classificação de TUOMINEN et al, 2011. Fonte: Autoria Própria.

Como o foco deste trabalho são as enzimas BAHD relacionadas com a biossíntese dos ácidos hidroxicinâmicos na parede celular de gramíneas, uma análise mais detalhada foi feita apenas com o clado que contém as proteínas identificadas na literatura com essa função. Mitchell et al. (2007) inicialmente identificaram esse grupo, com 12 proteínas em arroz, e Bartley et al. (2013) ampliaram a análise identificando 20 proteínas em arroz e verificaram que esse clado se subdivide em dois: Um contendo OsAt1 a OsAt10 (grupo A) e outro OsAt11 a OsAt20 (grupo B). Até o presente momento, todas as enzimas caracterizadas funcionalmente relacionadas com os ácidos hidroxicinâmicos pertencem ao primeiro grupo (Tabela 7). Ademais, no grupo B, a expressão de poucos genes de arroz foi confirmada. Nas análises filogenéticas de toda a família BAHD, inferidas pelo método da máxima verossimilhança, verificou-se que poucas proteínas das demais gramíneas pertencem ao grupo B. Dessa forma, foi representada uma filogenia construída a partir das proteínas preditas do genoma de cada uma das espécies analisadas, pertencentes ao grupo A do clado de Mitchell (Figura 13). No caso de cana-de-açúcar, foi escolhido um representante (Unigene) do cluster de grupos ortólogos (COG, Cluster of Orthologous Groups). Alguns unigenes do clado analisado foram encontrados apenas em bancos transcriptômicos ou de sequências expressas. De fato, os genomas disponibilizados recentemente ainda não representam a totalidade de genes de cana-de-açúcar, de forma que o uso de diferentes bancos de RNA-seq amplia a robustez das análises nesta espécie. Então, nestes casos, foi representado o transcrito cuja sequência fosse mais completa, quando comparada com o ortólogo em sorgo.



Figura 13 - Análise filogenética do clado A de Mitchell das BAHD aciltransferases

Legenda: A análise foi conduzida nas espécies: Cana-de-açúcar (verde), sorgo (vermelho), milho (amarelo), arroz (azul-marinho), Brachypodium (azul claro) e Setaria (azul escuro), utilizando o método da máxima verossimilhança, modelo JTT, no algoritmo PhyML 3.1 do software *SeaView*. O suporte da topologia indicado em cada nó foi calculado pelo método aLRT. Clados destacados com ramos coloridos contém sequências já caracterizadas. Fonte: Autoria própria.

No grupo A de Mitchell, existe apenas um representante de Arabidopsis (At3g62160.1), identificado inicialmente como único ortólogo de 12 genes de arroz (MITCHELL et al., 2007). Por outro lado, várias proteínas das espécies de gramíneas analisadas neste trabalho foram identificadas neste clado, corroborando com os dados de Bartley *et al.* (2013) de que esse grupo sofreu expansão gênica em gramíneas relativo às eudicotiledôneas (BARTLEY et al., 2013). Ademais, é possível observar em todos os clados que as sequências de cana-de-açúcar são mais próximas evolutivamente das de sorgo e milho, seguida das de setaria, enquanto Brachypodium e arroz formam um grupo separado, concordando com a relação filogenética entre as espécies.

Na filogenia, é possível identificar subclados contendo enzimas já caracterizadas funcionalmente em gramíneas. Os subclado contendo BdAt1, recentemente caracterizado com provável função de incorporação de FA à hemicelulose (BUANAFINA et al., 2016), foi destacado em vermelho. Nota-se que houve duplicação gênica em milho. O grupo At2, homólogo de At1, apresentou um representante de cada espécie. Ainda não se sabe se há redundância funcional de At2 e At1.

Também foi identificado o grupo contendo OsPMT, ZmpCAT e BdPMT, relacionados com a incorporação de pCA na lignina (clado em azul). Arroz possui dois parálogos nesse clado, OsAt4 (PMT) e OsAt3, sendo que apenas o primeiro foi caracterizado funcionalmente, de forma que ainda não foi verificado se há redundância funcional. Nota-se também que houve duplicação em Setaria, com três sequências neste grupo. Sorgo apresentou apenas um representante. As proteínas já caracterizadas funcionalmente desse grupo possuem nomenclaturas distintas (PMT ou At4 e pCAT). Neste trabalho, as sequências de cana-de-açúcar identificadas foram denominadas ScAt4, seguindo a nomenclatura das sequências utilizada por Bartley *et al.* (2013).

O subclado que contém *OsAt5*, recentemente identificado como responsável pela incorporação de FA ao monolignol (KARLEN et al., 2016), foi destacado em rosa. Nessa análise, não foram encontradas sequências genômicas, entretanto foram identificados transcritos no entrenó e raiz.

O subclado contendo *OsAt10*, relacionado com à incorporação de pCA à hemicelulose (BARTLEY et al., 2013), foi representado em laranja. Cada espécie de gramínea contém apenas um representante de *At10*. O grupo *At9*, cuja função foi recentemente descrita como associada à incorporação de FA à hemicelulose em Setaria e cana-de-açúcar (DE SOUZA et al., 2018, 2019), foi destacado em verde e também apresentou um representante de cada espécie.

Já para o grupo de genes *At6* e *At7*, ainda não existem estudos que atribuam uma função específica. Embora em estudos com arroz o silenciamento conjunto de *OsAt6* a *OsAt10* tenha levado à diminuição de FA e uma linhagem com alteração na expressão de *OsAt7* (*activation tagged*) tenha apresentado conteúdo modificado de FA (PISTON et al., 2008; BARTLEY et al., 2014), em Brachypodium o silenciamento e superexperssão (35S) de *BdAt7* não resultou em alteração no conteúdo de ácidos hidroxicinâmicos na parede celular (BUANAFINA et al., 2016). Na filogenia, é possível observar que houve duplicação gênica em milho, com dois genes *At6* e dois *At7*, e em Brachypodium, com dois genes *At6*. É interessante observar que Setaria apresentou apenas uma sequência *At6*, sugerindo que durante a evolução essa espécie de gramínea perdeu um representante *At7*. É possível que *At6* e *At7* apresentem a mesma função, e em Setaria exista apenas um gene para essa atividade. Em cana-de-açúcar novamente não foi encontrada a sequência nos genomas parciais, mas nos transcriptomas de plântula, entrenó e raiz.

Para as sequências At8, foi sugerido recentemente que BdAt8 ou BdPMT2(Bradi1g36980) tenha também papel na incorporação de pCA na lignina (LAPIERRE et al., 2016), entretanto mais estudos de caracterização funcional desse gene são requeridos. Na filogenia, nota-se que houve duplicação gênica em Brachypodium, e ainda não foi verificado se a sequência Bradi5g15600 pode também estar relacionada com a atividade PMT.

5.2 ANÁLISE DE EXPRESSÃO *IN SILICO* DOS GENES DE CANA-DE-AÇÚCAR DO GRUPO A DE MITCHELL EM DIFERENTES TECIDOS

Para analisar a expressão dos genes *ScAt1* a *ScAt10* em diferentes tecidos de canade-açúcar, os dados de abundância dos transcritos provenientes das análises de RNA-seq e da quantidade de sequências expressas em cada biblioteca do SUCEST fornecidos pelos colaboradores foram agrupados (*clusterizados*) e comparados por meio de *heatmaps* (Figura 14).



Figura 14 – Análise *in silico* do padrão de expressão dos genes *ScAt1* a *ScAt10* em diferentes transcriptomas.

5.5 12 5 11 4.5 At2 Sc entreno.2 At8_Sc_raiz At5_Sc_entreno 10 At4_Sc_raiz.1 At7_Sc_folha At10_Sc_entreno 4 At5 Sc raiz.1 At9 Sc folha At2_Sc_entreno.1 Q At5_Sc_raiz.2 3.5 At1 Sc folha At4_Sc_entreno At6_Sc_raiz.1 At8_Sc_folha 2 At7_Sc_entreno 8 3 At6_Sc_raiz.2 At4_Sc_folha.1 At8_Sc_entreno At4_Sc_raiz.3 At2 Sc folha At1_Sc_entreno 0 2.5 SASO 5345 3050 627 (*Horina) At1 Sc raiz 810C O65 (-IIGIINA) At9 Sc entreno At9 Sc raiz S, S \$ 5

Legenda: A- Dados de expressão do SUCEST (VETTORE et al., 2003). B- Expressão dos transcritos em folha das variedades IAC-SP96-3046, SP81-3250, SP80-3280, RB83-5486, RB92-5345, IAC-SP95-3018 (CARDOSO-SILVA et al., 2014). C- Expressão de transcritos em entrenós das variedades IAC-SP04627 (maior lignina) e IAC-SP04065 (menor lignina), com conteúdo contrastante de lignina (VICENTINI et al., 2015). D- Expressão em quatro regiões diferentes de raiz (S1 a S4) (PIOVEZANI, 2017). A análise foi feita no *software* R a partir dos dados de expressão fornecidos pelos colaboradores Dr. Renato Vicentini e Dr. Marcos Silveira Buckeridge. Foi feito agrupamento hierárquico pelo método euclidiano e distância completa. Fonte: Autoria Própria.

Nesta análise, os dados de expressão de cada biblioteca do SUCEST foram agrupados por tecido (entrenó, folha, raiz, inflorescência, meristemas (apical e lateral), sementes e calo), e o agrupamento hierárquico foi feito somente entre as diferentes sequências, o que permitiu visualizar o padrão de expressão dos genes analisados nos tecidos (Figura 9A), e comparar com o padrão de expressão em folha e entrenó nas distintas variedades (Figura 9B), bem como o de raiz nas distintas regiões (Figura 9C).

Um aspecto relevante é que os genes *ScAt1* e *ScAt9* agruparam em todos os bancos, indicando padrão de expressão elevado e similar nos tecidos analisados. Esse perfil sugere co-expressão desses genes em alguns tecidos, indicando que podem ter função similar ou relacionada. Isso corrobora com o possível papel de *ScAt1* na incorporação de FA à hemicelulose (BUANAFINA, 2016), o mesmo já demonstrado para *ScAt9* (DE SOUZA et al., 2019).

Adicionalmente, o gene *ScAt6* foi altamente expresso apenas em raiz, verificado tanto nos dados de RNA-seq como nas bibliotecas do SUCEST, enquanto o homólogo *ScAt7* foi mais expresso em entrenó e pouco em folha. Essa informação suporta a hipótese de que *ScAt6* poderia ser funcional em raiz, o que precisaria ser verificado em estudos de caracterização funcional desses genes.

Para o gene *ScAt10*, notou-se níveis baixos de expressão no entrenó, um pouco maior no cultivar com maior quantidade de lignina, e expressão apenas em inflorescência da biblioteca do SUCEST. Os níveis baixos de expressão suportam a função de incorporação de *p*CA à hemicelulose, encontrada para o ortólogo em arroz O*sAt10* (BARTLEY et al., 2013), uma vez que a maior parte do *p*CA das gramíneas encontra-se ligado à lignina (HATFIELD et al., 2009; LU; RALPH, 1998).

A análise dos *heatmaps* demonstrou que o gene *ScAt4* foi altamente expresso em folha, raiz e entrenó, esse último verificado tanto pelo RNA-seq como nas bibliotecas de colmo do SUCEST. Nos dados de RNA-seq de entrenó, também foi verificado maior expressão na variedade que contém mais lignina. Assim, os dados de expressão corroboram com o possível papel de *ScAt4* na síntese de lignina, de incorporação de pCA ao monolignol, já identificada nos ortólogos *OsPMT*, *BdPMT* e *ZmpCAT*. Ademais, uma observação interessante é que os genes *ScAt8* agruparam com *ScAt4* em quase todos os *heatmaps*, e também apresentaram maior nível de expressão na variedade com maior quantidade de lignina. Esse dado corrobora com a possibilidade dos genes *ScAt8* também possuírem atividade *p*-cumarato monolignol transferase, sugerida pelo estudo em Brachypodium (LAPIERRE et al., 2016).

5.3 ANÁLISE DE EXPRESSÃO DOS GENES DE CANA-DE-AÇÚCAR DO GRUPO A DE MITCHELL AO LONGO DO DESENVOLVIMENTO DO COLMO DE HÍBRIDOS CONTRASTANTES

Uma análise mais refinada de expressão por PCR em tempo real (RT-qPCR) foi feita para o grupo de genes de Mitchel identificados em cana-de-açúcar, em dois híbridos com teor de ácidos hidroxicinâmicos distintos, H89 (1.1% FA e 3.9% *p*CA) e H321 (1.4% FA e 6.8% *p*CA) (MASARIN et al., 2011), e em diferentes estágios de desenvolvimento do colmo. Foram analisados entrenós representando estágios progressivos de desenvolvimento (1,5,11) separados em córtex e medula. Com isso, objetivou-se encontrar padrões de expressão que pudessem fornecer indícios da potencial relação de cada um desses genes com a incorporação dos ácidos hidroxicinâmicos na hemicelulose ou lignina. Os resultados de expressão normalizados pelo gene de referência foram agrupados (*clusteirzados*) em um *heatmap* (Figura 15)

Figura 15 – Análise do padrão de expressão por RT-qPCR dos genes *ScAt1* a *ScAt10* em diferentes tecidos de híbridos de cana-de-açúcar H89 e H321



Legenda: Dados de expressão normalizados com o gene de referência (ΔCt equivalente, corrigido pela eficiência dos iniciadores) e convertidos para uma escala de 0 a 10. Os valores foram agrupados hierarquicamente pelo método euclidiano e distância completa. A- Valores para os entrenós dos híbridos H89 e H321, separados em córtex e medula. EC1, EC5 e EC9 representam os valores do córtex dos entrenós 1,5 e 9 respectivamente. EM1, EM5 e EM9 representam os valores da medula dos entrenós 1,5 e 9, respectivamente. B- Expressão em folha jovem de dois indivíduos do híbrido H89(H89.1 e H89.3) Fonte: Autoria Própria.

Corroborando com a análise *in silico* dos dados de RNA-seq e sequências expressas (item 5.2), os genes *ScAt1*, *ScAt4* e *ScAt9* apresentaram altos níveis de expressão. Em uma análise por RT-qPCR em *Brachypodium*, os homólogos desses 3 genes (*BdAt1*, *BdAt4* e *BdAt9*) também foram os mais expressos em diferentes tecidos em distintos estágios de desenvolvimento (MOLINARI et al., 2013). Esses resultados condizem com o envolvimento desses genes identificados nas principais atividades de incorporação dos ácidos hidroxicinâmicos na parede celular, a de ligação de *p*CA à lignina (*At4*) e FA à GAX (*At1* e *At9*). Ainda confirmando a análise *in silico*, os genes *ScAt1* e *ScAt9* também agruparam no *heatmap*, fornecendo mais um indicativo de que ambos estão relacionados com a incorporação de FA à hemicelulose.

Já os níveis de expressão de ScAt10 e ScAt5 não foram detectáveis, em contraste com o encontrado na análise de RNA-seq do entrenó de cultivares de cana (Figura 14, seção 5.2), em que houveram níveis baixos de expressão. Como esses genes não foram expressos no entrenó, foi feita uma análise de expressão em folha do H89 (Figura, B). Nessa análise em folha, ScAt10 e ScAt5 foram expressos. Novamente, houve contraste com as análises de RNA-seq de folha, que não detectaram expressão desses genes. Provavelmente, esse contraste ocorreu porque, embora ambas as análises tenham sido executadas em folha, os estágios de desenvolvimento da planta eram distintos (folha adulta e folha jovem). Dessa forma, nota-se que os níveis de expressão de ScAt10 e ScAt5 são baixos e detectáveis apenas em tecidos e estágios específicos. Esse padrão condiz com a suposta função de ScAt10 de incorporação de pCA à hemicelulose, já que é baixa a abundância desse tipo de ligação dos ácidos hidroxicinâmicos, predominando a ligação à lignina. Ademais, justamente nos tecidos jovens de folha, menos lignificados, que se espera que a maior parte do pCA esteja incorporado à hemicelulose (BARTLEY et al., 2013). A baixa expressão de ScAt5 também é condizente com a provável função de incorporação de FA éster à lignina, já que muito pouco FA-monolignol foi detectado em gramíneas (KARLEN et al., 2016).

Para melhor visualizar a tendência de expressão de cada gene nos diferentes tecidos e híbridos, os níveis relativos ao gene de referência foram representados no gráfico de dispersão (Figura 17). A análise estatística comparativa ao longo do desenvolvimento do entrenó, entre córtex e medula, e entre os híbridos foi representada (Tabela 11).

Tecido	ScAt1	ScAt2	ScAt4	ScAt6	ScAt7	ScAt8	ScAt9
H321EC1	2.136 ^{a,1,*}	7.860 ^{a,1,*}	1.318 ^{a,1}	4.327 ^{a,1}	4.037 ^{a,1,*}	5.381 ^{a,1,*}	1.586 ^{a,1,*}
H321EC5	1.530 ^{a,1}	7.262 ^{a,1}	2.871 ^{a,2}	5.693 ^{a,1}	6.210 ^{a,1}	7.008 a,2	2.247 ^{a,2}
H321EC9	2.478 a,1,*	7.043 ^{a,1}	5.599 ^{a,3,*}	6.487 ^{a,1}	7.001 ^{a,1}	7.016 ^{a,2}	4.556 ^{a,2}
H321EM1	4.191 ^{b,1,*}	8.278 ^{a,1}	2.902 a,1,*	8.050 ^{b,2}	6.400 ^{b,1,*}	7.969 ^{b,1,*}	4.229 ^{b,1,*}
H321EM5	3.968 ^{b,1}	8.853 ^{b,1}	4.453 ^{b,12}	6.397 ^{a,1}	6.561 ^{a,1}	7.709 ^{a,1}	4.554 ^{a,1}
H321EM9	3.533 ^{a,1,*}	7.257 ^{a,1}	5.507 a,2,*	9.765 ^{a,3,*}	6.879 ^{a,1,*}	7.520 ^{a,1,*}	4.576 ^{a,1}
H89EC1	0.943 ^{a,1,*}	2.938 a,1,*	-0.208 ^{a,1}	1.624 ^{a,1}	1.122 ^{a,1,*}	4.033 a,1,*	0.666 ^{a,1,*}
H89EC5	3.669 a,1	6.061 ^{a,1}	4.199 ^{a,1}	5.290 ^{a,1}	4.811 ^{a,2}	4.819 ^{a,1}	3.270 ^{a,1}
H89EC9	4.003 a,1,*	5.871 ^{a,1}	3.362 a,1,*	4.744 ^{a,1}	5.49 ^{a,2}	4.985 ^{a,1}	3.118 ^{a,1}
H89EM1	1.454 ^{a,1,*}	3.534 ^{a,1}	1.071 ^{a,1,*}	2.679 ^{a,1}	3.625 ^{b,1,*}	1.982 ^{b,1,*}	1.093 a,1,*
H89EM5	1.134 ^{a,1}	4.310 ^{a,1}	3.542 ^{a,2}	4.840 ^{a,1}	3.315 ^{a,1}	3.779 ^{a,1}	1.366 ^{a,1}
H89EM9	2.083 ^{b,1,*}	6.089 ^{a,1}	3.255 ^{a,2,*}	5.186 ^{a,1,*}	4.328 a,1,*	4.758 a,1,*	2.420 a,1

Tabela 11– Análise estatística dos dados de expressão dos genes ScAt1 a ScAt10 por RTqPCR em diferentes tecidos de colmo dos híbridos H321 e H89

Notas: Valores de ΔCt para os entrenós dos híbridos H89 e H321, separados em córtex e medula. EC1, EC5 e EC9 representam os valores do córtex dos entrenós 1,5 e 9 respectivamente.

EM1, EM5 e EM9 representam os valores da medula dos entrenós 1,5 e 9, respectivamente. Os híbridos foram indicados pelo prefixo H89 ou H321.

Valores com a mesma letra não foram significativamente distintos na comparação entre o córtex e medula do mesmo entrenó, do mesmo híbrido, e.g EC1 contra EM1 (teste T p<0.05).

Valores com o mesmo número não foram significativamente distintos na comparação entre córtex dos entrenós 1,5,9, ou medula 1,5,9 (ANOVA seguida de teste *post-hoc* de Tukey, p<0.05). Valores com asterisco foram significativamente distintos na comparação do mesmo tecido, entre os dois híbridos (teste T, p<0.05).

Fonte: Autoria própria.



Figura 17 – Análise da tendência de expressão por RT-qPCR dos genes *ScAt1* a *ScAt10* s ao longo do desenvolvimento do colmo dos híbridos de cana-de-açúcar H89 e H321

Legenda: Dados de expressão normalizados com o gene de referência (Valor de negativo ΔCt). Valores para os entrenós dos híbridos H89 e H321, separados em córtex e medula. EC1, EC5 e EC9 representam os valores do córtex dos entrenós 1,5 e 9 respectivamente. EM1, EM5 e EM9 representam os valores da medula dos entrenós 1,5 e 9, respectivamente. Os híbridos foram indicados pelo prefixo H89 ou H321. Fonte: Autoria Própria.

Observou-se que, em muitos casos, o padrão de expressão não foi significativamente distinto, a ponto de indicar uma óbvia correlação entre tecido ou híbrido (Figura 17, Tabela 11). De qualquer forma, vários genes apresentaram maior expressão no córtex do entrenó 1 (*At4, At6, At7, At8* e *At9*), mais jovem. Molinari et al., 2013 encontraram padrão similar em uma análise também por RT-qPCR em Brachypodium. O gene *BdAt4* foi mais expresso no entrenó mais jovem em plantas já na fase reprodutiva. Analisando a planta toda, os genes *BdAt4* e *BdAt9* também foram mais expressos no início da fase vegetativa, do que nos estágios mais avançados da fase vegetativa (MOLINARI et al., 2013). Em uma análise também em cana-de-açúcar, por RT-qPCR, Souza et al., 2019 encontraram maiores níveis de expressão para *ScAt9* em tecidos foliares jovens do que em maduros (DE SOUZA et al., 2019). Dessa forma, a expressão desses genes parece ser maior no início do desenvolvimento, sugerindo que tanto os genes com atividade na lignina como os com atividade na hemicelulose são expressos antes da lignificação.

Notou-se que o gene *ScAt1* foi mais expresso no córtex do que na medula do híbrido 321. Porém, no híbrido H89, não houveram diferenças entre córtex e medula. Ademais, os níveis na medula de H89 foram maiores do que na medula do H321. Isso pode significar que há diferença na forma como esses genes são regulados, e isso pode estar relacionado à diferença fenotípica entre os dois híbridos.

De uma forma geral, não foram observadas muitas diferenças significativas nos valores de expressão dos genes do clado de *Mitchell* entre H89 e H321. É possível que a diferença entre a composição de ácidos hidroxicinâmicos dos dois híbridos não tenha sido grande o suficiente a ponto de revelar padrões contrastantes na expressão gênica. Entretanto, alguns genes apresentaram mais expressão na medula e no córtex 1 do H89 do que no H321 (*ScAt1, ScAt7, ScAt8* e *ScAt9*). Isso não era esperado, considerando que o teor de ácidos hidroxicinâmicos é maior no H321. Porém, é possível que os genes sejam regulados de forma diferente nos dois híbridos, o que levou ao padrão observado. Ademais, outra possibilidade é que embora os genes do clado de Mitchell sejam mais expressos no H89, exista uma limitação de substrato para esses genes (FA e *p*CA), devido ao fato de que esse híbrido possui menor teor de lignina do que H321(16.8g/100g para o H89 contra 20.2 g/100g para o H321) (MASARIN et al., 2011). Assim, os genes da via de síntese dos monolignois podem ser menos expressos resultando em uma menor disponibilidade de ácidos ferúlico e *p*-cumárico.

De todos os genes no clado de Mitchell, *ScAt4* e *ScAt10* foram selecionados para posterior caracterização funcional em linhagens da planta modelo milho (*Zea mays*), devido ao seu potencial em aplicações para diminuição da recalcitrância da biomassa e porque os padrões de expressão detectados tanto nas análises *in silico* como por RT-qPCR condizem com o suposto papel na incorporação de *p*CA na lignina e na hemicelulose, respectivamente. Ademais, nenhum desses genes foi estudado em cana-de-açúcar.

5.4 CLONAGEM DOS GENES DE BIOSSÍNTESE DE pCA NOS VETORES DE SUPER-EXPRESSÃO (Ox) E SILENCIAMENTO GÊNICO (RNAi)

As sequências codificantes (CDS) de *ScAt10* e *ScAt4*, potencialmente envolvidas com a incorporação de *p*CA à hemicelulose e lignina, respectivamente, foram selecionadas para clonagem em vetores de super-expressão e silenciamento usando o sistema *gateway* (*Invitrogen*[®]), para posterior investigação do papel desses genes de cana-de-açúcar na parede celular.

Inicialmente, os produtos de PCR possuindo as extremidades recombinantes *attB* foram inseridos no plasmídeo *pDONR221*, gerando os vetores de entrada denominados *pDONR:RNAi:At10*, *pDONR:RNAi:At4*, *pDONR:Ox:ScAt10* e *pDONR:Ox:ScAt4*, (Anexo A).

A fim de confirmar os vetores clonados, foi feita digestão com enzima de restrição. Inicialmente, as sequências desses vetores foram analisadas no *software Snapgene* para identificação dos sítios de restrição e predição do padrão de bandas obtido para a digestão com as enzimas selecionadas. Esse padrão foi comparado com o observado após o ensaio de digestão dos vetores *pDONR:RNAi:ScAt10* e *pDONR:RNAi:ScAt4* com *PvuI* (Figura 18); *pDONR:Ox:ScAt10* com *KpnI* (Figura 19) e *pDONR:Ox:ScAt4* com dupla digestão *KpnI* e *PvuII* (Figura 20).

Figura 18 – Confirmação dos vetores de entrada *pDONR:RNAi:ScAt10* e *pDONR:RNAi:ScAt4* por digestão com a enzima de restrição *PvuI*



Legenda: Da esq. para dir: Padrão de bandas esperado pela análise das sequências no *software Snapgene*, resultado da digestão do vetor *pDONR221* e resultado da digestão dos vetores *pDONR:RNAi:ScAt10* e *pDONR:RNAi:ScAt4*. MW – Marcador de peso molecular, 1000kb (Neb[®]). A análise do padrão de bandas foi conduzida em gel de agarose 0,8% com corante *Unisafe dye*[®] (Uniscience). As imagens foram capturadas utilizando transiluminador.

Fonte: Autoria própria.

Figura 19 – Confirmação dos vetores de entrada *pDONR:Ox:Sc:At10* por digestão com enzima de restrição KpnI



Legenda:Da esq. para dir: Padrão de banda esperado pela análise das sequências no *software Snapgene*, resultado da digestão do vetor *pDONR221* e resultado da digestão dos vetores de duas colônias distintas (*pDONR:Ox:ScAt10 1 e 2*). MW – Marcador de peso molecular, 1kb (Neb[®]) e MW.1- Marcador de peso molecular, 1Kb (*Invitrogen[™]*). A análise do padrão de bandas foi conduzida em gel de agarose 0,8% com corante *Unisafe dye[®]* (*Uniscience*). As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria.





Legenda: Da esq. para dir: Padrão de bandas esperado pela análise das sequências no *software Snapgene*, resultado da dupla digestão do vetor *pDONR221* com *KpnI* e *PvuII* e resultado da digestão dos vetores provenientes de quatro colônias distintas (*pDONR:Ox:ScAt4* 1, 2). MW – Marcador de peso molecular, 1000kb (Neb[®]) A análise do padrão de bandas foi conduzida em gel de agarose 0,8% com corante brometo de etídeo. As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria.

Para os vetores *pDONR:RNAi:ScAt10* e *pDONR:RNAi:ScAt4*, o padrão obtido indicou que de fato houve inserção dos fragmentos de *RNAi*, com cerca de 160 pb, gerando um vetor menor que o *pDONR221*, com aproximadamente 2.760 pb (Figura 18). Observou-se também para os vetores *pDONR:Ox:ScAt10* e *pDONR:OxScAt4*, que o padrão de bandas correspondeu ao esperado, confirmando a clonagem (Figuras 19 e 20).

Adicionalmente, todos os vetores de entrada foram confirmados por sequenciamento.

Após a confirmação, os vetores foram recombinados com os vetores de destino, obtendo as construções finais *pzp221b:RNAi:At4*, *pzp221b:RNAi:At10*, *pzp221b:Ox:ScAt10* e *pzp221b:Ox:ScAt4* (Anexo B). Para confirmar a inserção das sequências, os vetores *pzp221b:RNAi:At4*, *pzp221b:RNAi:At10* e *pzp221b:Ox:ScAt10* foram inicialmente submetidos à digestão com *KpnI* (Figuras 21 e 22). O vetor *pzp221b:Ox:ScAt4* foi digerido com *PvuII* (Figura 23) Figura 21 - Confirmação dos vetores de destino *pzp221b:RNAi:ScAt10* e *pzp221b:RNAi:ScAt4* por digestão com a enzima de restrição *KpnI*



Legenda: Da esq. para dir: Padrão de bandas esperado pela análise das sequências no *software Snapgene*, resultado da digestão do vetor *pzp221b:RNAi* e resultado da digestão dos vetores provenientes de duas colônias distintas (*pDONR:RNAi:ScAt10* 1 e 2, *pDONR:RNAi:ScAt4* 1 e 2). MW – Marcador de peso molecular, 1000 kb (Neb[®]). A análise do padrão de bandas foi conduzida em gel de agarose 0,8% com corante *Unisafe dye[®]* (*Uniscience*). As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria.

Figura 22 - Confirmação dos vetores de destino *pzp221b:Ox:ScAt10* por digestão com a enzima de restrição *KpnI*



Legenda: Da esq. para dir: Padrão de bandas esperado pela análise das sequências no *software Snapgene*, resultado da digestão do vetor *pzp221b:Ox* e resultado da digestão dos vetores provenientes de duas colônias distintas (*pDONR:Ox:ScAt10* 1 e 2). MW – Marcador de peso molecular, 1000kb (Neb®). A análise do padrão de bandas foi conduzida em gel de agarose 0,8% com corante *Unisafe dye*® (*Uniscience*). As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria.



Figura 23 – Confirmação dos vetores de destino *pzp221b:Ox:ScAt4* por digestão com a enzima de restrição *PvuII*

Legenda: Da esq. para dir: Padrão de bandas esperado pela análise das sequências no *software Snapgene*, resultado da dupla digestão do vetor *pDONR221* com *KpnI* e *PvuII* e resultado da digestão dos vetores provenientes de quatro colônias distintas (*pDONR:Ox:ScAt4* 1, 2). MW – Marcador de peso molecular, 1000kb (Neb[®]) A análise do padrão de bandas foi conduzida em gel de agarose 0,8% com corante brometo de etídeo. As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria.

Para a digestão com *KpnI* e *PvuII* o padrão de bandas correspondeu exatamente ao esperado para todos os vetores analisados, confirmando a inserção dos fragmentos. Para os vetores de silenciamento, o padrão de bandas confirma que as sequências *RNAi:At4* e *RNAi:At10* foram inseridas nos dois pontos de recombinação, o que é fundamental para o posterior efeito de silenciamento do gene na planta.

Adicionalmente, todos os vetores também foram confirmados por digestão com *BamHI* (Figuras 24, 25 e 26). Para as construções de silenciamento, o padrão de bandas após digestão com *BamHI* do vetor vazio *pzp221b:RNAi* também foi comparado com o obtido experimentalmente pelo colaborador PhD. David Rancour (*USDA*) (MARITA et al., 2014), um dos pesquisadores que participou da construção dos vetores *pzp221b RNAi* e *Ox* utilizados neste trabalho (Figura 24).

Figura 24 - Confirmação dos vetores de destino *pzp221b:RNAi:ScAt10* e *pzp221b:RNAi:ScAt4* por digestão com a enzima de restrição *BamHI*



Legenda: Da esq. para dir: Padrão de bandas esperado pela análise das sequências no *software Snapgene*, resultado da digestão do vetor Pzp221b:RNAi realizada pelo colaborador Dr. David Rancour, resultado da digestão do vetor pzp221b:RNAi deste trabalho, e resultado da digestão dos vetores provenientes de duas colônias distintas (*pDONR:RNAi:ScAt10* 1 e 2, *pDONR:RNAi:ScAt4* 1 e 2). MW – Marcador de peso molecular, 1000kb (Neb®). A análise do padrão de bandas foi conduzida em gel de agarose 0,8% com corante *Unisafe dye*® (*Uniscience*). As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria e colaboradores.

Figura 25 - Confirmação dos vetores de destino *pzp221b:Ox:ScAt10* por digestão com a enzima de restrição *BamHI*



Legenda: Da esq. para dir: Padrão de bandas esperado pela análise das sequências no *software Snapgene*, resultado da digestão do vetor *pzp221b:Ox* e resultado da digestão dos vetores provenientes de duas colônias distintas (*pDONR:Ox:ScAt10* 1 e 2). MW – Marcador de peso molecular, 1000 kb (Neb[®]). A análise do padrão de bandas foi conduzida em gel de agarose 0,8% com corante *Unisafe dye[®]* (*Uniscience*). As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria.

Figura 26 - Confirmação dos vetores de destino *pzp221b:Ox:ScAt4* por digestão com a enzima de restrição *BamHI*



Legenda: Da esq. para dir: Padrão de bandas esperado pela análise das sequências no *software Snapgene*, resultado da digestão do vetor *pzp221b:Ox* e resultado da digestão dos vetores provenientes de duas colônias distintas (*pzp221b:Ox:ScAt4* 1 e 2). MW – Marcador de peso molecular, 1 kb (Neb[®]). A análise do padrão de bandas foi conduzida em gel de agarose 0.8% corado com brometo de etídeo. As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria.

Para o vetor pzp221b:RNAi, notou-se que o padrão diferiu do esperado pela análise da sequência no *snapgene*. Uma banda que deveria possuir entre 1,5 e 2,0 kb (1.780 pb), apresentou tamanho entre 1,0 e 1,5 kb. Entretanto, no resultado obtido pelo colaborador PhD David Rancour, notou-se que o mesmo padrão deste trabalho foi encontrado (Figura 24). Provavelmente, existe alguma diferença na sequência do vetor nessa região, entre o arquivo que foi fornecido pelo colaborador e a sequência real do vetor. O fato de este trabalho ter obtido o mesmo resultado que o do PhD David Rancour demonstra que o vetor pzp221bRNAi utilizado corresponde corretamente ao que foi enviado pelos colaboradores. Analisando o mapa desses plasmídeos, nota-se que essa sequência com tamanho distinto está na parte que foi recombinada, de forma que isso não influenciou na análise do padrão obtido para a digestão dos vetores finais clonados. Os fragmentos obtidos após a digestão de pzp221b:RNAi:At4 e pzp221b:RNAi:At10 confirmam a inserção das sequências nos plasmídeos, suportando que a recombinação ocorreu corretamente nos dois pontos necessários. Já o padrão dos vetores pzp221b:Ox, pzp221b:Ox:ScAt10 e pzp221b:Ox:ScAt4 correspondeu exatamente ao esperado, confirmando a inserção da sequência (Figuras 25 e 26).

Todas as sequências dos vetores de destino foram confirmadas por sequenciamento.

5.5 TRANSFORMAÇÃO DAS CONSTRUÇÕES DE SUPER-EXPRESSÃO E SILENCIAMENTO EM LINHAGENS DA PLANTA MODELO ZEA MAYS

Os vetores de clonagem de super-expressão e silenciamento foram enviados para *Plant Transformation Facillity* (PTF) em *Iowa State University* (IOWA STATE UNIVERSITY, 2019), para transformação na planta modelo *Zea mays* (milho). As linhagens provenientes dos eventos de transformação com as construções de super-expressão de *ScAt10* e *ScAt4* foram denominadas A871 e A885, respectivamente. Já as linhagens de silenciamento de *ScAt10* e *ScAt4* foram denominadas A872 e A873, respectivamente. Foram obtidos 12 eventos de transformação independentes para A871 e A873, 11 eventos para A872 e 20 eventos para A885.

5.6 CONFIRMAÇÃO DA PRESENÇA DO INSERTO POR GENOTIPAGEM DAS LINHAGENS DE MILHO TRANSGÊNICAS

Conduziu-se a genotipagem por PCR das linhagens de milho, de maneira a verificar a presença do inserto nos transgênicos (Figura 27). Cada linhagem proveniente de um evento independente de transformação recebeu um nome distinto, indicado na Figura 27.

Figura 27 - Genotipagem das linhagens A871 (*Ox:ScAt10*), A872 (*Ox:ScAt4*) A873 (*At4:RNAi*) e A885 (*Ox:ScAt4*) por PCR



Legenda: A – Análise dos eventos A871 (*Ox:ScAt10*) B - Análise dos eventos A872 (*ScAt10:RNAi*), exceto A872 4b. C – Análise dos eventos A873 (*ScAt4:RNAi*), exceto A873 1C. D – Análise dos eventos A885 (*Ox:ScAt4*), exceto A885 11A. MM – Marcador de peso molecular, 100 pb (Neb[®]). A análise do foi conduzida em gel de agarose 3% corado com brometo de etídeo. As imagens foram capturadas utilizando transiluminador. Fonte: Autoria própria.

A maior parte dos transgênicos foi confirmado quanto à presença do inserto, exceto as linhagens A872-4B, A873-1C e A885-11A (Figura 25, B, C e D), as quais foram consideradas "escapes" e utilizadas como controle nas análises dos transgênicos. Também foi verificado que os iniciadores foram de fato específicos para as sequências de cana-de-açúcar, uma vez que não houve amplificação a partir do DNA genômico do milho selvagem B73, usado como controle.

5.7 ANÁLISE DA EXPRESSÃO GÊNICA DAS LINHAGENS DE MILHO TRANSGÊNICAS POR RT-qPCR

A análise de expressão de todas as linhagens transgênicas (A872, A873, A871 e A885), foi feita para a base da primeira folha completamente desenvolvida acima do órgão reprodutor feminino (orelha), em uma tentativa de utilizar tecidos aproximadamente no mesmo estágio desenvolvimento, permitindo a comparação entre os níveis de expressão das amostras transgênicas de cada construção e selvagem.

Para as linhagens transgênicas de silenciamento A872 (*ScAt10:RNAi*) e A873 (*ScAt4:RNAi*) foram verificados os níveis de expressão dos genes endógenos *ZmAt10* e

ZmAt4, respectivamente, por PCR em tempo real (RT-qPCR), a fim de observar se de fato houve silenciamento desses genes pelo *RNAi* como esperado.

Na análise de expressão de ZmAt10 nas linhagens A872 (*RNAi:ScAt10*), não foi possível detectar expressão nos eventos transgênicos, nem nos controles "escape"A872-4B e selvagem B73. Assim, nesse estágio foliar (após o desenvolvimento dos órgãos reprodutivos), observou-se que o nível de expressão de ZmAt10 era muito baixo para o selvagem, de forma que não foi possível detectar diferenças de expressão decorrentes do silenciamento nas plantas transgênicas A872. No entanto, foi levantada a hipótese de que o gene fosse mais expresso em fases de desenvolvimento mais jovens, explicando os níveis de *p*CA encontrados na hemicelulose de folha de milho, ainda que em pequenas quantidades (LAPIERRE et al., 2018).

Para verificar em quais tecidos ou estágios de desenvolvimento ZmAt10 era expresso, inicialmente foi feita uma análise utilizando a ferramenta Maize efp browser do site (The *Bio-Analytic* for Plant Bar utoronto Resource Biology, http://bar.utoronto.ca/efp_maize/cgi-bin/efpWeb.cgi). Essa ferramenta permite a visualização dos níveis de expressão gênica em diferentes tecidos de milho B73, baseados nos dados de Atlas transcriptômicos já publicados. Utilizando o atlas de Hoopes et al. (HOOPES et al., 2019), observou-se que ZmAt10 tem maiores níveis de expressão na ponta da folha imatura, no estágio de desenvolvimento V9 (Fase vegetativa com um total de 9 folhas completamente desenvolvidas com colar a partir da primeira folha verdadeira, Figura). Alguns outros estágios foliares também apresentaram expressão de ZmAt10, ainda que em menores níveis (Figura 28).

Figura 28 - Análise dos dados de expressão de milho B73 provenientes do atlas de RNAseq



Legenda: Análise conduzida utilizando a ferramenta *Maize eFP Browser*, do site *BAR utoronto*, utilizando *ZmAt10* como sonda para busca no atlas (Hoopes et al., 2018). Valores de expressão em FPKM (Fragmentos Por kilobase de transcrito por Milhão de *reads* mapeados). Folha imatura corresponde à ponta da folha. V3,V9 e V18 são os estágios de desenvolvimento em que 3, 9 e 18 folhas estavam com colar completamente desenvolvido, respectivamente. VT corresponde ao estágio de transição do crescimento vegetativo para o reprodutivo. Fonte: Autoria própria.

Após essa análise inicial *in silico*, foi feita também uma análise da expressão de *ZmAt10* em milho selvagem por PCR em tempo real, utilizando a ponta de folhas na fase vegetativa (folhas 9, 10 e 12, numeradas da base para o topo, estágio V9). Foi encontrada expressão na folha 9 (Figura 29), similar ao encontrado para análise *in silico*. No entanto, esperava-se detectar alguma expressão na folha 11 por exemplo, ainda que em níveis mais baixos.

Figura 29 – Expressão Relativa de *ZmAt10* (RT-qPCR) nas amostras da nona folha de dois indivíduos de uma linhagem de milho selvagem



Legenda: Expressão relativa calculada pelo método 2[^]-ΔΔCt, utilizando os Ct corrigidos pela eficiência dos iniciadores, amostra de referência Fl 9.1. Valores representam a média e os limites superiores e inferiores determinados pelo desvio padrão (n=3). Fonte: Autoria própria.

Assim, o gene *ZmAt10* parece ter baixos níveis de expressão e em estágios muito específicos do desenvolvimento, o que dificultou a detecção dos efeitos do silenciamento nas linhagens *ScAt10:RNAi*.

Já para as linhagens A873 (*ScAt4:RNAi*), a maior parte dos eventos apresentaram diminuição significativa (p<0.05) na expressão de *ZmAt4* quando comparada com o controle representativo, a linhagem escape A872-4B (Figura 30).

Figura 30 – Expressão Relativa de *ZmAt10* (RT-qPCR) nas linhagens de milho transgênicas A873 (*ScAt4:RNAi*) e nos controles linhagens escape e selvagem



Legenda: Expressão relativa na base foliar acima do órgão reprodutor feminino. Valores calculados pelo método $\Delta\Delta$ Ct incorporando a eficiência dos iniciadores, usando como amostra de referência A873-3B. Linhagens transgênicas foram representadas em vermelho (A873-). Os controles foram representados em azul, linhagens "escape" A872 4B e selvagem B73. Valores representam a média e os limites superiores e inferiores determinados pelo desvio padrão (n=3). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Apenas a linhagem A873-3B não apresentou redução significativa. Esses resultados demonstram a eficiência da construção de silenciamento por *RNAi*, principalmente para as linhagens A873-5C, A873-17B e A873-4C, as quais tiveram mais de 7 vezes de redução na expressão comparadas com o controle escape A872-4B. A maior

diminuição ocorreu para as linhagens A873-5C, as quais tiveram nível de expressão 16 vezes menor do que o controle A872-4B.

Para as construções de super-expressão A871 (*Ox:ScAt10*) e A885 (*Ox:ScAt4*), foram verificados os níveis de expressão dos genes inseridos *ScAt10* e *ScAt4*, bem como dos endógenos de milho *ZmAt10* e *ZmAt4*, respectivamente.

Para as linhagens A871, detectou-se expressão de *ScAt10* em todos os eventos (Figura 31). De todas as linhagens, A871-14B apresentou o maior nível de expressão e A871-2A o menor, sendo a diferença bastante expressiva (30 x maior para A871-14B). Isso mostra que a eficiência da construção de super-expressão com o promotor Ubiquitina de milho foi variável entre os eventos independentes, embora todos eles estejam expressando *ScAt10*. Os iniciadores foram específicos para o cDNA de cana-de-açúcar, não havendo amplificação nos controles linhagem escape A872-4B e selvagem B73.

Figura 31 – Expressão Relativa de *ScAt10* (RT-qPCR) nas linhagens de milho transgênicas A871 (*Ox:ScAt10*) e nos controles linhagens escape e selvagem



Legenda: Expressão relativa na base foliar acima do órgão reprodutor feminino. Valores calculados pelo método $\Delta\Delta$ Ct, incorporando a eficiência dos iniciadores, usando como amostra de referência A871-2A. Linhagens transgênicas foram representadas em vermelho (A871-). Não foi detectada amplificação para as linhagens "escape" A872 4B e selvagem B73, as quais foram representadas como "zero". Valores representam a média e os limites superiores e inferiores determinados pelo desvio padrão (n=3). Fonte: Autoria própria.
Nenhuma amplificação foi detectada para o endógeno *ZmAt10* nas linhagens A871, corroborando com o que foi encontrado para as linhagens selvagens e A872. Dessa forma, observou-se que há super-expressão de *ScAt10* com pouca influência do endógeno *ZmAt10*, uma vez que este último possui níveis de expressão muito baixos na planta selvagem.

Para as linhagens A885 o gene *ScAt4* também foi expresso em todos os eventos analisados, em níveis diferentes (Figura 32). Os maiores valores foram para A885-5C e 16A, cerca de 16 e 10 vezes mais expressos que A885-9B, evento com o menor nível de expressão. Entretanto, os níveis de expressão do endógeno *ZmAt4* também foram variáveis entre os eventos (Figura 33), com A885-8B e A885-9B possuindo os maiores níveis, 12 e 10 vezes mais que a amostra menos expressa A885-16A.

Figura 32– Expressão Relativa de *ScAt4* (RT-qPCR) em folha nas linhagens de milho transgênicas A885 (*Ox:ScAt4*) e nos controles linhagens escape e selvagem



Legenda: Expressão relativa na base foliar acima do órgão reprodutor feminino. Valores calculados pelo método $\Delta\Delta$ Ct, incorporando a eficiência dos iniciadores, usando como amostra de referência A885 9B. Linhagens transgênicas foram representadas em vermelho (A885-). Os controles foram representados em azul, linhagens "escape" A872 4B e selvagem B73. Valores representam a média e os limites superiores e inferiores determinados pelo desvio padrão (n=3). Fonte: Autoria própria.



Figura 33– Expressão Relativa de *ZmAt4* (RT-qPCR) nas linhagens de milho transgênicas A885 (*Ox:ScAt4*) e nos controles linhagens escape e selvagem

Legenda: Expressão relativa na base foliar acima do órgão reprodutor feminino. Valores calculados pelo método $\Delta\Delta$ Ct, incorporando a eficiência dos iniciadores, usando como amostra de referência A885 9B. Linhagens transgênicas foram representadas em vermelho (A885-). Os controles foram representados em azul, linhagens "escape" A872 4B e selvagem B73. Valores representam a média e os limites superiores e inferiores determinados pelo desvio padrão (n=3). Fonte: Autoria própria.

Como todos os eventos A885 analisados apresentaram expressão de *ScAt4*, todos "super- expressam" esse gene de cana-de-açúcar, ausente no selvagem B73. Entretanto, se ambas as enzimas *ScAt4* e *ZmAt4* possuírem a mesma função, é possível que os níveis de expressão de *ScAt4* em determinados eventos não sejam suficientes para aumentar significativamente a atividade At4 (*p*-coumaril monolignol transferase) na planta com relação à já existente naturalmente, promovida pelo endógeno *ZmAt4*. Dessa forma, a fim de se comparar os níveis de expressão de *ScAt4* e *ZmAt4* em cada um dos eventos A885, os valores de Δ Ct foram representados em um *heatmap* (Figura 34)



Figura 34 – Análise de expressão por RT-qPCR dos genes *ScAt4* e *ZmAt4* nos eventos A885 (*Ox:ScAt10*)

Legenda: Expressão na base foliar acima do órgão reprodutor feminino. Valores normalizados pela média geométrica dos genes endógenos (Δ Ct), utilizando os Ct corrigidos pela eficiência dos iniciadores. A885-representam as linhagens transgênicas, enquanto B73 e A872 4B são os controles linhagens selvagem B73 e "escape", respectivamente. Eventos agrupados hierarquicamente pelo método euclidiano e distância completa. Fonte: Autoria própria.

Embora os valores de ΔCt não sejam quantitativamente comparáveis entre genes diferentes, dadas as diferenças nos iniciadores e nos alvos (tamanho, conteúdo GC) que influenciam a cinética da reação de PCR, a análise pode ser semi-quantitativa, a fim apenas de indicar um padrão de expressão. No caso das linhagens A885, nota-se uma tendência clara de maior expressão de *ScAt4* do que *ZmAt4* principalmente nos eventos A885-16A, A885-5C, A885-8B.

5.8 ANÁLISE FENOTÍPICA DA COMPOSIÇÃO QUÍMICA DA PAREDE CELULAR DAS LINHAGENS DE MILHO TRANSGÊNICAS

5.8.1 Análise do impacto do silenciamento ou super-expressão no conteúdo de ácidos hidroxicinâmicos das linhagens transgênicas

Com o objetivo de verificar o impacto da alteração da expressão gênica de *At10* e *At4* na composição de ácidos hidroxicinâmicos, o conteúdo total de ácidos ferúlico (FA) e *p*-cumárico (*p*CA) da parede celular liberado por tratamento alcalino brando foi determinado para o colmo maduro de todos os eventos (A872, A873, A871 e A885), buscando-se identificar as linhagens com alterações significativas. Tendo em vista que tanto as ligações de FA à arabinose como as de pCA à arabinose e lignina são do tipo éster, nesta análise inicial não foi feita a quantificação dos fenólicos liberados por tratamento severo. Nas análises de parede celular de milho, o tratamento brando é suficiente para quebrar a maior parte das ligações ésteres (GRABBER et al., 1995).

Nas linhagens A872 (*At10:RNAi*), não foi observada diferença significativa entre os níveis de *p*CA(Figura 35) ou FA (Figura 36) nos transgênicos com relação aos controles linhagens escape A872-4B, A873-1C e os selvagens B73. Esse resultado corrobora com o encontrado para a expressão gênica de *ZmAt10*. Como os níveis de expressão desse gene são muito baixos no milho selvagem, o silenciamento não foi suficiente para provocar alterações significativas no conteúdo de ácidos hidroxicinâmicos.

Figura 35– Quantidade de ácido *p*-cumárico liberada pelo tratamento brando no colmo das linhagens A872 (*At10:RNAi*) e controle linhagens escape e selvagem B73



Legenda: As linhagens transgênicas foram representadas em vermelho (A872-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria



Figura 36– Quantidade de ácido ferúlico liberada pelo tratamento brando no colmo das linhagens A872 (*At10:RNAi*) e controle linhagens escape e selvagem B73

Legenda: As linhagens transgênicas foram representadas em vermelho (A872-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Para as linhagens A873 (*At4:RNAi*), apenas o evento A873-5C apresentou redução significativa (p<0.001) no conteúdo de *p*CA éster liberado pelo tratamento brando (Figura 37). Esta linhagem também foi a que apresentou menor nível de expressão do gene *ZmAt4* (16 x, Figura 30), sugerindo uma correlação entre o nível de expressão gênica de *ZmAt4* e o conteúdo de ácido *p*-cumárico. Tal correlação suporta a função de *ZmAt4* relacionada à incorporação de *p*CA. No entanto, outras linhagens que também tiveram redução significativa da expressão gênica, ainda que em menor nível do que A873-5C (Figura 30), não apresentaram efeito significativo no conteúdo de *p*CA éster. É possível que a redução na expressão nesses eventos não tenha sido suficiente para suprimir a atividade de *At4* a ponto de reduzir os níveis de *p*CA, sendo necessária uma redução mais expressiva como a de A873-5C.



Figura 37– Quantidade de ácido *p*-cumárico liberada pelo tratamento brando no colmo das linhagens A873 (*At4:RNAi*) e controle linhagens escape e selvagem B73

Legenda: As linhagens transgênicas foram representadas em vermelho (A873-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Algumas linhagens A873 apresentaram redução no conteúdo de FA éster liberado pelo tratamento alcalino brando (Figura 38), embora menos significativa do que a redução de *p*CA observada para A873-5C (p<0.05 para A873-17B e A873-3B, p<0.01 para A873-5C). Para A873-5C, com menor expressão de *ZmAt4* e menor nível de *p*CA, a redução de FA foi de aproximadamente 2,5x comparada com a média dos controles. Embora exista a possibilidade de que tal diminuição no teor de FA éster esteja relacionada com a expressão reduzida desse gene, mais estudos seriam necessários para revelar se de fato trata-se de uma resposta direta ao silenciamento. Resultado similar foi encontrado por Marita *et al.* (2014), ao também silenciarem *ZmAt4* por *RNAi* em milho, em que alguns eventos apresentaram alteração no conteúdo de FA, de redução ou aumento, entretanto pouco expressiva, não sendo possível afirmar relação com o silenciamento de *ZmAt4*.



Figura 38– Quantidade de ácido ferúlico liberada pelo tratamento brando no colmo das linhagens A873 (*At4:RNAi*) e controle linhagens escape e selvagem B73

Legenda: As linhagens transgênicas foram representadas em vermelho (A873-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Na análise dos eventos A871 (Ox:ScAt10), as linhagens A871-11B e A871-4A apresentaram aumento significativo no conteúdo de pCA éster liberado pelo tratamento alcalino brando (Figura 39). A871-11B foi o evento que apresentou maiores níveis de expressão de ScAt10 (Figura 31), indicando relação entre a expressão gênica e a quantidade de pCA. Esse aumento suporta a relação de ScAt10 com os níveis de pCA na parede celular.



Figura 39– Quantidade de ácido *p*-cumárico liberada pelo tratamento brando no colmo das linhagens A871 (Ox:Sc*At10*) e controle linhagens escape e selvagem B73

Legenda: As linhagens transgênicas foram representadas em vermelho (A871-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Todos os eventos A871, com exceção de A871-7A, apresentaram redução significativa no conteúdo de ácido ferúlico (p<0.05, Figura 40), mais expressiva para A871- 11B, A871-14B, A871-2A e A871-6A (p<0.001). Este resultado corrobora com a hipótese de que *ScAt10* estaria envolvido com a incorporação de *p*CA à hemicelulose, uma vez que efeito similar foi obtido por Bartley *et al.* (2013) por meio da super-expressão de *OsAt10* em linhagens de arroz (BARTLEY et al., 2013). Efeito similar também foi obtido para a super-expressão de *OsAt10* na gramínea *Switchgrass*, que levou ao aumento de *p*CA e redução de FA (~28%) nas folhas verdes (LI et al., 2018). Entretanto, diferentemente do presente trabalho e de *Bartley et al.*, (2013), o conteúdo de *p*CA sofreu leve diminuição nas folhas senescentes (~15%), indicando algum efeito em *Switchgrass* que não foi verificado arroz e nem para a expressão de *ScAt10* no colmo senescente de milho.



Figura 40– Quantidade de ácido ferúlico liberada pelo tratamento brando no colmo das linhagens A871 (Ox:Sc*At10*) e controle linhagens escape e selvagem B73

Legenda: As linhagens transgênicas foram representadas em vermelho (A871-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Com o objetivo de melhor visualizar o impacto da super-expressão de *ScAt10* tanto no *p*CA como no FA, a razão *p*CA/FA foi representada (Figura 41). Os eventos A871- 11B, A871-14B, A871-2A e A871-6A apresentaram aumento mais significativo na razão pCA/FA (p<0.001). O evento com o maior aumento na razão *p*CA/FA foi A871-11B, com maior nível de expressão de *ScAt10*. No entanto, outros eventos com menor nível de expressão de *ScAt10* também apresentaram diferenças significativas, indicando que estes níveis foram suficientes para provocar efeitos significativos na parede celular.



Figura 41– Razão entre as quantidades *p*CA/FA liberadas pelo tratamento brando no colmo das linhagens A871 (Ox:Sc*At10*) e controle linhagens escape e selvagem B73

Legenda: As linhagens transgênicas foram representadas em vermelho (A871 -) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Para os eventos A885 (Ox:ScAt4), optou-se por fazer uma análise preliminar, sem as replicatas necessárias para as análises estatísticas, apenas para buscar eventos com grande alteração no conteúdo de ácidos hidroxicinâmicos. No entanto, tais análises preliminares não revelaram diferenças significativas no conteúdo de pCA (Figura 42) ou FA (Figura 43) entre as linhagens transgênicas e os controles linhagens escape A872-4B, A873-1C e selvagem B73.



Figura 42– Quantidade de ácido *p*-cumárico liberada pelo tratamento brando no colmo das linhagens A885 (Ox:Sc*At4*) e controle linhagens escape e selvagem B73

Legenda: As linhagens transgênicas foram representadas em vermelho (A885-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Fonte: Autoria própria.



Figura 43– Quantidade de ácido ferúlico liberada pelo tratamento brando para as linhagens A885 (Ox:Sc*At4*) e controle linhagens escape e selvagem (B73)

Legenda: As linhagens transgênicas foram representadas em vermelho (A885-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Fonte: Autoria própria.

Considerando a atividade proposta para a enzima *ScAt4*, de *p*-coumaril-CoA monolignol transferase, seria esperado um aumento expressivo na quantidade de *p*CA éster nas linhagens transgênicas com super expressão desse gene, devido a um aumento na incorporação de pCA à lignina. Tal resultado foi encontrado por Petrik *et al.* (2014) para a super-expressão do gene de Brachypodium *BdAt4* nesta espécie (PETRIK et al., 2014). Mesmo o evento A885-16A, com maior nível de expressão de *ScAt4*, apresentou apenas 19 g/kg de ácido *p*-cumárico liberado no tratamento alcalino brando, valor muito próximo do encontrado para os controles e do descrito para milho na literatura (HATFIELD et al., 2009; MARITA et al., 2014).

É possível que ScAt4 tenha função distinta de ZmAt4 em cana-de-açúcar, porém a similaridade das enzimas, o fato de um fragmento do gene de cana ter sido capaz de silenciar ZmAt4, bem como o de que os homólogos em outras gramíneas como Brachypodium e arroz apresentaram atividade similar sugerem função relacionada para esse grupo de ortólogos At4. Outra possibilidade é que a enzima de cana-de-açúcar não tenha apresentado atividade em milho, por ser de outra espécie e apresentar algum padrão distinto estrutural ou de enovelamento. Porém, a proximidade evolutiva das espécies sugere que a atuação da enzima em milho seria possível. Uma hipótese seria de que o nível de expressão de ScAt4 com relação ao do endógeno ZmAt4 não tenha sido suficiente para provocar aumentos significativos no teor de pCA éster, ainda que os níveis encontrados pela análise RT-qPCR tenham sido altos. No entanto, uma provável explicação é que exista um fator limitante para a atividade enzimática. As plantas transgênicas podem apresentar um excesso da enzima At4, porém algum fator limitante como a disponibilidade de substratos pCA ou monolignol podem estar limitando a atividade, fazendo com que a quantidade de produto final monolignol-pCA não seja significativamente diferente do selvagem. Estudos posteriores seriam necessários para avaliar essas hipóteses, e verificar se de fato a função de ScAt4 é a mesma de ZmAt4.

As análises do conteúdo total de ácidos hidroxicinâmicos liberada pelo tratamento alcalino brando revelaram impacto na parede celular apenas para as linhagens A871 (Ox:ScAt10) e A873 (At4:RNAi). No entanto, tais análises não são suficientes para determinar se as alterações ocorreram na hemicelulose ou na lignina. Assim, com esses resultados não é possível verificar a hipótese de participação de ScAt10 na incorporação de pCA à hemicelulose. Dessa forma, foram feitos ensaios específicos para verificar o conteúdo de ácidos hidroxicinâmicos em cada uma das frações (hemicelulose ou lignina).

Dada a ausência de diferença significativa os eventos A885 não foram incluídos nas análises posteriores.

5.8.2 Análise do impacto do silenciamento ou super-expressão gênica no conteúdo de ácidos hidroxicinâmicos especificamente unido à hemicelulose (GAX)

Para a análise do conteúdo de ácidos hidroxicinâmicos unidos à Glucuronoarabinoxilana (GAX), foram selecionados os eventos com diferença significativa com relação ao selvagem, a partir da análise do conteúdo de pCA e FA éster liberados pelo tratamento alcalino brando (seção 5.7.1).

Embora os eventos A872 (At10:RNAi) não tenham apresentado diferença significativa no teor de pCA ou FA éster, foi feita a determinação do teor exclusivamente unido a hemicelulose para os eventos A872-14A, A872-18B e A872-2B, apenas para verificar se alguma diferença era detectável nesta análise mais específica. Entretanto, confirmando as análises anteriores, não foi detectada diferença no conteúdo de pCA-Arabinose (pCA-Ara) entre as linhagens transgênicas A872 e os controles selvagem e escape (Figura 44). De fato, tanto nos transgênicos como nos controles, o conteúdo de pCA-Ara foi muito baixo ou quase indetectável. Esse resultado corrobora com os baixos níveis de pCA unidos à GAX descritos na literatura para milho, principalmente em colmo (HATFIELD et al., 2009; LAPIERRE et al., 2018), estando a grande maioria do pCA da parede celular unido à lignina.

Pouco efeito também foi verificado para o conteúdo de FA exclusivamente ligado à hemicelulose (Figura 45) Figura 44– Quantidade de ácido *p*-cumárico especificamente unido à hemicelulose, liberado pela hidrólise ácida branda no colmo das linhagens A872 (*At10:RNAi*) e controle linhagens escape e selvagem B73



Legenda: As linhagens transgênicas foram representadas em vermelho (A872-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Figura 45– Quantidade de ácido ferúlico especificamente unida à hemicelulose, liberado pela hidrólise ácida branda no colmo das linhagens A872 (*At10:RNAi*) e controle linhagens escape e selvagem B73



Legenda: As linhagens transgênicas foram representadas em vermelho (A872-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Para as linhagens A873 (*ScAt4:RNAi*), foram selecionados para esta análise o melhor evento A873-5C com menor conteúdo de pCA éster, bem como A873-13B, A873-3B e A873-4C, que tiveram um pouco menos pCA éster que os demais, embora não significativo. O conteúdo de pCA-Ara (Figura 46) e FA-Ara (Figura 47) obtido nessas linhagens foi muito parecido com o dos controles selvagem e escape, extremamente baixo. De fato, o silenciamento de *At4* não interferiu no conteúdo de pCA da hemicelulose.

Figura 46– Quantidade de ácido *p*-cumárico especificamente unido à hemicelulose, liberado pela hidrólise ácida branda, no colmo das linhagens A873 (*At14:RNAi*) e controle linhagens escape e selvagem B73



Legenda: As linhagens transgênicas foram representadas em vermelho (A873-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Figura 47– Quantidade de ácido ferúlico especificamente unido à hemicelulose, liberado pela hidrólise ácida branda, no colmo das linhagens A873 (*At14:RNAi*) e controle linhagens escape e selvagem B73



Legenda: As linhagens transgênicas foram representadas em vermelho (A873-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Já para a análise de A871(Ox:ScAt10) foram selecionados os eventos com diferença significativa na razão pCA/FA éster (p<0.001): A871-11B, A871-14B, A871-2A e A871-6A (Seção 5.7.1). A diferença entre o conteúdo de pCA-Ara entre as linhagens A871 e os controles selvagem e escape foi muito significativa (p<0.001) (Figura 48). Todos os eventos analisados apresentaram níveis de pCA-Ara muito maiores, composto praticamente não detectado nas linhagens controle, ou ainda A872 e A873. É notável que a super-expressão de ScAt10 resultou na incorporação de pCA à arabinose nos transgênicos, atividade praticamente inexistente no selvagem. Figura 48– Quantidade de ácido *p*-cumárico especificamente unido à hemicelulose, liberado pela hidrólise ácida branda, no colmo das linhagens A871 (*Ox:ScAt10*) e controle linhagens escape e selvagem B73



Legenda: As linhagens transgênicas foram representadas em vermelho (A871-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

A super-expressão de At10 também teve efeito significativo no teor de FA-Ara, com todas as linhagens apresentando expressiva redução (p<0.01) (Figura 49). Em arroz, a super-expressão do gene OsAt10 resultou em efeitos similares no conteúdo de pCA e FA incorporados especificamente à GAX (BARTLEY et al., 2013), sendo mais um indicativo que o gene identificado em cana-de-açúcar ScAt10 possui de fato função similar à de OsAt10. Figura 49– Quantidade de ácido ferúlico especificamente unido à hemicelulose, liberado pela hidrólise ácida branda, no colmo das linhagens A871 (*Ox:ScAt10*) e controle linhagens escape e selvagem B73



Legenda: As linhagens transgênicas foram representadas em vermelho (A871-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

A diminuição observada em FA pode ser devida a uma competição de FA e pCA pelos sítios de ligação à GAX. Outra hipótese é de que o aumento na incorporação de pCA na hemicelulose resulte em uma diminuição na quantidade de pCA disponível na via dos fenilpropanóides, diminuindo a síntese de FA pela enzima (BARTLEY et al., 2013).

5.8.3 Análise do impacto do silenciamento ou super-expressão gênica no conteúdo de ácidos hidroxicinâmicos especificamente unido à lignina

Embora as análises anteriores tenham mostrado que ScAt10 está envolvido na modificação do conteúdo de *p*CA da hemicelulose, isso não exclui a possibilidade de que este gene possua efeito também sobre o teor de ácidos hidroxicinâmicos na lignina. Ademais, faz se necessário testar a hipótese de que a diferença no conteúdo de *p*CA éster

total para A8735C de fato estava relacionada à lignina. Assim, foi feita a análise pelo método DFRC do conteúdo de pCA e FA proveniente especificamente da lignina, para os mesmos eventos selecionados descritos na seção 5.7.2. Foram comparados os teores de monolignois conjugados S-pCA, detectados pelo método, já que a principal incorporação de pCA ocorre nas unidades S e muito pouco G-pCA foi detectado.

Corroborando com o observado para todas as análises deste trabalho, as linhagens A872 não diferem dos controles selvagem e escape (Figura 50).

Figura 50– Quantidade de ácido *p*-cumárico conjugado a unidades S (S-*p*CA), proveniente da análise DFRC, no colmo das linhagens A872 (*At10:RNAi*) e controle linhagens escape e selvagem (B73)



Legenda: As linhagens transgênicas foram representadas em vermelho (A872-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Corroborando também com o resultado obtido na análise do pCA éster total das linhagens A873 (seção 5.7.1), apenas o evento A873-5C apresentou redução significativa no pCA unido às unidades S da lignina (Figura 51).



Legenda: As linhagens transgênicas foram representadas em vermelho (A873-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

Em conjunto, os resultados para A873, tanto das análises genotípicas e fenotípicas, mostram que os fragmentos senso e anti-senso do gene de cana-de-açúcar, inseridos por meio da construção de *RNAi*, foram eficientes no silenciamento gênico de *ZmAt4* de maneira mais significativa na linhagem A873-5C, que apresentou diminuição no teor de *p*CA especificamente da lignina, e não da hemicelulose. O gene *ZmAt4* já foi caracterizado na literatura (MARITA et al., 2014), em que o silenciamento resultou na diminuição no teor de *p*CA total. No entanto, no trabalho de Marita *et al.* (2014), não foi feita a análise para detectar se essa diferença de fato era específica da lignina. O presente trabalho, utilizando um fragmento de cana para silenciar *ZmAt4*, mostra que de fato essa atividade é específica ao monolignol, e não à hemicelulose, como encontrado para *BdAt4* (PETRIK et al., 2014).

A análise do conteúdo de S-pCA nas linhagens A871 revelou que não há diferença significativa entre os transgênicos e os controles selvagem e escape (Figura 52).

Assim, verificou-se que *ScAt10* não tem influência no conteúdo de ácido *p*-cumárico unido à lignina. Efeito similar também foi verificado em arroz (*OsAt10*, Bartley et al., 2013).

Figura 52– Quantidade de ácido *p*-cumárico conjugado a unidades S (S-*p*CA), proveniente da análise DFRC, no colmo das linhagens A871 (*Ox:ScAt10*) e controle linhagens escape e selvagem B73



Legenda: As linhagens transgênicas foram representadas em vermelho (A871-) e os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Valores com *** foram significativos a p<0.001, ** a p<0.01 e * a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

5.8.4 Análise do impacto do silenciamento ou super-expressão gênica na composição da hemicelulose (GAX) de linhagens A873 e A871

É comum observar efeitos compensatórios quando se modifica um componente da parede celular (WOLF; HÉMATY; HÖFTE, 2012). Como a construção *ScAt4:RNAi* resultou na diminuição significativa de pCA na lignina na linhagem A873-5C, algum efeito poderia ser detectado na composição de outros polímeros da parede celular, como celulose ou hemicelulose. Ademais, a super-expressão de *ScAt10* causou efeitos significativos na incorporação de pCA à Glucuronoarabinoxilana, sendo possível que algum efeito compensatório pudesse ocorrer na composição de açúcares da hemicelulose, ou ainda no conteúdo total de glicose. Dessa forma, foram medidos os níveis de glicose, arabinose e xilose na parede celular das linhagens transgênicas A873-5C, bem como nas linhagens A871 selecionadas e nos controles (Figura 53)

Nenhuma diferença significativa (p < 0.05, ANOVA seguida do teste de Dunnet) foi detectada entre os transgênicos e os controles selvagem e linhagens escape. O resultado encontrado para as linhagens A871 (Ox:ScAt10) difere do obtido por Bartley et al. (2013). Ao super-expressar o gene OsAt10 em arroz, eles verificaram um leve aumento no conteúdo de glicose na palha madura. No entanto, a superexpressão de OsAt10 em setaria não provocou mudanças significativas no teor de glicose na palha, de forma similar ao encontrado no presente trabalho (LI et al., 2018). Como discutido por Bartley et al. (2013), não se sabe o que levou à diferença no teor de glicose nas linhagens de super-expressão OsAt10, provavelmente algum efeito compensatório que não foi resultado direto da atividade At10. Esse efeito deve ter ocorrido por alguma resposta induzida em arroz, que não ocorreu para a super-expressão do gene ScAt10 em milho ou de OsAt10 em setaria. De qualquer forma, a super-expressão de OsAt10 tanto em arroz como em setaria levou a aumentos nos rendimentos de sacarificação, indicando que o teor de glicose não foi o responsável pela melhora na digestibilidade dessas linhagens. Em seu trabalho com setaria, Li et al. concluem que o principal fator relacionado à diminuição da recalcitrância das linhagens de super-expressão de OsAt10 é a diminuição no teor de ácido ferúlico, reduzindo as ligações entre lignina e hemicelulose (LI et al., 2018).





Legenda: A linhagem transgênica de silenciamento *ScAt4:RNAi* A873-5C foi representada em verde e as linhagens de super-expressão *Ox:ScAt10* (A871-) foram representadas em vermelho. Os controles selvagem (B73) e escape (A872 4B e A873 1C) em azul. Valores representam a média e o erro padrão da média (SEM, n=2). Não houve diferença significativa entre as linhagens A871 e a média dos controles a p<0.05 (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles). Fonte: Autoria própria.

5.8.5 Análise do impacto da super-expressão gênica de ScAt10 na composição da lignina

Verificou-se também se houveram efeitos compensatórios na composição da lignina para as linhagens A873-5C (*ScAt4:RNAi*) e as linhagens A871 selecionadas (Ox:ScAt10) por meio das informações obtidas para a análise DFRC. Essa análise permite a quantificação dos monolignois, diferenciando os acilados dos não acilados. Assim, foram representados na tabela 12 a quantidade de unidades S, G e H, que representam extremidades fenólicas livres, e os conjugados S-pCA e G-pCA (Tabela 12).

Tabela 12– Quantidades de unidades S, G, H, S-*p*CA e G-*p*CA detectadas pelo método DFRC para as linhagens A871 e controles escape e selvagem B73

Linhagem	Н	G	S	S-pCA	G-pCA
A873 5C	0.47 ± 0.01	$15.29 \pm 0.29^*$	$7.93 \pm 0.00^{*}$	5.27±0.24***	0.11±0.00
A871 11B	0.29±0.01	11.59±0.00	5.18±0.13	13.57±0.47	0.38±0.01
A871 14B	0.28 ± 0.00	9.52±0.03	4.34±0.01	12.12±0.38	0.29±0.01
A871 2A	0.37 ± 0.00	10.32 ± 0.02	4.99±0.06	15.64±1.05	0.25 ± 0.02
A871 6A	0.29 ± 0.00	9.92±0.18	5.06 ± 0.14	14.18±0.81	0.34±0.02
Escape A872 4B	0.29±0.01	10.18 ± 0.20	5.48 ± 0.20	18.03±0.18	0.38±0.05
Escape A873 1C	0.22 ± 0.00	7.69 ± 0.24	4.68 ± 0.00	12.92±0.11	0.24 ± 0.00
Selvagem B73	0.46 ± 0.01	12.56±0.06	7.21±0.14	16.62 ± 1.62	0.31±0.02

Nota: A873 5C linhagem transgênica de silenciamento *ScAt4:RNAi*. A871 11B, A871 14B, A871 2A e A871 6A linhagens transgênicas de super-expressão *Ox:ScAt10*.

Valores representam a média e o erro padrão da média (SEM, n=2).

*** Diferença significativa a p<0.001 e * diferença p<0.05 entre as linhagens transgênicas e a média dos controles (ANOVA, seguido do teste *post-hoc* de Dunnet contra a média dos controles).
Fonte: Autoria própria.

Para a linhagem A873-5C, observou-se um leve aumento na quantidade G e S de extremidade fenólica livre. No entanto, considerando a redução de unidades S aciladas por pCA, a quantidade total de unidades S diminuiu levemente. Esse resultado concorda com o obtido por Petrik *et al.* (2014), em linhagens mutantes com perda de função bem como silenciadas por *RNAi* para o gene *BdAt4* (PETRIK et al., 2014).

Para as linhagens A871, não houveram diferenças significativas no teor H, G e S livres ou acilados, entre as ligninas dos transgênicos e os controles selvagem e linhagens escape. Ainda que pCA seja um intermediário da via dos fenilpropanóides, de maneira que o aumento da incorporação de pCA à arabinose poderia desequilibrar a proporção de cada um dos monolignois, a composição da lignina não parece ter sido afetada. Bartley *et al.*

(2013) também não encontraram nenhum efeito compensatório na lignina pela superexpressão de *OsAt10* em arroz, embora tenham analisado folha e palha madura (BARTLEY et al., 2013).

Em suma, os resultados demonstram relação do gene *ScAt10* com o conteúdo de pCA e FA na hemicelulose, de forma similar ao encontrado na literatura para o gene *OsAt10* (BARTLEY et al., 2013; LI et al., 2018). O grupo de genes *At10* parece estar relacionado com a atividade de incorporação de pCA na arabinose. No entanto, novos estudos são necessários para elucidar o mecanismo por meio do qual essa incorporação ocorre. Ainda é um enigma o fato de que as BAHD aciltransferases em geral são enzimas citoplasmáticas, enquanto a esterificação das xilanas acontece no complexo de Golgi. Assim, foi sugerido que *At10* possa transferir FA ou pCA a um intermediário no citoplasma, que depois seria transportado para o Golgi (LI et al., 2018).

6 CONCLUSÃO

A anotação dos genes BAHD em cana-de-açúcar, análise filogenética e do padrão de expressão permitiu a identificação de candidatos com potencial função relacionada à incorporação de ácidos hidroxicinâmicos, que podem ser explorados em estudos futuros sobre a via de adição dos ácidos ferúlico e p-cumárico na parede celular e aumento da digestibilidade. O fragmento de ScAt4 clonado no vetor de RNAi foi capaz de provocar o silenciamento gênico de ZmAt4 na espécie modelo Zea mays, sendo que uma linhagem A873-5C apresentou redução significativa (\sim 3x) no conteúdo de *p*CA exclusivo da lignina. Esses resultados fornecem mais uma evidência do papel dos genes At4 na incorporação de pCA ao monolignol em gramíneas, e abre caminho para mais estudos funcionais e de aplicação biotecnológica de ScAt4. A super-expressão de ScAt10 em milho resultou em aumento expressivo de pCA especificamente ligado à Arabinose e diminuição no conteúdo de FA, consistente com os outros estudos na literatura com o gene de arroz OsAt10, sugerindo envolvimento de ScAt10 com a incorporação de pCA à GAX. Ademais, o efeito provocado pela super-expressão de ScAt10, de diminuição de FA e aumento de pCA, pode ser explorado em cana-de-açúcar para obter linhagens transgênicas com maior digestibilidade, visando maior produtividade em biorefinarias. Como perspectivas, ensaios de digestibilidade enzimática estão sendo realizados nas linhagens de milho A871 (Ox:ScAt10) a fim de verificar se houve aumento na sacarificação. Por fim, a identificação, clonagem e transformação de ScAt4 e ScAt10 em vetores de superexpressão e silenciamento possibilita estudos futuros, como geração de duplos transgênicos ou ensaios enzimáticos, que poderão contribuir com informações a respeito da via de incorporação de *p*CA na hemicelulose e lignina de gramíneas.

REFERÊNCIAS

AMORIM, H. V. et al. Scientific challenges of bioethanol production in Brazil. **Applied Microbiology and Biotechnology**, v. 91, n. 5, p. 1267–1275, 2011.

BARTLEY, L. E. et al. Overexpression of a BAHD Acyltransferase, OsAt10, Alters Rice Cell Wall Hydroxycinnamic Acid Content and Saccharification. **Plant Physiology**, v. 161, n. 4, p. 1615–1633, 2013.

BAUCHER, M. et al. Critical Reviews in Plant Sciences Biosynthesis and Genetic Engineering of Lignin Biosynthesis and Genetic Engineering of Lignin. **Critical Reviews in Plant Sciences**, v. 17, n. 2, p. 125–197, 1998.

BEHERA, S. et al. Importance of chemical pretreatment for bioconversion of lignocellulosic biomass. **Renewable and Sustainable Energy Reviews**, v. 36, p. 91–106, ago. 2014.

BENNETZEN, J. L. et al. Reference genome sequence of the model plant Setaria. **Nature biotechnology**, v. 30, n. 6, p. 555–61, 2012.

BYRT, C. S.; GROF, C. P. L.; FURBANK, R. T. C4 Plants as Biofuel Feedstocks: Optimising Biomass Production and Feedstock Quality from a Lignocellulosic Perspective. Journal of Integrative Plant Biology, v. 53, n. 2, p. 120–135, 2011.

BLAKENEY, A. B. et al. A simple and rapid preparation of alditol acetates for monosaccharide analysis. **Carbohydrate Research**, v. 113, n. 2, p. 291–299, Mar. 1983.

BOERJAN, W.; RALPH, J.; BAUCHER, M. Lignin Biosynthesis. Annual Review of Plant Biology, v. 54, n. 1, p. 519–546, 2003.

BONTPART, T. et al. BAHD or SCPL acyltransferase? What a dilemma for acylation in the world of plant phenolic compounds. **New Phytologist**, v. 208, n. 3, p. 695–707, 2015.

BOTTCHER, A. et al. Lignification in Sugarcane: Biochemical Characterization, Gene Discovery, and Expression Analysis in Two Genotypes Contrasting for Lignin Content. **Plant Physiology**, v. 163, n. 4, p. 1539–1557, 2013.

BUANAFINA, M. M. D. O. et al. Functional testing of a PF02458 homologue of putative rice arabinoxylan feruloyl transferase genes in Brachypodium distachyon. **Planta**, v. 243, n. 3, p. 659–674, Mar. 2016.

BUANAFINA, M. M. DE O. Feruloylation in Grasses: Current and Future Perspectives. **Molecular Plant**, v. 2, n. 5, p. 861–872, Sept. 2009.

BURTON, R. A.; FINCHER, G. B. Current challenges in cell wall biology in the cereals and grasses. **Frontiers in Plant Science**, v. 3, n. 130, p. 1–6, 2012.

CARDOSO-SILVA, C. B. et al. De Novo Assembly and Transcriptome Analysis of Contrasting Sugarcane Varieties. **PLoS ONE**, v. 9, n. 2, p. e88462, Feb. 2014.

CASLER, M. D.; JUNG, H.-J. G. Selection and Evaluation of Smooth Bromegrass Clones with Divergent Lignin or Etherified Ferulic Acid Concentration. **Crop Science**, v. 39, p. 1866–1873, 1999.

CESARINO, I. et al. Suspension cell culture as a tool for the characterization of class III peroxidases in sugarcane. **Plant Physiology and Biochemistry**, v. 62, p. 1–10, 2013.

CHANDEL, A. K. et al. Sugarcane bagasse and leaves: Foreseeable biomass of biofuel and bio-products. **Journal of Chemical Technology and Biotechnology**, v. 87, n. 1, p. 11–20, 2012.

COSTA, T. H. F. F. et al. The enzymatic recalcitrance of internodes of sugar cane hybrids with contrasting lignin contents. **Industrial Crops and Products**, v. 51, p. 202–211, 2013.

D'AURIA, J. C. Acyltransferases in plants: a good time to be BAHD. **Current Opinion in Plant Biology**, v. 9, n. 3, p. 331–340, 2006.

DE CARLI POELKING, V. G. et al. Analysis of a modern hybrid and an ancient sugarcane implicates a complex interplay of factors in affecting recalcitrance to cellulosic ethanol production. **PLoS ONE**, v. 10, n. 8, 2015.

DE SOUZA, A. P. et al. Sugarcane as a Bioenergy Source: History, Performance, and Perspectives for Second-Generation Bioethanol. **Bioenergy Research**, v. 7, n. 1, p. 24–35, 2014.

DE SOUZA, W. R. et al. Suppression of a single BAHD gene in Setaria viridis causes large, stable decreases in cell wall feruloylation and increases biomass digestibility. **New Phytologist**, v. 218, n. 1, p. 81–93, Apr. 2018.

DE SOUZA, W. R. et al. Silencing of a BAHD acyltransferase in sugarcane increases biomass digestibility. **Biotechnology for Biofuels**, v. 12, n. 111, p. 1–14, Dec. 2019.

DEL RÍO, J. C. et al. Differences in the chemical structure of the lignins from sugarcane bagasse and straw. **Biomass and Bioenergy**, v. 81, p. 322–338, 2015.

DEMIRBAS, A. Biorefineries: Current activities and future developments. **Energy Conversion and Management**, v. 50, n. 11, p. 2782–2801, 2009.

DIAS, M. O. S. et al. Biorefineries for the production of first and second generation ethanol and electricity from sugarcane. **Applied Energy**, v. 109, p. 72–78, 2013.

DOS SANTOS, L. V. et al. Second-Generation Ethanol: The Need is Becoming a Reality. **Industrial Biotechnology**, v. 12, n. 1, p. 40–57, 2016.

EL-GEBALI, S. et al. The Pfam protein families database in 2019. Nucleic Acids Research, v. 47, n. D1, p. D427–D432, 2019.

FAIK, A. Xylan Biosynthesis: News from the Grass. **Plant Physiology**, v. 153, n. 2, p. 396–402, 2010.

FERRAZ, A. et al. Mapping of Cell Wall Components in Lignified Biomass as a Tool to Understand Recalcitrance. In: DA SILVA, S. S.; CHANDEL, A. K. (Eds.). **Biofuels in Brazil**. Cham: Springer International Publishing, 2014. p. 173–202.

GARSMEUR, O. et al. A mosaic monoploid reference sequence for the highly complex genome of sugarcane. **Nature Communications**, v. 9, n. 1, 2018.

GOUY, M.; GUINDON, S.; GASCUEL, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. **Molecular biology and evolution**, v. 27, n. 2, p. 221–224, 2010.

GRABBER, J. H. et al. Ferulate cross-linking in cell walls isolated from maize cell suspensions. **Phytochemistry**, v. 40, n. 4, p. 1077–1082, 1995.

GRABBER, J. H. et al. Genetic and molecular basis of grass cell-wall degradability. I. Lignin–cell wall matrix interactions. **Comptes Rendus Biologies**, v. 327, n. 5, p. 455–465, May 2004.

GRABBER, J. H.; HATFIELD, R. D.; RALPH, J. Diferulate cross-links impede the enzymatic degradation of non-lignified maize walls. **Journal of the Science of Food and Agriculture**, v. 77, n. 2, p. 193–200, June 1998.

GRABBER, J. H.; RALPH, J.; HATFIELD, R. D. Ferulate Cross-Links Limit the Enzymatic Degradation of Synthetically Lignified Primary Walls of Maize. Journal of Agricultural and Food Chemistry, v. 46, n. 7, p. 2609–2614, July 1998.

GUINDON, S. et al. New Algorithms and Mehtods to Estimate Maximum-Likelihood Phylogenies: Asessing the Performance of PhyML 2.0. **Systematic Biology**, v. 59, n. 3, p. 307–321, 2010.

HARTLEY, R. D. p-Coumaric and ferulic acid components of cell-walls of ryegrass and their relationship with lignin and digestibility. **Journal of the Science of Food and Agriculture**, v. 23, p. 1347, 1972.

HATFIELD, R. D. et al. Grass lignin acylation: P-coumaroyl transferase activity and cell wall characteristics of C3 and C4 grasses. **Planta**, v. 229, n. 6, p. 1253–1267, 2009.

HATFIELD, R. D.; MARITA, J. M.; FROST, K. Characterization of p -coumarate accumulation, p -coumaroyl transferase, and cell wall changes during the development of corn stems. **Journal of the Science of Food and Agriculture**, v. 88, n. 14, p. 2529–2537, Nov. 2008.

HATFIELD, R. D.; RANCOUR, D. M.; MARITA, J. M. Grass Cell Walls: A Story of Cross-Linking. **Frontiers in Plant Science**, v. 7, Jan. 2017.

HATFIELD, R.; RALPH, J.; GRABBER, J. H. A potential role for sinapyl p-coumarate as a radical transfer mechanism in grass lignin formation. **Planta**, v. 228, n. 6, p. 919–928, 2008.

HE, L.; TERASHIMA, N. Formation and Structure of Lignin in Monocotyledons. III. Heterogeneity of Sugarcane (Saccharum officinarum L.) Lignin with Respect to the Composition of Structural Units in Different Morphological Regions. **Journal of Wood Chemistry and Technology**, v. 10, n. 4, p. 435–459, 1990.

HIMMEL, M. E. et al. Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. **Science**, v. 315, n. 5813, p. 804–807, Feb. 2007.

HOOPES, G. M. et al. An updated gene atlas for maize reveals organ-specific and stressinduced genes. **Plant Journal**, v. 97, n. 6, p. 1154–1167, 2019.

IBGE. Produção Agrícola Municipal, 2017. Disponível em: https://sidra.ibge.gov.br/pesquisa/pam/tabelas. Acesso em: 15 jul. 2019.

INOUE, H.; NOJIMA, H.; OKAYAMA, H. High efficiency transformation of Escherichia coli with plasmids. **Gene**, v. 96, n. 1, p. 23–28, Jan. 1990.

IOWA STATE UNIVERSITY. Plant Transformation Facility, 2019. Disponível em: http://agron-www.agron.iastate.edu/ptf/. Acesso em: 15 jul. 2019.

JEOH, T. et al. Cellulase digestibility of pretreated biomass is limited by cellulose accessibility. **Biotechnology and Bioengineering**, v. 98, n. 1, p. 112–122, Sept. 2007.

JUNG, H.-J. G.; RALPH, J.; HATFIELD, R. D. Degradability of phenolic acidhemicellulose esters: A model system. **Journal of the Science of Food and Agriculture**, v. 56, n. 4, p. 469–478, 1991.

KARLEN, S. D. et al. Monolignol ferulate conjugates are naturally incorporated into plant lignins. **Science advances**, v. 2, p. 1–9, 2016.

KELLOGG, E. A. Update on Evolution Evolutionary History of the Grasses 1. **Plant physiology**, v. 125, p. 1198–1205, 2001.

KIM, C.; TANG, H.; PATERSON, A. H. Duplication and divergence of grass genomes: Integrating the chloridoids. **Tropical Plant Biology**, v. 2, n. 1, p. 51–62, 2009.

LAM, T. B. T.; IIYAMA, K.; STONE, B. A. Hot alkali-labile linkages in the walls of the forage grass Phalaris aquatica and Lolium perenne and their relation to in vitro wall digestibility. **Phytochemistry**, v. 64, n. 2, p. 603–607, 2003.

LAPIERRE, C. et al. Evaluation of Feruloylated and p-Coumaroylated Arabinosyl Units in Grass Arabinoxylans by Acidolysis in Dioxane/Methanol. Journal of Agricultural and Food Chemistry, v. 66, n. 21, p. 5418–5424, 2018.

LI, G. et al. Overexpression of a rice BAHD acyltransferase gene in switchgrass (Panicum virgatum L.) enhances saccharification. **BMC Biotechnology**, v. 18, n. 1, p. 1–10, 2018.

LINGLE, S. E.; THOMSON, J. L. Sugarcane Internode Composition During Crop Development. **BioEnergy Research**, v. 5, n. 1, p. 168–178, 2012.

LU, F.; RALPH, J. Efficient Ether Cleavage in Lignins: The Derivatization Followed by Reductive Cleavage Procedure as a Basis for New Analytical Methods. In: LEWIS, N.G.; SARKANEN, S. (Eds.). Lignin and lignan biosynthesis. Washington, D.C: American Chemical Society, 1998. p. 294–322.

LU, F.; RALPH, J. Detection and determination of p-coumaroylated units in lignins. **Journal of Agricultural and Food Chemistry**, v. 47, n. 5, p. 1988–1992, 1999.

MACADAM, J. W.; GRABBER, J. H. Relationship of growth cessation with the formation of diferulate cross-links and p-coumaroylated lignins in tall fescue leaf blades. **Planta**, v. 215, n. 5, p. 785–793, 2002.

MACRELLI, S.; MOGENSEN, J.; ZACCHI, G. Techno-economic evaluation of 2nd generation bioethanol production from sugar cane bagasse and leaves integrated with the sugar-based ethanol process. **Biotechnology for Biofuels**, v. 5, n. 1, p. 22, 2012.

MADEIRA, F. et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. **Nucleic Acids Research**, v. 47, n. W1, p. W636–W641, 2019.

MANOLI, A. et al. Evaluation of candidate reference genes for qPCR in maize. Journal of Plant Physiology, v. 169, n. 8, p. 807–815, 2012.

MARITA, J. M. et al. Variations in the cell wall composition of maize brown midrib mutants. **Journal of Agricultural and Food Chemistry**, v. 51, n. 5, p. 1313–1321, 2003.

MARITA, J. M. et al. Identification and suppression of the p-coumaroyl CoA:Hydroxycinnamyl alcohol transferase in Zea mays L. **Plant Journal**, v. 78, n. 5, p. 850–864, 2014.

MASARIN, F. et al. Chemical composition and enzymatic digestibility of sugarcane clones selected for varied lignin content. **Biotechnology for Biofuels**, v. 4, n. 1, p. 55, 2011.

MATTIELLO, L. et al. Physiological and transcriptional analyses of developmental stages along sugarcane leaf. **BMC Plant Biology**, v. 15, n. 1, p. 1–21, 2015.

MITCHELL, R. A. C.; DUPREE, P.; SHEWRY, P. R. A Novel Bioinformatics Approach Identifies Candidate Genes for the Synthesis and Feruloylation of Arabinoxylan. **Plant Physiology**, v. 144, n. 1, p. 43–53, 2007.

MOLINARI, H. B. C. et al. Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in Brachypodium distachyon. **Frontiers in Plant Science**, v. 4, n. 50, p. 1–14, 2013.

MOTTIAR, Y. et al. Designer lignins: Harnessing the plasticity of lignification. **Current Opinion in Biotechnology**, v. 37, p. 190–200, 2016.

MUELLER-HARVEY, I. et al. Linkage of p-coumaroyl and feruloyl groups to cell-wall polysaccharides of barley straw. **Carbohydrate Research**, v. 148, n. 1, p. 71–85, 1986.

OLIVEIRA, D. M. et al. Ferulic acid: A key component in grass lignocellulose

recalcitrance to hydrolysis. Plant Biotechnology Journal, v. 13, n. 9, p. 1224–1232, 2015.

PATERSON, A. H.; MOORE, P. H.; TEW, T. L. The Gene Pool of Saccharum Species and Their Improvement. In: PATERSON, A. H. (Ed.). Genomics of the Saccharinae. New York: Springer, 2013. p. 43–71.

PAULY, M. et al. Hemicellulose biosynthesis. Planta, v. 238, n. 4, p. 627–642, 2013.

PAULY, M.; KEEGSTRA, K. Cell-wall carbohydrates and their modification as a resource for biofuels. **Plant Journal**, v. 54, n. 4, p. 559–568, 2008.

PETRIK, D. L. et al. P-Coumaroyl-CoA: Monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in Brachypodium distachyon. **Plant Journal**, v. 77, n. 5, p. 713–726, 2014.

PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic acids research**, v. 29, n. 9, p. e45, 2001.

PIOVEZANI, A. R. Systems Integration Tool: uma ferramenta para integração e visualização de dados em larga escala e sua aplicação em cana-deaçúcar.2017.134p.Tese (Doutorado em Bioinformática)-Universidade de São Paulo, São Paulo, 2017.

PISTON, F. et al. Down-regulation of four putative arabinoxylan feruloyl transferase genes from family PF02458 reduces ester-linked ferulate content in rice cell walls. **Planta**, v. 231, n. 3, p. 677–691, 2010.

RAE, A. L.; MARTINELLI, A. P.; DORNELAS, M. C. Anatomy and Morphology. In: MOORE, P. H.; BOTHA, F. C. (Eds.). **Sugarcane: Physiology, Biochemistry and Functional Biology**. 1st ed. Oxford: Wiley-Blackwell, 2014. p. 19–33.

RAGAUSKAS, A. J. et al. The Path Forward for Biofuels and Biomaterials. Science, v. 311, n. 5760, p. 484–489, 2006.

RALPH, J. Hydroxycinnamates in lignification. **Phytochemistry Reviews**, v. 9, n. 1, p. 65–83, 2010.

RALPH, J.; GRABBER, J. H.; HATFIELD, R. D. Lignin-ferulate cross-links in grasses: active incorporation of ferulate polysaccharide esters into ryegrass lignins. **Carbohydrate Research**, v. 275, n. 1, p. 167–178, Sept. 1995.

RAMAKERS, C. et al. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. **Neuroscience Letters**, v. 339, n. 1, p. 62–66, Mar. 2003.

RANCOUR, D. M. et al. Cell wall composition and digestibility alterations in Brachypodium distachyon achieved through reduced expression of the UDP-arabinopyranose mutase. **Frontiers in Plant Science**, v. 6, n. 446, p. 1–20, June 2015.

RANCOUR, D. M.; MARITA, J. M.; HATFIELD, R. D. Cell wall composition throughout development for the model grass Brachypodium distachyon. **Frontiers in Plant Science**,

RANI, S. H. et al. Defective in Cuticular Ridges (DCR) of Arabidopsis thaliana, a gene associated with surface cutin formation, encodes a soluble diacylglycerol acyltransferase. **Journal of Biological Chemistry**, v. 285, n. 49, p. 38337–38347, 2010.

R CORE TEAM. R: A language and environment for statistical computing. Vienna, 2013. Disponível em: http://www.R-project.org/. Acesso em 15 jul. 2019.

REGNER, M. et al. Reductive Cleavage Method for Quantitation of Monolignols and Low-Abundance Monolignol Conjugates. **ChemSusChem**, v. 11, n. 10, p. 1580, 2018.

RIAÑO-PACHÓN, D. M.; MATTIELLO, L. Draft genome sequencing of the sugarcane hybrid SP80-3280. **F1000Research**, v. 6, n. 861, p. 1–7, June 2017.

RUIJTER, J. M. et al. Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. **Nucleic Acids Research**, v. 37, n. 6, 2009.

SAEMAN, J. F.; MOORE, W. E.; MILLETT, M. A. Sugar units present. Hydrolysis and quantitative paper chromatography. **Methods in carbohydrate chemistry**, v. 3, p. 54–69, 1963.

SANJUÁN, R. et al. Morphological and Chemical Composition of Pith and Fibers from Mexican Sugarcane Bagasse. **Holz als Roh- und Werkstoff**, v. 59, n. 6, p. 447–450, 2001.

SCHELLER, H. V.; ULVSKOV, P. Hemicelluloses. Annual Review of Plant Biology, v. 61, n. 1, p. 263–289, 2010.

SIMS, R. E. H. et al. An overview of second generation biofuel technologies. **Bioresource Technology**, v. 101, n. 6, p. 1570–1580, 2010.

SINGH, J.; SUHAG, M.; DHAKA, A. Augmented digestion of lignocellulose by steam explosion, acid and alkaline pretreatment methods: A review. **Carbohydrate Polymers**, v. 117, p. 624–631, 2015.

SIQUEIRA, G. et al. Topochemical distribution of lignin and hydroxycinnamic acids in sugar-cane cell walls and its correlation with the enzymatic hydrolysis of polysaccharides. **Biotechnology for biofuels**, v. 4, n. 1, p. 7, 2011.

SMITH, R. A et al. Engineering monolignol p-coumarate conjugates into Poplar and Arabidopsis lignins. **Plant Physiology**, v. 169, n. 4, p. 2992–3001, Oct. 2015.

SOCCOL, C. R. et al. Bioethanol from lignocelluloses: Status and perspectives in Brazil. **Bioresource Technology**, v. 101, n. 13, p. 4820–4825, 2010.

SOUZA, G. M. et al. The Sugarcane Genome Challenge: Strategies for Sequencing a Highly Complex Genome. **Tropical Plant Biology**, v. 4, n. 4, p. 145–156, 2011.

ST-PIERRE, B.; LUCA, V. DE. Chapter Nine Evolution of acyltransferase genes: Origin and diversification fo the BAHD superfamily of acyltransferases involved in secondary

metabolism. Recent Advances in Phytochemistry, v. 34, p. 285-315, 2000.

STRICKLER, S. R.; BOMBARELY, A.; MUELLER, L. A. Designing a transcriptome next-generation sequencing project for a nonmodel plant species. **American Journal of Botany**, v. 99, n. 2, p. 257–266, 2012.

SWISS INSTITUTE OF BIOINFORMATICS. Expasy Bioinformatics Resource Portal, 2019. Disponível em: https://web.expasy.org/translate/. Acesso em 15 jul. 2019.

TAKAHAMA, U.; ONIKI, T. Enhancement of peroxidase-dependent oxidation of sinapyl alcohol by an apoplastic component, 4-coumaric acid ester isolated from epicotyls of Vigna angularis L. **Plant and Cell Physiology**, v. 38, n. 4, p. 456–462, 1997.

TAMURA, K. et al. MEGA6: Molecular evolutionary genetics analysis version 6.0. **Molecular Biology and Evolution**, v. 30, n. 12, p. 2725–2729, 2013.

TIMPANO, H. et al. Brachypodium Cell Wall Mutant with Enhanced Saccharification Potential Despite Increased Lignin Content. **Bioenergy Research**, v. 8, n. 1, p. 53–67, 2015.

TUOMINEN, L. K.; JOHNSON, V. E.; TSAI, C.-J. J. Differential phylogenetic expansions in BAHD acyltransferases across five angiosperm taxa and evidence of divergent expression among Populus paralogues. **BMC Genomics**, v. 12, n. 1, p. 236, 2011.

VAN DER WEIJDE, T. et al. The potential of C4 grasses for cellulosic biofuel production. **Front Plant Sci**, v. 4, n. 107, p. 1-18, 2013.

VANDESOMPELE, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome biology**, v. 3, n. 7, p. 1–12, June 2002.

VETTORE, A. L. et al. The libraries that made SUCEST. **Genetics and Molecular Biology**, v. 24, n. 1–4, p. 1–7, 2001.

VETTORE, A. L. Analysis and Functional Annotation of an Expressed Sequence Tag Collection for Tropical Crop Sugarcane. **Genome Research**, v. 13, n. 12, p. 2725–2735, Dec. 2003.

VICENTINI, R. et al. Large-scale transcriptome analysis of two sugarcane genotypes contrasting for lignin content. **PLoS ONE**, v. 10, n. 8, p. 1–19, 2015.

VOGEL, J. Unique aspects of the grass cell wall. **Current Opinion in Plant Biology**, v. 11, n. 3, p. 301–307, 2008.

WENDE, G.; FRY, S. C. 2-o- β -d-xylopyranosyl-(5-o-feruloyl)-l-arabinose, a widespread component of grass cell walls. **Phytochemistry**, v. 44, n. 6, p. 1019–1030, 1997.

WITHERS, S. et al. Identification of grass-specific enzyme that acylates monolignols with p-coumarate. **Journal of Biological Chemistry**, v. 287, n. 11, p. 8347–8355, 2012.

WOLF, S.; HÉMATY, K.; HÖFTE, H. Growth Control and Cell Wall Signaling in Plants. **Annual Review of Plant Biology**, v. 63, n. 1, p. 381–407, 2012.

WYMAN, C. E. Cellulosic Ethanol: A Unique Sustainable Liquid Transportation Fuel. **MRS Bulletin**, v. 33, n. 04, p. 381–383, 2008.

YANG, B.; WYMAN, C. E. Effect of Xylan and Lignin Removal by Batch and Flowthrough Pretreatment on the Enzymatic Digestibility of Corn Stover Cellulose. **Biotechnology and Bioengineering**, v. 86, n. 1, p. 88–95, 2004.

ZHANG, J. et al. Allele-defined genome of the autopolyploid sugarcane Saccharum spontaneum L. **Nature Genetics**, v. 50, n. 11, p. 1565–1573, 2018.
pDONR:RNAi:ScAt10, PDONR:RNAi:ScAt4. OS SÍTIOS DE RESTRIÇÃO DAS ENZIMAS USADAS PARA ANEXO A – MAPAS DOS VETORES GATEWAY® pDONR221, pDONR:OX:ScAt10, pDONR:OX:ScAt4, CONFIRMAÇÃO DOS VETORES FORAM INDICADOS NA FIGURA.





USADAS PARA CONFIRMAÇÃO FORAM INDICADOS NA FIGURA. PZP221B:OX:ScAt4, PZP221B:RNAI:ScAt4 E PZP221B:RNAI:ScAt10. OS SÍTIOS DE RESTRIÇÃO DAS ENZIMAS ANEXO B – MAPAS DOS VETORES DE DESTINO GATEWAY® PZP221B:OX, PZP221B:RNAi, PZP221B:OX:ScAt10,



ANEXO C – CURVA DE MELTING DOS GENES BAHD DE CANA-DE-AÇÚCAR EM ANÁLISES DE RT-qPCR NOS HÍBRIDOS H89 E H321 DE CANA- DE -AÇÚCAR



Legenda: A - Curva de melting de cada gene indicado, nas amostras do córtex e medula dos híbridos H89 e H321, diluídas de 1:50. B - Curva de melting nas amostras provenientes da folha jovem do H89, diluídas de 1:50 para *ScAt10* e 1:25 para *ScAt5*. As curvas de melting com pico alto, definido e em colorido representam as amostras (A e B). Os picos ou linhas na cor bege claro representam o controle negativo (reação sem cDNA). Fonte: Autoria própria.

ANEXO D – CURVA DE MELTING DOS GENES ENDÓGENOS DE MILHO E OS TRANSGENES DE CANA-DE-AÇÚCAR NAS ANÁLISES DE RT-qPCR DAS LINHAGENS TRANSGÊNICAS DE MILHO A871, A873 E A885



Legenda: A- Curva de melting do transgene de cana-de-açúcar *Sc:At10* e dos genes constitutivos de milho *ZmLEU* e *ZmMEP*. Análises nas amostras das linhagens A871 (*Ox:ScAt10*). B- Curva de melting do gene endógeno de milho *ZmAt4* e dos genes constitutivos de milho *ZmCUL* e *ZmLEU*. Análises nas amostras das linhagens A873 (*At4:RNAi*).C- Curva de melting do transgene de cana-de-açúcar *ScAt4* e dos genes constitutivos de milho *ZmCUL* e *ZmLEU*. Análises nas amostras das linhagens A873 (*At4:RNAi*).C- Curva de melting do transgene de cana-de-açúcar *ScAt4* e dos genes constitutivos de milho *ZmCUL* e *ZmLEU* e endógeno *ZmAt4*. Análises nas amostras das linhagens A885 (*Ox:ScAt4*). As curvas em colorido representam as amostras A, B e C. O pico ou linha na cor bege claro representa o controle negativo. Fonte: Autoria própria.