TESE DE DOUTORADO

**REGULAÇÃO DA DEGRADAÇÃO DA PAREDE CELULAR DURANTE A FORMAÇÃO DO
AERÊNQUIMA EM RAÍZES DE CANA-DE-AÇÚCAR**

MSc. Eveline Queiroz de Pinho Tavares

Orientador: Dr. Marcos S. Buckeridge

São Paulo, maio de 2015

TESE DE DOUTORADO

**REGULAÇÃO DA DEGRADAÇÃO DA PAREDE CELULAR DURANTE A FORMAÇÃO DO
AERÊNQUIMA EM RAÍZES DE CANA-DE-AÇÚCAR**

Regulation of cell wall degradation during aerenchyma development in roots of sugarcane

MSc. Eveline Queiroz de Pinho Tavares

Orientador: Dr. Marcos S. Buckeridge

Tese apresentada ao Instituto de Biociências
da Universidade de São Paulo, para a
obtenção de Título de Doutor em Ciências, na
Área de Biologia de Sistemas

São Paulo, maio de 2015

FICHA CATALOGRÁFICA

TAVARES, Eveline Queiroz de Pinho

Regulação da degradação da parede celular durante a formação do aerênquima em raízes de cana-de-açúcar

Tese (Doutorado) – Instituto de Biociências da Universidade de São Paulo. Departamento de Botânica.

1. Aerênquima 2. Cana-de-açúcar 3. Parede celular 4. Regulação transcricional 5. Etanol 2G.

I. Universidade de São Paulo. Instituto de Biociências. Departamento de Botânica. Maio de 2015.

Comissão julgadora:

Prof(a) Dr(a):

Prof(a) Dr(a):

Prof(a) Dr(a):

Prof(a) Dr(a):

Prof. Dr. Marcos Silveira Buckeridge
Orientador

É sempre difícil a gente falar das qualidades da gente, mas acho que uma qualidade que a gente deve cultivar, não é que eu a tenha, mas que eu almejo tê-la, é saber-me frágil, saber-me incompleta, saber-me faltosa, dependente desse outro que me sustenta, que me completa, que me acalma e que me inquieta.

Marina Silva

AGRADECIMENTOS

Agradeço ao Prof. Dr. Marcos Buckeridge, orientador desta tese e grande mentor da idéia por trás deste trabalho;

Agradeço à FAPESP pela bolsa concedida e aos membros da banca pelo tempo dedicado a leitura desta tese e aprimoramento deste documento;

Preciso agradecer também aos inúmeros co-orientadores oficiais ou não, temporários ou não, dos quais precisei ao longo da execução do meu trabalho:

Amanda Pereira de Souza, que encontrou tempo em sua rotina corrida e leu todas as linhas que escrevi em todos relatórios, manuscritos, resumos de congresso e também nesta tese – mesmo depois de já ter deixado nosso grupo. Você foi uma grande amiga e inspiração desde que cheguei aqui e te desejo toda sorte nos próximos capítulos da sua vida.

Marie-Anne van Sluys, que abriu as portas de seu laboratório e me deu acesso a tudo que eu precisasse como se de seu grupo eu também fosse. Sua equipe é feita de pessoas tão competentes quanto prestativas e simpáticas, e em algum momento do doutorado tive o imenso prazer de poder contar com cada um deles: Jonas, Tati, Paula, Andréia, Érika Jesus, Andrés, Guilherme.

Jacqueline Grimma-Pettenati, une personne de routine lourde mais avec l'âme légère. J'espère encore que j'aurai une deuxième fois la chance de travailler dans ton groupe. Isabelle Truchet, la co-superviseur très bordelique sur la paillasse mais qui souriait avec les yeux. Vous deux, il faut que vous sachez que j'ai eu en France des jours magiques, presque insolites. Quand je pense au stage, je me rappelle des matins jaunes à Castanet, en faisant du vélo pour arriver au labo, le vent frais contre le visage.... et bien sur les PCR que ne marchait jamais. Anna, Chien, Marçal, Hong, Nathalie, Fabien, Annabelle, Audrey, Chantal, Christiane, Hua, Bruno, Raphael, gros bisous et un très grand merci!

Nathália de Setta e seu eterno sorriso no rosto, que me ajudou muito na reta final, sempre se prontificando a discutir, corrigir e me conduzir nas análises de sequências de uma plantinha tão complicada.... Foi diante dos seus olhos que um TE se inseriu em minha tese.

Eduardo Purgatto, uma enciclopédia ambulante de literatura de regulação hormonal. Obrigada por me ceder espaço na sua agenda e em seu laboratório. Sua contribuição e fé no meu trabalho me traz um imenso sentimento de dívida. Que venham publicações em retribuição!

Gláucia Mendes Souza, por ceder os iniciadores constitutivos e espaço em seu laboratório. Em especial agradeço à Carolina Lembke pelo microarranjo, parte fundamental desta tese.

Aos amigos que conheci no LAFIECO, que atormento diariamente: Eagle, Menina Greuce, Débora, Dri, Amanda Rusiska, Vivi, Vinícius, Luís, Giovanna, Ivan. Com a companhia de vocês era sempre mais fácil ver o lado bom de tudo (até das infundáveis coletas de raiz). *Minhas raízes, eu balanço, eu balanço...*

Por mais importante que seja essa etapa em minha vida acadêmica, agradeço acima de tudo aqueles que me fizeram pensar em outra coisa que não BACs, pectina ou sinalização por etileno. Gente que perto ou longe ajuda a lapidar a pessoa por trás da pesquisadora. Gente que inclusive sequer sabe o que são BACs, pectina ou sinalização por etileno, aliás. Prefiro que minha mãe continue sabendo é que o frango com quiabo que ela faz é minha comida preferida, que meu pai saiba que gostaria de dividir uma cerveja com ele mais vezes e que a Karla saiba que pode fazer mais pulseirinhas para mim que vou usar! Que a Keila saiba que o jeito que a cabeça dela funciona me fascina, então nem faz mal que ela não saiba o que é um fator de transcrição. E à Carla, cujos cafunés nunca são suficientes... te devo mais do que poderia retribuir a alguém, sobretudo pelo suporte emocional, pela torcida e por ainda insistir em valorizar meu trabalho quando eu mesma não o faço. Por fim, agradeço ao Floriano por pular no meu colo e * * * pelo teclado do computador nos momentos em que a inspiração fugia e o cansaço chegava... A essas pessoas (de duas ou quatro patinhas), preciso dizer que é um prazer enorme ser neste momento motivo de orgulho de vocês.

ÍNDICE

RESUMO	1
ABSTRACT	2
INTRODUÇÃO	3
1. Uso de biomassa vegetal para bioenergia	3
2. Parede celular: composição e degradação endógena	4
3. Aerênquima	5
3.1 Formações induzida e constitutiva	6
3.1.1 Hipóxia	7
3.1.2 Estresse nutricional	9
4. Cascatas de sinalização e fatores de resposta ao etileno	10
4.1 Efeito do etileno na formação do aerênquima	14
5. Morte celular programada	16
6. Degradação da parede celular na formação do aerênquima	17
7. Viabilizando o pré-tratamento biológico em cana-de-açúcar	20
OBJETIVOS GERAL E ESPECÍFICOS	22
DESENHO EXPERIMENTAL	23
METODOLOGIA	24
Cultivo de cana-de-açúcar e coleta de raiz	24
<u>PARTES A & B:</u> Cultivo de cana-de-açúcar em estresse nutricional e na presença de 1-MCP	25
1. Cultivo de cana-de-açúcar irrigada com distintas soluções de nutrientes	25
2. Cultivo de cana-de-açúcar em 1-metilciclopropeno	26
3. Preparação das amostras para microscopia de luz e cálculo da área do aerênquima	27
4. Dosagem de AIA livre	28
5. Dosagem de etileno	28
6. Escolha de sequências-alvo e desenho de oligonucleotídeos específicos	29
7. Extração de RNA de raízes de cana-de-açúcar e síntese de cdna para qPCR e microarranjo	31
8. Análise dos padrões de expressão gênica por pcr em tempo real	32
9. Hibridização em oligoarray e processamento da imagem	32
<u>PORTE C:</u> Seleção e sequenciamento de genes alvo	33
1. Triagem das sequências alvo na biblioteca de BACs de <i>Saccharum hybrid</i> (SHCRBa) da linhagem R570	33
2. Extração de DNA plasmidial em média escala	35
3. Preparo da biblioteca, sequenciamento na plataforma 454/Roche e predição de genes	35
4. Avaliação da predição automática e caracterização gênica das sequências candidatas	35
<u>PORTE D:</u> Análise de sintenia, filogenia e estrutura dos genes alvo sequenciados	36

ÍNDICE

1. Avaliação da colinearidade e sintonia dos BACs de cana-de-açúcar em relação a <i>S. bicolor</i>	37
2. Análise da estrutura dos genes alvo	37
<u>PARTE E: Seleção da sequência para clonagem e análise <i>in silico</i> dentre as versões do gene alvo</u>	37
1. Mapeamento dos reads do RNAseq/aerênquima nas sequências codificantes dos genes alvo	37
2. Análise <i>in silico</i> das proteínas preditas a partir da versão mais expressa dos genes alvo	38
<u>PARTE F: Avaliação da interação dos fatores de transcrição RAV e ERF com os promotores das glicosil hidrolases</u>	39
1. Identificação <i>in silico</i> de elementos em <i>cis</i> nos promotores de <i>ScEPG1a</i> e <i>ScARA1a</i>	39
2. Método de choque térmico para obtenção de células competentes de <i>Escherichia coli</i> e transformação	39
3. <i>Agrobacterium tumefaciens</i> eletrocompetentes e transformação	39
4. Síntese de oligonucleotídeos para vetores para transformação de <i>A. tumefaciens</i>	40
5. Amplificação das sequências alvo	40
6. Construção de vetores intermediários e finais	41
7. Sequenciamento	42
8. Infiltração de folhas de <i>Nicotiana benthamiana</i>	42
9. Extração de proteínas de discos foliares de <i>N. benthamiana</i> e ensaio de β -glucuronidase	43
ANÁLISE ESTATÍSTICA	43
RESULTADOS	44
<u>PARTES A & B: Cultivo de cana-de-açúcar em estresse nutricional e na presença de 1-MCP</u>	44
1. Área ocupada pelo aerênquima no córtex da raiz	44
2. Perfil transcricional diante de distintas disponibilidades de nutrientes	46
3. Formação do aerênquima e dosagem hormonal diante da aplicação de 1-MCP	50
4. Expressão diferencial de genes avaliada por microarranjo após tratamento com 1-MCP	52
5. Validação do microarranjo por qPCR	54
<u>PARTE C: Seleção e sequenciamento de genes alvo</u>	56
1. Características gerais dos BACs sequenciados	56
<u>PARTE D: Análise de sintonia, filogenia e estrutura dos genes alvo sequenciados</u>	36
1. Análises de colinearidade e sintonia dos BACs homeólogos de cana-de-açúcar em relação a <i>S. bicolor</i>	58
2. Caracterização da estrutura gênica dos genes alvo homeólogos em relação a <i>S. bicolor</i>	69
<u>PARTE E: Seleção da sequência para clonagem e análise <i>in silico</i> dentre as versões do gene alvo</u>	73

ÍNDICE

1. Análise do perfil transcricional das cópias de <i>ScEPG1</i> , <i>ScRAV1</i> , <i>ScARA1</i> e <i>ScERF1</i>	73
2. Caracterização das proteínas preditas a partir da sequência dos genes-alvo	76
<u>PARTE F: Avaliação da interação dos fatores de transcrição RAV e ERF com os promotores</u>	86
das glicosil hidrolases	
1. Mapeamento de elementos em cis nos promotores de <i>ScEPG1a</i> and <i>ScARA1a</i>	86
2. Construção dos vetores para transformação de <i>A. tumefaciens</i>	87
3. Regulação transcricional mediada por RAV1a e ERF1a em folhas de <i>N. bethamiana</i>	88
DISCUSSÃO	102
1. A formação constitutiva do aerênquima em cana-de-açúcar	90
2. Padrão transcricional decorrente da perturbação no âmbito hormonal	94
3. Caracterização dos <i>loci</i> ocupados pelos genes alvo sequenciados	99
4. Análise comparativa das proteínas codificadas por <i>ScARA1a</i> e <i>ScEPG1a</i> e o possível	101
emprego biotecnológico	
5. Aspectos estruturais e funcionais preditos dos fatores de transcrição sequenciados	107
6. Regulação transcricional de <i>ScEPG1a</i> por <i>ScRAV1a</i>	108
7. A estrutura e funcionalidade do genoma de cana-de-açúcar para os quatro genes alvo	110
CONCLUSÃO	112
REFERÊNCIAS	114
ANEXOS	134
ARTIGO ACEITO PARA PUBLICAÇÃO	154

A fim de contornar a problemática imposta pela recalcitrância da parede celular à hidrólise, o processo de produção de etanol a partir de biomassa vegetal requer um passo denominado de pré-tratamento, que consiste em tornar a biomassa mais acessível à ação de glicosil hidrolases que atacam a parede celular. O melhor conhecimento de processos de degradação de parede operados pela própria planta tem o potencial de direcionar as pesquisas em bioenergia, indicando quais os mecanismos mais eficientes para desmontar o complexo de polímeros da parede celular. Cinhado pré-tratamento biológico, este consiste na hidrólise de polissacarídeos estruturais empregando mecanismos e elementos chave de processos de degradação de parede celular que ocorrem na própria planta. Um exemplo de evento com esta característica é a formação de aerênquima lisígeno, que consiste na abertura de espaços de gás em tecidos parenquimáticos. A formação do aerênquima em cana-de-açúcar se dá de forma modular, sendo o processo constituído por seis etapas: 1) percepção do sinal inicial; 2) separação celular; 3) expansão celular; 4) morte celular programada; 5) hidrólise de hemiceluloses e 6) hidrólise de celulose. O presente trabalho teve como principal foco estudar a regulação das duas etapas iniciais, visando obter conhecimentos que possibilitem para tornar a biomassa de cana de açúcar menos resistente à penetração de enzimas.

O aerênquima ocorre na raiz de cana-de-açúcar por um processo constitutivo. Sua independência de um indutor externo foi corroborada mediante tratamento com nutrientes e inibidor da percepção por etileno (1-MCP) pela análise dos cinco centímetros mais apicais da raiz. O atraso na formação do aerênquima após tratamento com nutrientes levou à expressão diferencial de genes relacionados à degradação de parede celular, morte celular programada e sinalização por etileno. Por outro lado, o 1-MCP não afetou visivelmente a formação do aerênquima, porém alterou o balanço hormonal na raiz. O padrão transcricional das raízes tratadas com 1-MCP revelou aumento da expressão de genes relacionados à expansão celular e estresse oxidativo. Tais padrões são discutidos à luz da regulação hormonal da formação do aerênquima através do estabelecimento de um balanço hormonal definido entre auxina e etileno. Ambos os experimentos levaram à seleção de quatro genes candidatos: dois fatores de transcrição (*ScRAV1* e *ScERF1*) e duas glicosil hidrolases (*ScEPG1* e *ScARA1*), sequenciados e analisados quanto à diversidade hom(e)óloga, sequências promotoras e estrutura gênica em comparação a *S. bicolor*. A topologia das reconstruções filogenéticas e a distribuição diferencial de sítios para RAV e ERF nos promotores de *ScEPG1* e *ScARA1* não parece refletir padrões de expressão distintos, visto que os diferentes hom(e)ólogos demonstraram ser igualmente expressos.

Em sistema heterólogo, o fator de transcrição RAV apresentou atividade repressora sobre o promotor de *ScEPG1*. Tal papel de RAV sugere regulação negativa sobre a degradação da lamela média e sobre o processo de perda de adesão e separação celular. A identificação de reguladores-chave da formação do aerênquima consiste em importante elo entre a sinalização hormonal e a degradação de parede celular, possibilitando a melhor compreensão dos mecanismos de controle da degradação endógena de parede celular. Os dados produzidos possibilitam a aplicação biotecnológica dos genes sequenciados, além de consistirem em significativo avanço no conhecimento sobre a fisiologia do processo estudado e na dinâmica da regulação da expressão gênica acoplada a aspectos filogenéticos das sequências alvo.

Palavras-chave: aerênquima; cana-de-açúcar; parede celular; regulação transcricional; etanol 2G.

ABSTRACT

In order to circumvent the problems imposed by the cell wall recalcitrance to hydrolysis, the process underlying biofuel production requires a step called pretreatment, aimed at making biomass more accessible to the action of glycosyl hydrolases that attack the cell wall. The increased knowledge of wall degradation processes operated by the plant itself has the potential to influence research in the bioenergy field, pointing out the most efficient mechanisms to disassemble the cell wall complex polymeric structure. The term coined biological pretreatment means taking advantage of key elements and mechanisms of cell wall degradation processes occurring in the plant itself in order to disassemble the entangled polysaccharide structure. One example of endogenous cell wall degradation process is the lysigenous aerenchyma formation, which consists in the opening of gas spaces in parenchyma tissue. The aerenchyma formation in sugarcane is thought to be a modular process that occurs in six steps: 1) target cells perception; 2) cell separation; 3) cell expansion; 4) programmed cell death; 5) hemicellulose hydrolysis and 6) cellulose hydrolysis. This work has as the main objective to study the regulation of the two initial steps, aiming at acquiring knowledge that would afford to turn biomass less resistant to enzyme penetration.

The aerenchyma develops within the roots of sugarcane as a constitutive process. Its independence from an external inducer was corroborated in this work by subjecting sugarcane plants to treatment with nutrients and one inhibitor of ethylene perception (1-MCP) when the five more apical centimeters of the root were analyzed. After treatment with nutrients, the delayed aerenchyma formation led to differential expression of cell wall degradation-, programmed cell death- and ethylene signalling-related genes. Under visual inspection, 1-MCP did not show effect on aerenchyma development. Instead, it changed the hormone balance mainly in the two most apical root segments. The transcriptional pattern of 1-MCP treated roots revealed increased expression of cell expansion- and oxidative stress-related genes. Such patterns are discussed in the light of the hormone regulation of the aerenchyma development through the establishment of a balance between auxin and ethylene within root segments. Both experiments led to the selection of four two candidate genes, two transcription factors (*ScRAV1* and *ScERF1*) and two glycosyl hydrolases (*ScEPG1* and *ScARA1*), that were sequenced and analyzed regarding homologous diversity, promoter sequences and gene structure compared to *S. bicolor*. The topology of the phylogenetic reconstructions and the differential distribution of sites for RAV and ERF within *ScEPG1* and *ScARA1* promoters did not reflect distinct expression patterns. Different hom(e)ologous of each target gene were equally expressed.

In an heterologous system, RAV transcription factor interacted with *ScEPG1* promoter, reducing the activity encoded by the reporter gene. This suggests the repressive role of RAV on pectin degradation within the middle lamella, leading to reduced cell adhesion and cell separation, possibly regulated by this glycosyl hydrolase. The identification of key regulators of the aerenchyma formation is an important link between hormone signaling and the wall degradation within the sugarcane roots. It enables a better understanding of control mechanisms underlying wall modifications when performed by the plant itself. Altogether, the data produced in this work allows biotechnological application of sequenced genes. Moreover, the produced data unveil key aspects regarding physiological aspects of aerenchyma development and highlights features related to the dynamics of gene expression in sugarcane coupled to the phylogenetic aspects of the four target sequences.

Keywords: aerenchyma; sugarcane; cell wall; transcriptional regulation; ethanol 2G.

1. Uso de biomassa vegetal para bioenergia

Estima-se que a produção anual de biomassa vegetal no mundo seja na faixa de 170 a 200 x 10⁹ toneladas por ano (Lieth, 1975), dos quais cerca de 70-75% correspondem à parede celular vegetal (Gomez *et al.*, 2008). Apesar de ser a matéria prima vegetal mais abundante do planeta, ela é também largamente desperdiçada, visto que somente 2% é utilizada na indústria têxtil, produção de papel e geração de energia (Pauly & Keegstra, 2008).

A principal fonte energética mundial atual, principalmente a relacionada com o setor de transportes, é derivada do petróleo (Gomez *et al.*, 2008). O uso contínuo dessa matriz energética tem contribuído fortemente para aumento de gases estufa na atmosfera, que por sua vez são considerados fatores de grande impacto para mudanças no clima (IPCC, 2014). Ao mesmo tempo, o desenvolvimento sócio-econômico acarreta em ascendente demanda por energia. Desse cenário emerge a necessidade da produção de uma quantidade maior de energia ao mesmo tempo em que esta deve ser produzida de maneira sustentável. Portanto, torna-se imperativa a elaboração de tecnologias que visem a maior participação da biomassa vegetal na indústria de combustíveis.

A tradição sucroalcooleira brasileira e o amadurecimento das indústrias e refinarias em nosso país nos consolidou como referência na área de biocombustíveis e segundo maior produtor de etanol, ultrapassando os 27 bilhões de litros em 2013 (UNICA, 2015). Já nos Estados Unidos, maior produtor de etanol da atualidade, os 47 bilhões de litros de etanol produzidos a partir de amido de milho em 2013 foram responsáveis pela significativa independência em relação à importação de combustíveis fósseis (RFA, 2014). Entretanto, níveis de cerca de 75 bilhões de litros de etanol foram projetados para este país até 2022 (RFS, 2005), demonstrando claramente a necessidade de investimento em tecnologia e iniciativas voltadas para a otimização da produção.

O combustível produzido a partir da fermentação da sacarose extraída da cana-de-açúcar no Brasil e do amido de milho nos Estados Unidos é conhecido como etanol de primeira geração (1G). Previsões apontaram para a perspectiva que, na ausência de subsídios, o etanol 1G não será capaz de substituir completamente os combustíveis fósseis (Trépant *et al.*, 2009). Além disso, o uso de sacarose e amido pela indústria de combustíveis têm gerado intensas discussões diante do aumento do preço dos alimentos e do crescimento da população mundial que não tem pleno acesso à alimentação (Tenenbaum *et al.*, 2008). Nesse contexto, a produção de etanol de segunda geração (2G) tem sido colocada em evidência, visto que a mesma se dá a partir de material vegetal não aproveitado como alimento e muitas vezes produto da tecnologia 1G. Além disso, o etanol 2G tem o potencial de substituir a tecnologia anterior no prazo de 10 a 15 anos (IEA, 2011) e/ou complementá-la (Macreli *et al.*, 2014).

Gramíneas como *Sorghum bicolor*, *Miscanthus sinensis*, *Panicum virgatum*, *Zea mays* e *Saccharum* sp. (cana-de-açúcar), apresentam forte potencial biotecnológico voltado ao uso como matéria prima na produção de etanol 2G (De Souza *et al.*, 2014). Visto que no Brasil a tecnologia de primeira geração encontra-se amadurecida e otimizada no que diz respeito ao uso de cana-de-açúcar (De Souza *et al.*, 2014), o advento de estratégias para produção de etanol 2G a partir da mesma planta emerge como

forte alternativa para contornar a crise energética iminente, sem necessidade de aumentar a área de plantio. O etanol 2G é resultado da produção de açúcares simples gerados pela degradação de polissacarídeos complexos que estruturam a parede celular vegetal. Tendo isso em vista, tanto o conhecimento de genes chave da biossíntese e degradação de polissacarídeos estruturais (Carpita *et al.*, 2001; Tavares *et al.*, 2015) quanto da arquitetura da parede celular (Buckeridge *et al.*, 2015) são fundamentais para o advento de rotas alternativas para desconstrução desta estrutura no contexto da bioenergia.

2. Parede celular: composição e degradação endógena

No caso de gramíneas sabe-se que a parede celular é característica do tipo II, com arabinoxilanos (AXs) como hemicelulose prevalente (Carpita & Gibeaut, 1993). Tratam-se de uma cadeia principal de resíduos de D-xilose unidos por ligações do tipo β -1,4 decorada por cadeias laterais de L-arabinose que se unem à xilose por ligações do tipo α -1,3. Com base nos dados da composição da parede celular de cana-de-açúcar (De Souza *et al.*, 2013) estima-se que este polissacarídeo responda por cerca de 40% desta estrutura. Acredita-se que este apresente interligações por meio de pontes de ferúlicos e ligações cruzadas com a celulose (Kozlova *et al.* 2014; Buckeridge *et al.*, 2015), um polímero linear de D-glicose β -1,4 ligada (Carpita & McCann, 2000). Uma outra região da molécula de AX interage mais intimamente com o β -glucano (BG) ou glucano de ligação mista (Kozlova *et al.*, 2014), predominante em monocotiledôneas e com estrutura similar à celulose, no qual as ligações β -1,3 surgem a cada três ou quatro ligações β -1,4 (Carpita & McCann, 2000;). Alternativamente, outras linhas de evidência atribuem ao xiloglucano (XG) o papel de hemicelulose mais firmemente ligado à celulose (De Souza *et al.*, 2013; Buckeridge *et al.*, 2015), sendo este uma cadeia linear de (1 $\rightarrow$ 4) β -D-glucano com ramificações de α -D-xilose (Carpita & McCann, 2000) que responde por 8% da parede celular de cana-de-açúcar (de acordo com estimativas a partir dos dados de De Souza *et al.*, 2013). Em cana-de-açúcar existem evidências que suportam uma forte interação entre xiloglucano e arabinoxilano com a microfibrila de celulose, dita “unidade básica” da parede celular (Buckeridge *et al.*, 2015). Envolvendo esta unidade está a lignina, uma complexa rede de fenilpropanóides (Vanholme *et al.*, 2010).

Por fim, pectinas são os polissacarídeos estruturalmente mais diversos e complexos, consistindo em uma família de polímeros cujos principais membros são o homogalacturonano (HG) e ramnogalacturonanos (RGs) I e II. Homogalacturonanos são as pectinas mais abundantes, consistindo em polímeros lineares oriundos da união de resíduos de ácido galacturônico (α -D-GalA) unidos por ligações α -1,4, atingindo até 100 resíduos de comprimento e contabilizando cerca de 65% da pectina (Carpita & McCann, 2000). HG pode ser parcialmente metil esterificado na posição C-6, podendo também ser O-acetilado nas posições O-2 ou O-3 ou conter outros ésteres em ligação cruzada, com estrutura pouco conhecida (Caffall & Mohnen, 2009). As demais pectinas consistem em HGs ramificados como o RG-I, que representa de 20 a 35% desta classe de polissacarídeos. Este apresenta uma cadeia principal formada pela repetição do dissacarídeo α -D-GalA-1,2 α -L-Rha-1,4, sendo que os resíduos de ramnose (Rha) apresentam cadeias laterais contendo um ou mais resíduos de α -L-Araf e β -D-Galp, lineares ou ramificados

(Carpita & McCann, 2000). O polímero mais complexo desta família, RGII, responde por 10% da pectina total, consistindo em cerca de oito resíduos de α -D-GalA 1,4 ligados decorados com cadeias laterais contendo até doze diferentes tipos de açúcares com cerca de vinte distintos tipos de ligações. RG-II comumente existe em paredes celulares na forma de dímeros estabilizados por um éster diol 1:2 borato, estruturando redes macromoleculares de pectinas, cruciais para o crescimento e funcionamento normais desta estrutura (Mohnen *et al.*, 2008).

A otimização de um processo que aumente o acesso de enzimas a estes polissacarídeos é um dos gargalos visualizados para a produção de etanol 2G em escala industrial. Embora exista a pesquisa em pré-tratamentos químicos e físicos visando aumentar o acesso do substrato às glicosil hidrolases, seu sucesso tem sido relativo uma vez que as metodologias testadas são onerosas e poluentes e as altas quantidades de enzimas elevam o custo desta tecnologia, inviabilizando a produção em larga escala (Margeot *et al.*, 2009).

Tendo em vista que os mais eficientes mecanismos de afrouxamento e degradação de parede celular são aqueles operados pelas próprias plantas, tem sido discutido que o estudo de aspectos chave destes mecanismos poderia melhor direcionar a pesquisa em bioenergia (Grandis *et al.*, 2014). Sabe-se que a degradação de pectinas é um dos pontos cruciais do amadurecimento de frutos (Brummell, 2006; Giovannoni, 2004), bem como a etapa inicial da formação de aerênquimas (Gunawardena *et al.*, 2001b; Leite, 2013). A consequente perda de adesão celular e aumento da porosidade desencadeiam alterações na parede que favorecem os passos subsequentes da hidrólise dos outros polissacarídeos da parede (Grandis *et al.*, 2014; Tavares *et al.*, 2015). Similarmente, o ataque a outros polissacarídeos como XG e BG também é descrito no contexto da degradação endógena de parede celular. A mobilização do xiloglucano do cotilédone que se dá no estabelecimento da plântula em *Hymenaea courbaril* provê açúcares e energia que possibilitam o crescimento (Tiné *et al.*, 2000). Por sua vez, as etapas requeridas para a hidrólise de β -glucano foram já descritas na germinação de cereais e gramíneas (Buckeridge *et al.*, 2004).

Tendo em vista a posição já consolidada da cana-de-açúcar na indústria sucroalcooleira brasileira, a viabilidade de tal intento é aumentada caso invistamos na busca de mecanismos endógenos de degradação de parede celular nesta planta. Notavelmente, em cana-de-açúcar a formação de aerênquima foi identificada no colmo e na raiz (Tetsushi & Karim, 2007) em plantas alagadas ou em terra firme (Begum *et al.*, 2013) e a caracterização da degradação de parede e identificação de proteínas envolvidas neste processo se deu em nosso grupo (Leite, 2013; Grandis, 2015).

3. Aerênquima

Raízes, rizomas e outros órgãos em plantas normalmente capturam O_2 para respiração diretamente da atmosfera que os cercam. Entretanto, durante períodos de alagamento ou em solos muito compactados, a oxigenação dos tecidos internos da planta pode ser gravemente reduzida. Caso não haja reabastecimento e restabelecimento das concentrações normais de O_2 na atmosfera circundante, todo o oxigênio remanescente será consumido pela planta e por microrganismos e o solo será em breve incapaz

de suprir os organismos nele presentes. Neste contexto, o aerênquima – tecido que compreende alta proporção de espaços preenchidos por gás, resultantes da morte ou crescimento diferencial de grupos de células – propicia à planta um método de oxigenação alternativo, consistindo em uma resposta adaptativa a circunstâncias de hipóxia (Colmer 2003; Evans 2003). A depender da espécie e condições ambientais, o aerênquima é formado ao longo de todo o corpo de uma planta, com a exceção de meristemas, feixes vasculares e tecidos esclerenquimáticos e de revestimento (Raven *et al*, 2007). Canais de ar característicos do aerênquima são mais comumente descritos no córtex das raízes, rizomas e caules (revisado por Yamauchi *et al* 2013). Nas raízes, o aerênquima é formado em regiões próximas ao meristema apical (Malik *et al*, 2003).

Descrito pela primeira vez por Sachs (1882), a formação do aerênquima pode se dar de modo esquizógeno, o qual consiste na separação de células, gerando lacunas de gás sem a necessidade de ocorrência de morte celular. Em contrapartida, o aerênquima que se desenvolve por meio de processo lisígeno é resultado da morte celular de grupos específicos de células cujas paredes celulares são atacadas, sendo o espaço posteriormente preenchido por gás (Evans, 2003). Em razão da hidrólise da parede celular se dar no processo lisígeno, a investigação de sua formação é de valioso interesse para a pesquisa em bioenergia.

A formação de aerênquima lisígeno é descrita no córtex de raízes de *Zea mays* (Drew *et al.*, 1989), *Oryza sativa* (Justin & Armstrong, 1991b), *Hordeum vulgare* (Arikado & Adachi, 1955), *Triticum aestivum* (Trought & Drew, 1980), *Sorghum bicolor* (Promkhambut *et al.*, 2011) e *Sagittaria lancifolia* (Shussler & Longstreth, 1996). Este é detectado em suas etapas iniciais em regiões do córtex próximas ao ápice radicular, podendo ocorrer tanto em raízes seminais jovens (Konings, 1982), bem como em raízes adventícias (Drew *et al.*, 1981). O desenvolvimento se dá acropetalmente, simultaneamente à extensão da raiz em direção ao solo (Evans, 2003).

O aerênquima desenvolvido por meio de lisigenia fornece não só um caminho interno para transporte de O₂, mas também simultaneamente reduz o número de células consumidoras deste gás (Drew *et al* 2000; Colmer, 2002). A espessa camada de suberina produzida em decorrência da formação do aerênquima lisígeno em raízes atua como uma barreira à perda de oxigênio para a rizosfera e maior oxigenação do ápice radicular (Jackson & Armstrong, 1999; Colmer, 2003). Vários estudos focam no desenvolvimento de aerênquima lisígeno em raiz e seu desenvolvimento constitutivo ou induzido, sobretudo graças às vantagens agronômicas no que diz respeito ao maior desempenho de plantas em ambientes com baixos níveis de oxigênio, como solos alagados ou muito compactados.

3.1 Formações induzida e constitutiva

A formação do aerênquima como uma resposta adaptativa em *Z. mays* se dá em consequência do cultivo em hipóxia, estresse nutricional e osmótico, além de impedância mecânica e aplicação de etileno (Drew *et al.*, 2000; Yamauchi *et al.*, 2013). Por outro lado, em espécies filogeneticamente próximas a esta (*Tripsacum* sp. e *Z. luxurians*) e espécies de regiões alagadas como *O.*

sativa (Kawai *et al.*, 1999) e *S. lancifolia* (Shussler & Longstreth, 1996) e na espécie aquática *Juncus effusus* (Visser & Bögemann, 2006) o aerênquima se forma sem que haja algum tipo de indução externa. Morfologicamente, tecidos apresentando formação de aerênquima constitutiva ou induzida são idênticos.

A formação de aerênquima constitutivo em plantas de ambiente não alagado também foi reportada. Em diversas espécies, tais como *S. bicolor* (Promkhambut *et al.*, 2011), *Musa* spp (Aguilar *et al* 1998), *H. vulgare* (Garthwaite *et al* 2003) e cana-de-açúcar (Glaz *et al.*, 2004; Gilbert *et al.*, 2007; Tetsushi & Karim, 2007; Begum *et al.*, 2013; Leite, 2013), foi reportada a formação de aerênquima mesmo sem a presença de indutores externos, ainda que na presença desses indutores a área ocupada pelo aerênquima possa ter aumentado. Especificamente para cana-de-açúcar, na caracterização da formação e aerênquima formado por lisigenia em raízes, executada por Leite (2013) na variedade SP80-3280 cultivada em vasos aerados com vermiculita, a formação de aerênquima ocupou aproximadamente 40% do córtex radicular e se deu sem o requerimento de nenhum indutor ambiental.

Embora a participação de indutores ambientais como hipóxia, estresse nutricional ou etileno exógeno (aspectos discutidos a seguir) não seja requerida na formação constitutiva do aerênquima, não há um consenso na literatura a respeito da sua regulação endógena. Inicialmente foi sugerido que a formação constitutiva do aerênquima em *O. sativa* não envolvia o etileno, visto que tanto inibidores de sua síntese quanto de sua percepção não bloquearam este processo. Tampouco a aplicação de etileno exógeno foi capaz de influenciar a formação de aerênquima (Jackson *et al.*, 1985). Embora estudos posteriores apresentem ressalvas à independência do aerênquima constitutivo em relação ao etileno (Justin & Armstrong, 1991b; Yukiyoishi & Karahara, 2014), a regulação por este hormônio ainda é um aspecto controverso pelo menos em algumas variedades de *O. sativa* (Colmer *et al.*, 2006). Já estudos empregando a planta aquática *J. effusus* na análise da formação constitutiva do aerênquima também sugerem que esta se dê independentemente do etileno, visto que o bloqueio de sua percepção não altera a porosidade medida nas raízes (Visser & Bögemann, 2006).

Notavelmente, a formação constitutiva do aerênquima não significa que este se forme sempre de forma constante. Pelo contrário, experimentos com *S. bicolor*, *Musa* spp, arroz e cana-de-açúcar demonstraram que a extensão do aerênquima aumenta radial e/ou longitudinalmente na raiz em plantas alagadas (Justin & Armstrong, 1991b; Aguilar *et al.*, 1998; Promkhambut *et al.*, 2011; Begum *et al.*, 2013) ou mantidas sob deficiência de nitrogênio (Abiko & Obara, 2013). Logo, embora a natureza constitutiva manifeste-se na constante presença de aerênquima, sua intensidade pode ser modulada por estresses ambientais como alagamento e estresse nutricional.

3.1.1 Hipóxia

A hipóxia é o termo que descreve o *status* de O₂ de células ou tecidos em condições em que os níveis de este gás são reduzidos a ponto de impor um fator limitante para a produção de ATP através da fosforilação oxidativa. Isso ocorre em níveis de oxigênio inferiores a 50 mmol/m³. Dennis e colaboradores (2000) dividiram a resposta à hipóxia em três etapas. Após a percepção dos baixos níveis de oxigênio, o

primeiro estágio (0-4h) consiste na rápida indução ou ativação de componentes de vias de transdução de sinal, ativando o segundo estágio (4-24h), que consiste na adaptação metabólica, incluindo indução de vias de fermentação que garantem a manutenção dos níveis energéticos. É neste estágio que, segundo os autores, a via de biossíntese do etileno é ativada. O terceiro estágio (24-72h), decisivo para sobrevivência da planta, envolve a formação do aerênquima.

Foi proposto que as baixas concentrações de oxigênio são inicialmente percebidas pelo metabolismo mitocondrial (Subbaiah & Sachs, 2003) por meio da queda na produção de ATP, acidólise citoplasmática ou estresse oxidativo (Bailey-Serres & Chang, 2005). Várias alterações no metabolismo já foram descritas como resposta à hipóxia, dentre elas: aumento da expressão do gene da hemoglobina e de enzimas do metabolismo fermentativo (Silva-Cardenas *et al.*, 2003), liberação de citocromo c pela mitocôndria (Virolainen *et al.* 2002) e aumento na concentração de citosólica de Ca^{2+} (Drew, 1997). Análises de expressão gênica demonstraram o acúmulo de mRNAs de enzimas envolvidas em vias de detoxificação a espécies reativas de oxigênio (*reactive oxygen species*, ROS) em *Arabidopsis thaliana* em hipóxia (Klok, 2002). Além disso, plantas cultivadas nestas condições apresentaram atividade aumentada das enzimas da via de biossíntese de etileno (He *et al.*, 1994), elevando a concentração de precursores deste hormônio (Atwell *et al.* 1988) e também de sua forma ativa (He *et al.*, 1996a). Adicionalmente, foi observado o aumento do número de transcritos de receptores deste hormônio e de fatores de resposta ao etileno (*ethylene response factors*, ERFs) em raízes de *Gossypium hirsutum* em hipóxia (Liu *et al.*, 2005; Christianson *et al.*, 2010). Conforme será mais explorado adiante, a deflagração de vias de detoxificação de ROS e a biossíntese e transdução de sinal do etileno estão intimamente relacionadas à formação do aerênquima.

Nesse sentido, o desenvolvimento do aerênquima é claramente demonstrado como uma alternativa metabólica para evitar o metabolismo fermentativo e anóxia (Aschi-Smiti, 2003). Assim salienta-se mais uma vez o papel da raiz, visto que a percepção da hipóxia pode ocorrer se apenas uma parte deste órgão for exposta a esta condição, porém o monitoramento local provoca uma resposta ao longo de toda sua extensão (Malik *et al.*, 2003). Frequentemente a formação de aerênquima em raízes foi acessada por meio do cultivo de plantas em hipóxia, seja produzida por meio de alagamento, no cultivo em soluções nutritivas sem agitação durante o crescimento hidropônico ou em vasos contendo ar empobrecido em oxigênio (3-7%). Quaisquer que fossem as condições experimentais empregadas para reduzir a concentração de oxigênio em torno das raízes, as respostas foram similares. Em *Z. mays*, foi reportado o surgimento de grandes espaços de gás em raízes, ocupando até 50% da área do córtex (Dantas *et al.* 2001) e estendendo-se de duas camadas de células após a epiderme até três camadas de células antes da endoderme (Gunawardena *et al.* 2001a).

A resposta à hipóxia em plantas é um fenômeno complexo e ainda pouco conhecido, levando a alterações no padrão metabólico, morfológico e fisiológico. Respostas adaptativas à hipóxia consistem em rápida alteração na produção, localização e/ou ativação de fatores de transcrição (Bailey-Serres & Chang, 2005). Os principais reguladores que sinalizam para a elevação dos níveis endógenos de etileno ainda não foram identificados. Há provas indiretas que o sensor de oxigênio putativo é uma heme-proteína,

mas sua identidade molecular ainda permanece por ser determinada (Geigenberger, 2003). Embora a transdução de sinal que leva a uma rápida inibição da respiração e dos processos que consomem ATP e de várias mudanças na expressão gênica devam ainda ser elucidados, alguns fatores de transcrição e proteínas responsivas à cálcio parecem ter um papel fundamental neste processo. Dentre eles estão os fatores de transcrição das famílias MYB e NAC, cujos transcritos se acumulam em *A. thaliana* (Hoeren *et al.*, 1998), *T. aestivum* (Lee *et al.*, 2007), *Z. mays* (Zhang *et al.*, 2006) e *G. hirsutum* (Christianson *et al.*, 2010) em resposta à hipóxia. Inclusive transcritos relativos a estas proteínas encontram-se mais abundantes em *Z. mays* com formação de aerênquima induzida por alagamento (Rajhi *et al.*, 2011).

É interessante ressaltar que as alterações morfológicas observadas na raiz diante da escassez de nutrientes são semelhantes às relatadas para plantas alagadas (He *et al.*, 1992; He *et al.*, 1994). A presença de ROS, acidificação do pH, influxo de cálcio e sinalização por etileno também são aspectos comuns a plantas submetidas a ambos estresses (Bouranis *et al.*, 2006). Como resultado, disponibilidade de nutrientes e/ou sua restrição são sinais suficientes para alteração da expressão de genes e, portanto, candidatos potenciais para estimulação de algumas das respostas ao alagamento.

3.1.2 Estresse nutricional

Avaliando a formação de aerênquima em *Z. mays* em relação à disponibilidade de várias fontes de nitrogênio, Konings & Vershuren (1980) demonstraram que a integridade das células do córtex de raiz é diretamente proporcional à presença de nitrato e amônio, independente das condições de aeração às quais as plantas foram submetidas. De fato, plantas crescendo na ausência de nutrientes e de oxigênio apresentam desenvolvimento de cavidades do aerênquima cerca de 6 vezes maior quando comparadas a plantas pouco aeradas e cultivadas com ampla disponibilidade de nutrientes. Mais tarde outros trabalhos confirmaram esses resultados diante do bloqueio da formação de aerênquima em *Z. mays* e *P. vulgaris* em plantas em hipóxia e crescidas na presença de fontes não só de nitrogênio como também de fósforo, enxofre e potássio (Fan *et al.*, 2003; Postma & Lynch, 2011a; Siyiannis *et al.*, 2011). Também em *O. sativa* a área ocupada pelo aerênquima foi aumentada em plantas cultivadas em normóxia associada à escassez de fontes de fósforo (Lu *et al.*, 1999), embora o mesmo não tenha sido observado diante da omissão de fontes de nitrogênio (Yang *et al.*, 2012).

Já foram descritas alterações morfo-fisiológicas oriundas da presença de aerênquima em raízes de plantas cultivadas em concentrações sub-ótimas de nutrientes, como o aumento no número de raízes laterais em *Z. mays* deficiente em enxofre (Bouranis *et al.*, 2006) e reduzidas taxas respiratórias em *P. vulgaris* deficiente em fósforo (Fan *et al.*, 2003). Nas simulações realizadas por Postma & Lynch (2011a) em *Z. mays* foi demonstrado que a presença de aerênquima beneficiou substancialmente plantas submetidas à deficiência em fontes de N, P e K. Estes autores observaram que a formação de aerênquima estimulou o crescimento das plantas em 55 a 72% (em termos de peso seco), a depender do nutriente omitido no cultivo. No caso de plantas de *P. vulgaris* e *Z. mays* carentes em fontes de fósforo, o desenvolvimento do aerênquima levou a um aumento de 14 e 70% em peso seco, respectivamente. A

presença do aerênquima consiste em um traço adaptativo para maior aquisição de fósforo devido à mobilização deste nutriente como resultado da morte seletiva de células corticais na raiz, reduzindo os custos metabólicos para exploração do solo (Postma & Lynch, 2011b). Em outras palavras, a formação de aerênquima leva ao menor investimento de material orgânico para crescimento da raiz enquanto permite a manutenção da mesma área de absorção de nutrientes (Fagerstedt, 2010).

No desenvolvimento do aerênquima em condições de estresse nutricional a atividade da sintase de ácido 1-carboxílico-1-aminociclopropano (ACC), precursor de etileno, e a concentração de seu produto foram igualmente diminuídas, em oposição ao que se deu em hipóxia. Houve desenvolvimento do aerênquima em tratamentos carentes em fontes de N ou P, sendo que a menor adição de etileno testada induziu formação máxima de espaços de gás. Conforme esperado, as raízes cultivadas em solução nutritiva completa só apresentaram pontos de ruptura no córtex após indução com maiores doses de etileno (He *et al.*, 1994). Concluiu-se que em carência nutricional o tecido da raiz torna-se mais sensível a este hormônio enquanto as taxas de síntese de etileno declinam. É relevante salientar que o bloqueio da síntese deste hormônio impede a formação de aerênquima mesmo em condições de carência de nutrientes (He *et al.*, 1992), confirmando que o etileno é indispensável na formação do aerênquima, independente da natureza do efeito indutor do ambiente. Conforme sugerido por He e colaboradores (1992), tais mudanças na sensibilidade ao etileno em tecidos carentes de nutrientes podem indicar tanto uma alteração na concentração ou afinidade dos receptores ao etileno e/ou regulação positiva da cascata de transdução de sinal de ação do etileno diante do estresse nutricional.

4. Cascatas de sinalização e fatores de resposta ao etileno

O etileno é o fator comum na formação do aerênquima na maioria das plantas em que este evento foi investigado (Yamauchi *et al.*, 2013). O etileno é produzido a partir de S-adenosil-metionina (SAM), subsequentemente convertida em ACC pela sintase do ACC (ACS), sendo este substrato para a enzima subsequente, oxidase do ACC (ACO) (revisado por Rzewuski & Sauter, 2008), que produz a forma ativa do hormônio. Inibidores da biossíntese de etileno atuam interferindo na conversão de SAM a ACC (aminoetoxivinilglicina, AVG e ácido amino oxiaacético, AOA) ou de ACC a etileno (Co^{2+}) (Yu & Yang, 1979).

Uma vez produzido, o etileno é detectado por um receptor específico posicionado na membrana do retículo endoplasmático. Em *A. thaliana* a família conta com 5 membros, ETHYLENE RECEPTOR 1 e 2 (ETR1 e 2), ETHYLENE RESPONSE SENSOR 1 e 2 (ERS1 e 2) e ETHYLENE INSENSITIVE 4 (EIN4) (Bleecker *et al* 1998). A estrutura de ETR1 é complexa e foi a primeira a ser descrita. Este consiste em um homodímero estabilizado por pontes dissulfeto, apresentando de três a quatro domínios transmembrana N-terminais, que adotam uma conformação em hélice formando um bolso no qual se liga o etileno. Esta proteína assemelha-se à histidina-quinases de dois componentes de bactérias, apresentando um domínio hidrofóbico, posicionado na membrana do retículo endoplasmático, um domínio histidina quinase (sensor), responsável pela transferência de um grupo fosfato para um resíduo de aspartato no domínio regulador de resposta (revisado por Kieber, 1997). Também foram detectados

INTRODUÇÃO

domínios GAF na estrutura destes receptores, indicando que possivelmente este se liga a cAMP ou cGMP. Os demais componentes da família de receptores são estruturalmente similares, embora não apresentando todos domínios (Bleecker *et al*, 1998).

Outras proteínas atuam no reconhecimento e ligação do etileno a seu receptor. Uma delas é RESPONSIVE TO ANTAGONIST 1 (RAN1), que bombeia íons cobre no sítio de reconhecimento do receptor, evento requerido para ligação ao hormônio (Hall *et al.*, 2007). Neste ponto atuam os íons prata, conhecidos inibidores da percepção de etileno. Provavelmente por similaridades no tamanho, estado de oxidação e capacidade de interação com o etileno, íons prata competem com íons cobre, impedindo a interação do hormônio com seu receptor (Beyer *et al* 1976). Sais de prata foram por muito tempo empregados como inibidores da percepção deste hormônio, entretanto seu emprego tem sido gradativamente abandonado em razão de efeitos adversos e toxicidade (Drew *et al.*, 1981).

Outro antagonista, 1-metilciclopropeno (1-MCP), também atua de modo a bloquear a percepção do etileno. Embora seja conhecido que o receptor tem maior afinidade pelo 1-MCP que pelo etileno (Jiang *et al* 1999) e que a interação com o antagonista se dá de forma irreversível, resultando da endocitose e síntese de novos receptores (Blackenship *et al* 2003), pouco se sabe sobre o funcionamento do 1-MCP em si.

Na ausência do etileno estes receptores estão permanentemente inibindo a via de sinalização à jusante. Tais receptores estruturam um complexo juntamente com CONSTITUTIVE TRIPLE RESPONSE 1 (CTR1), cuja região C-terminal apresenta assinaturas de quinases de Ser/Thr e apresentando sequência primária similar a Raf quinases (Kieber *et al*, 1993). Uma vez que o etileno se liga ao seu receptor, este último altera sua conformação tridimensional e desliga a atividade repressora do complexo receptor-CTR1, de-reprimindo as respostas ao hormônio (Fig. 1) (Rzewuski & Sauter, 2008; Stepanova & Alonso, 2009). Isso resulta na ativação de demais proteínas envolvidas na transdução de sinal à jusante do receptor, culminando na regulação da expressão de ETHYLENE RESPONSE FACTOR 1 (ERF1), fatores de transcrição da família AP2/EREBP que controlam a expressão de genes regulados pelo etileno (Solano *et al* 1998). Estes e outros elementos de menor relevância a este trabalho estão ilustrados na Figura 1.

Juntamente com fatores de transcrição do tipo MYB-(R1)R2R3 e bHLH (*basic helix-loop-helix*), a superfamília AP2/ERF (às vezes referida como AP2/EREBP, de APETALA2/*ethylene responsive element binding protein*) é uma das maiores dentre fatores de transcrição de plantas (Ohme-Takagi e Shinshi, 1995; Riechmann *et al.*, 2000), apresentando 147 membros em *A. thaliana* (Nakano *et al* 2006). Essa superfamília é definida pela presença de ao menos um domínio AP2/ERF (às vezes referido somente como AP2), pela primeira vez identificada em *A. thaliana* como envolvido no desenvolvimento floral (Jofuku *et al.*, 1994).

Figura 1. Representação esquemática da via de sinalização do etileno. Receptores na membrana do retículo endoplasmático apresentam estrutura similar às quinases de dois componentes de bactérias. Receptores e CTR1 funcionam como reguladores negativos na ausência do hormônio. A ligação do etileno ao receptor se dá mediante interação de íons cobre disponibilizados por RAN1. Desse modo os receptores transitam para um estado inativo, em um processo mediado por RTE1 (REVERSION TO ETHYLENE SENSITIVITY 1). EIN2 e 3 se acumulam mediante exposição a etileno em razão da regulação negativa de ETP1 e 2 (que ubiquitinam EIN2) e EBF1 e 2 (que ubiquitinam EIN3). Tal regulação negativa de ETP1 e 2 (EIN2 TARGETING PROTEIN 1 e 2) e EBF 1 e 2 (EIN3-BINDING F-box 1 e 2) é mediada pela exoribonuclease EIN5. EIN3 e EIL1 regulam a expressão de genes alvo, dentre eles ERF1. Retirado de Stepanova & Alonso (2009).

O domínio AP2/ERF pode ocorrer em três subfamílias. A sub-família AP2 contém dois domínios AP2/ERF, enquanto na sub-família ERF este domínio aparece só uma vez. Por fim, a menor sub-família compreende os fatores de transcrição do tipo RAV, que apresentam, em adição a um único domínio AP2/ERF, também um domínio B3, conservado em outras famílias de fatores de transcrição de plantas (Nakano *et al.*, 2006). Estes fatores contêm um domínio de 60 a 70 aminoácidos conservados que exerce a função de ligação ao DNA. Este assume uma estrutura secundária composta por uma α hélice e 3 folhas β (Cao *et al.*, 2006). A folha β mais C-terminal liga-se à sequência alvo no promotor dos genes regulados de forma monomérica (Allen *et al.*, 1998). Tem sido observada a ligação de domínios AP2 a um sítio de ligação GCC box, com sequência GCCGCC altamente conservada para modulação de genes de resposta ao etileno (Cao *et al.*, 2006) e denominado GCCCORE (Chakravarthy *et al.*, 2003). Já fatores de transcrição da subfamília RAV se ligam a uma sequência CAACA (RAV1AAT) e CACCTG (RAV1BAT) por meio dos domínios AP2 e B3, respectivamente (Kagaya *et al.*, 1999).

A caracterização funcional de fatores de resposta a etileno apontou para papéis na regulação das respostas a estresses bióticos e abióticos, além de atuação em processos como acumulação de ceras, identidade de meristema floral, desenvolvimento de pecíolos e biossíntese de alcalóides (Nakano *et al.*, 2006). No caso de estresses abióticos, embora tenha sido mostrado o aumento no número de transcritos de ERFs em *A. thaliana*, *O. sativa* e *G. hirsutum* submetidos a tratamentos de hipóxia, seca, frio e estresse salino (Fujimoto *et al.*, 2000; Cao *et al.*, 2006; Jin & Liu, 2008; Yang *et al.* 2011), há escassos relatos de quais seriam os possíveis genes regulados. Um deles codifica para a proteína ADH1 de *A. thaliana*, cuja expressão é aumentada pelos ERFs RAP2.2 (Hinz *et al.* 2010) e AtERF73 (Yang *et al.* 2011) em plantas cultivadas em hipóxia. No caso específico de AtERF73, tanto o bloqueio da biossíntese como da sinalização por etileno impedem o aumento nos níveis de transcritos referentes ao fator de transcrição e ADH1. Isso sugere um elo entre a sinalização por etileno e o metabolismo fermentativo na resposta à hipóxia, favorecendo a sobrevivência da planta (Yang *et al.*, 2011).

Já no que diz respeito à sinalização em estresse biótico, a expressão de aumentada de AtERF1, 2, 5 e 14 em *A. thaliana* foi relacionada ao acúmulo de transcritos referentes a defensinas e uma quitinase (Fujimoto *et al.*, 2000; Oñate-Sánchez *et al.*, 2007). Notavelmente, foi demonstrado que as proteínas ERF1 de *A. thaliana* e EREBP-2 de *Nicotiana benthamiana* se ligam ao GCCCORE no promotor dos genes b-Chi e Gln2, respectivamente, ambas codificando para quitinases (Ohme-Takagi & Shinshi, 1995; Solano *et al.* 1998). Este resultado representa uma evidência que fatores de transcrição da família ERF podem atuar na regulação da expressão de glicosil hidrolases.

No que diz respeito ao papel desempenhado por fatores de transcrição desta família em processos que envolvem degradação de parede celular vegetal, salientam-se aqueles relacionados ao amadurecimento de frutos. Em *S. lycopersicum* a queda nos níveis de transcritos do ERF SIAP2a por RNAi resultou no aumento da expressão de PG2A, que codifica para uma endopoligalacturonase, sugerindo que este ERF atue como regulador negativo durante o amadurecimento (Chung *et al.* 2010). A fim de explorar a relação de ERFs com a expressão de genes relacionados ao amadurecimento em *Actinidia deliciosa*, os promotores de genes de enzimas modificadoras de parede celular foram fusionados a genes repórteres. Dentre eles foi identificado o promotor da AdXET5, que codifica para uma xiloglucano transglicosilase/hidrolase, cuja regulação negativa por AdERF9 foi reportada (Yin *et al.* 2010).

Ainda no que concerne ao amadurecimento de frutos, o papel fundamental não só do etileno mas também da auxina por trás da indução de ERFs já foi reportado (Given *et al.*, 1988; Jones *et al.* 2002; Balbi & Lomax, 2003), sobretudo no que diz respeito às modificações de parede celular (Davies *et al.*, 1997; Sozzi *et al.* 2000; Aharoni *et al.* 2002). Analisando o perfil transcricional de sete ERFs em dois cultivares distintos de *Prunus domestica* durante o amadurecimento, El-Sharkawy e colaboradores (2009) concluíram que a diversidade de padrões de expressão se deve não só às variações na concentração de etileno, como também de outros hormônios. Estes autores sugerem que a auxina é responsável pela atividade de divisão e expansão celular em fases do amadurecimento em que o acúmulo de ERFs foi alto mesmo diante de baixos níveis de etileno. De fato, outros trabalhos demonstram acúmulo de transcritos relativos a ERFs e

também de proteínas de resposta à auxina em plantas submetidas à hipóxia (Nanjo *et al.*, 2011; Rajhi *et al.*, 2011). Sendo assim, essas proteínas poderiam se acumular tanto em resposta ao etileno como em resposta à ação combinatória de ambos hormônios. Como consequência, o balanço hormonal seria então o responsável pela indução da expressão de ERFs em etapas de intensas expansão e divisão celular. Entretanto, a relevância de ERFs na formação do aerênquima ainda não foi confirmada.

4.1 Efeito do etileno na formação do aerênquima

As vias de sinalização do etileno que culminam na formação do aerênquima ainda são pouco conhecidas, embora alguns elementos-chave já tenham sido identificados em trabalhos empregando agonistas e antagonistas da formação de aerênquima (Fig. 2) (He *et al.*, 1996a).

Em condições hipóxia o etileno se acumula nos tecidos, ativando uma via de transdução de sinal que envolve proteínas G, cujo provável alvo é fosfolipase C (PLC). Esta proteína hidrolisa fosfatidilinositol 4,5-bisfosfato, gerando dois segundos mensageiros: inositol 1,4,5-trifosfato, IP3, indutor de liberação de Ca^{2+} , e 1,2-diacylglicerol, ativador de quinases C (Drew *et al.*, 2000; Shiono *et al.*, 2008). Por meio de proteínas quinase os processos de degradação de parede celular e morte celular programada são ativados (Fig. 2).

Embora as precisas vias pelas quais genes alvo na degradação de parede celular são expressos permaneçam desconhecidas (Fig. 2), é importante notar que o processo de formação de aerênquima é indissociável da morte celular programada, ambos requerendo a sinalização por etileno e a participação do cálcio como importante segundo mensageiro (He *et al.*, 1996b).

Figura 2. Processos supostamente envolvidos na formação de aerênquima lisígeno. ACC: ácido 1-carboxílico-1-aminociclopropano; AdoMet: adenosil metionina; CTR1: CONSTITUTIVE TRIPLE RESPONSE 1, regulador negativo da resposta ao etileno; EIN2: ETHYLENE INSENSITIVE 2, repressor da ubiquitinação dos fatores de transcrição EIN3/EIL3 (ETHYLENE INSENSITIVE 3 e EIN-LIKE) por EBF1/2 (EIN3-BINDING F-box 1 e 2); PLC: fosfolipase C; IP3: inositol 1,4,5-trifosfato; CaCM: calmodulina dependente de cálcio; XTH: xiloglucano transglicosilase/hirolase. Para explicação completa, vide texto da seção 4. Fonte: Shiono e colaboradores (2008).

5. Morte celular programada

A formação de aerênquima em raízes é um importante modelo de estudo da morte celular programada (Kawai *et al.*, 1999; Gunawardena *et al.*, 2001a), sendo etapa obrigatória da formação lisígena. Em *O. sativa* e *Z. mays* um dos primeiros sinais de lise celular é a acidificação do citoplasma. (Kawai *et al.* 1999; Siyiannis *et al.*, 2009). Além disso, o influxo de cálcio e a liberação de enzimas hidrolíticas supostamente aprisionadas no vacúolo parecem ser as etapas indispensáveis para provocar a morte das células sensibilizadas pelo etileno. O Ca^{2+} tem sido apontado como um provável segundo mensageiro na transdução de sinal que ligam os mecanismos sensores de hipóxia ao processamento dos restos celulares ao final da formação do aerênquima (Fig. 2), sendo um dos aspectos mais abordados da morte celular programada. A detecção de cálcio *in situ*, bem como o acúmulo de transcritos de proteínas ligantes a este íon ao longo da formação do aerênquima em *Z. mays* (Bouranis *et al.*, 2006; Rahji *et al.*, 2011) corroboram esta observação.

O íon cálcio é requerido para renaturação e atividade proteolítica de cisteíno-proteases produzida em condições de baixa oxigenação. Esta classe de enzimas é descrita como participante de vários processos nos quais ocorre a morte celular programada (Jones, 2000), estando provavelmente armazenadas em vacúolos. Sabe-se que estes são locais de estocagem de enzimas hidrolíticas como proteases e nucleases em tecidos vegetativos (Lam, 2005). Um ponto importante de execução da morte celular programada é o rompimento destes vacúolos, resultante da atividade de uma família de cisteíno-proteases caspases-like chamadas VPEs (do inglês *vacuolar processing enzymes*). Embora o alvo de VPE durante a deflagração da morte celular seja ainda desconhecido, houve uma dramática redução no colapso do vacúolo em plantas após silenciamento desta protease (Hatsugai *et al.* 2006), o que pode ocorrer também após tratamento com inibidores de caspases (Hiraiwa *et al.*, 1999).

Em *Z. mays* foi evidenciado que, nas etapas iniciais da formação do aerênquima, vacúolo e tonoplasto tornam-se elétron-densos, enquanto organelas como mitocôndrias, retículos endoplasmáticos e complexos de Golgi mantiveram-se intactos. Numerosas vesículas, possivelmente secretoras, encontravam-se associadas ao Golgi. Por fim foi observada a perda definitiva do conteúdo das células, e perda de integridade da parede celular e do vacúolo (Campbell & Drew, 1983; Gunawardena *et al.*, 2001a; Leite, 2013). Além disso, a morte celular programada pode ser acompanhada pelo monitoramento de células com DNA nuclear degradado, conforme observado a partir do segundo centímetro de raiz de cana-de-açúcar após o ápice (Leite, 2013).

No que diz respeito ao ataque à parede celular nas etapas mais avançadas da formação de aerênquima e morte celular programada, micrografias eletrônicas evidenciam extensões de parede celular menos elétron densas (Campbell & Drew 1983), aparentando enrugamento (Lenochová *et al.*, 2009). Restaram, ao final do processo, remanescentes de parede celular circundando os espaços de gás (Gunawardena *et al.* 2001b; Leite, 2013) e cordões de células colapsadas cruzando o córtex (Leite, 2013). Além disso, acredita-se que a modificação de parede seja operada por mecanismos da própria célula que se encontra em morte celular programada (Gunawardena *et al.* 2001b; Grandis, 2015). Coincidentemente,

as modificações observadas por micrografias eletrônicas na parede celular evidenciaram também o acúmulo de organelas e pequenas vesículas próximas às bordas das células (Gunawardena *et al* 2011a), podendo as mesmas consistirem em pontos de acúmulos de enzimas hidrolíticas. É, portanto, razoável assumir que um arranjo ordenado de enzimas de degradação da parede é requerido.

6. Degradação da parede celular na formação do aerênquima

Um conjunto limitado de trabalhos enfocaram a identificação de glicosil hidrolases durante a formação de aerênquima, a maioria empregando plantas cultivadas em condições de hipóxia (He *et al.*, 1996b; Rahji *et al.*, 2011). Uma enzima que é descrita como sendo intimamente envolvida na modificação de parede celular durante a formação do aerênquima é a endo- β -1,4-endoglicanase, cuja atividade aumenta com a elevação dos níveis endógenos de etileno (Bragina *et al.*, 2003). Em contrapartida, a atividade desta enzima diminui com a utilização de bloqueadores da morte celular (He *et al.*, 1996a) ou da biossíntese de etileno (He *et al.*, 1994). No entanto, este efeito pode ser revertido pela exposição simultânea a etileno exógeno, indicando o papel essencial deste hormônio na regulação positiva desta enzima. Um outro aspecto que reforça o papel indutor do etileno na degradação de parede durante o desenvolvimento do aerênquima é o fato do aumento da atividade da ACC sintase preceder o da atividade celulásica em regiões próximas ao meristema apical (He *et al.*, 1992).

Segundo a literatura hemiceluloses também são atacadas durante a formação de aerênquima. Isto é evidente à medida que a atividade de xilanase foi observada em extratos protéicos de raiz de *Z. mays* após alagamento (Bragina *et al.*, 2001; 2003). Outra enzima potencialmente associada à formação do aerênquima é a xiloglucano transglicosilase/hidrolase (XTH), tendo sido identificada durante a formação de aerênquima em raízes de *Z. mays* após alagamento (Saab & Sachs, 1996; Rajhi *et al.*, 2011). Embora a participação desta enzima tenha sido inferida por meio do acúmulo de transcritos em plantas alagadas e sua atividade nunca tenha sido comprovada, a indução por etileno também foi observada visto que o bloqueio da percepção deste hormônio por 1-MCP leva à queda na quantidade deste mRNA (Rajhi *et al.*, 2011).

Por fim, o ataque às pectinas durante a formação de aerênquima foi sugerido em *Z. mays* mediante a detecção dos transcritos de endopoligalacturonases (Rahji *et al.*, 2011), da atividade de pectinase (Bragina *et al.*, 2003) e alteração no padrão de ligação de anticorpos anti-pectina na lamela média (Gunawardena *et al.*, 2001b). O uso de anticorpos específicos para pectina metilesterificada e não-metilesterificada demonstram que as alterações na parede celular ocorreram em fases muito precoces da morte celular programada. Estas mudanças ocorrem na fase em que vesículas abundantes estão presentes na membrana plasmática e associadas a várias organelas (Gunawardena *et al.*, 2001a).

Em cana-de-açúcar, embora a degradação de celulose e xilanos não tenha sido observada, a degradação de pectinas da lamela média foi claramente identificada. A ligação do anticorpo anti-homogalacturonano demonstrou um padrão de marcação pontilhado na lamela média na região mais próxima do ápice radicular, na qual o aerênquima ainda não estava formado. A intensidade da marcação

foi reduzida conforme as células se expandiram e posteriormente o aerênquima se desenvolveu, indicando a degradação da pectina presente na lamela média. Além disso, a marcação do córtex de raízes de cana-de-açúcar com o anticorpo anti- β -glucano indicou a remoção deste polissacarídeo durante o processo de abertura de espaços de gás no córtex (Leite, 2013). No que diz respeito às proteínas identificadas durante este processo destacam-se pectinases, arabinofuranosidases, manosidases, β -1,3-glucanases, liquenases, xilosidases, inibidores de xilanase e celulases (Grandis, 2015). Juntamente com a caracterização bioquímica (Leite, 2013) a identificação destas proteínas corrobora a hidrólise do homogalacturonano na lamela média, com perda de adesão celular.

Sugere-se que posteriormente outros componentes da parede celular sejam atacados, dentre eles a celulose e o arabinoxilano. Possivelmente regiões amorfas da celulose são hidrolisadas pela endoglucanase e β -glucosidase enquanto o arabinoxilano é desramificado pela arabinofuranosidase, graças a uma das atividades previstas para esta enzima. A interação entre as cadeias de xilano e as microfibrilas de celulose (sobretudo regiões de celulose cristalina remanescentes) resultaria num compósito altamente resistente à hidrólise que compõe os restos de parede celular observados ao final da formação do aerênquima (Leite *et al.*, em preparação). Por fim, o sequenciamento do transcriptoma de raízes em formação de aerênquima resultou na identificação de uma média de 560 transcritos a cada centímetro de raiz analisado (Tavares *et al.*, em preparação). Em adição às atividades enzimáticas previamente reportadas, transcritos referentes a expansinas, transglicosilases/hidrolases de xiloglucano, β -galactosidases e feruloil esterases foram identificados, corroborando a alta atividade de expansão celular e sugerindo redução das interligações do arabinoxilano. Os principais elementos relevantes para a degradação da parede celular e morte celular programada descritos no aerênquima de cana-de-açúcar estão resumidos na Figura 3.

Ainda que existam diferenças entre as descrições dos eventos iniciais na formação de aerênquima entre *O. sativa* e *Z. mays* e provavelmente entre ambas espécies e cana-de-açúcar, parece não haver dúvida que a degradação da parede celular se inicia com a separação na lamela média, seguido pelo colapso das células e da parede que as envolve, deixando os espaços de gás (Evans, 2003). Outro aspecto relevante ainda não explorado em cana-de-açúcar é a regulação por etileno, que em *Z. mays* influencia glicosil hidrolases tanto no âmbito transcricional como na atividade enzimática durante a formação no aerênquima (Bragina *et al.*, 2003; Rahji *et al.*, 2011).

Assim como todo processo biológico que tem como alvo a hidrólise de parede celular, a formação do aerênquima parece proceder de forma similar no que diz respeito à sequência na qual os polissacarídeos são atacados (Grandis *et al.*, 2014; Tavares *et al.*, 2015). Visto que a recalcitrância que a biomassa vegetal impõe à hidrólise é um resultado direto de como os polímeros se agregam e mantêm interações entre si estruturando a parede celular (McCann & Carpita, 2008; Buckeridge *et al.*, 2015), a compreensão de mecanismos endógenos de degradação ou modificação desta estrutura apresenta grande potencial biotecnológico, conforme discutido anteriormente.

Figura 3. Elementos chave identificados para elaboração do modelo de degradação de parede celular na formação do aerênquima em raízes de cana-de-açúcar. Enzimas como expansinas e XTH possivelmente atuam reduzindo interações entre xiloglucano e celulose, enquanto a endopoligalacturonase ataca o HG após remoção de grupamentos acetil pela pectina acetil esterase. Conjuntos de células expandidas e separação celular decorrente da hidrólise de HG na lamela média foram eventos observados *in situ*. O aumento da porosidade possivelmente se dá pela degradação de RG, cujas cadeias laterais de arabinano e arabinogalactano podem ser removidas pela α -arabinofuranosidase e β -galactosidase, respectivamente. Por fim, a hidrólise de β -glucano ocorreria mediante ação de liquenases, enquanto feruloil esterase e α -arabinofuranosidase reduziriam interligações e ramificações do AX. A proteína inibidora de xilanases provavelmente atua impedindo a despolimerização da cadeia principal do AX, que é mantida em um suposto compósito em interação com a celulose. Tais compósitos são provavelmente responsáveis pela redução da sacarificação observada após decorridas as atividades enzimáticas supracitadas. Especula-se que a atividade de celulases (endoglucanase e β -glucosidase) se dê sobre a celulose amorfa, o que também corroboraria a hipótese relacionada à queda na sacarificação. Resultados oriundos dos trabalhos de Leite (2013); Grandis e colaboradores (2014); Grandis (2015); Leite e colaboradores (em preparação); Tavares e colaboradores (em preparação).

Nesse sentido o estudo da formação do aerênquima consiste em uma valiosa oportunidade para a prospecção de glicosil hidrolases de plantas bem como sua possível aplicação voltada à produção de etanol de segunda geração. Além disso, o enfoque na regulação de glicosil hidrolases possibilita a compreensão dos mecanismos pelos quais a planta opera a degradação de sua própria parede celular, jogando luz sobre pontos relacionados à recalcitrância à hidrólise, aspecto este considerado o grande entrave para a produção de combustíveis de segunda geração a partir de biomassa vegetal (Himmel *et al.*, 2009). Um método inteligente de ataque à biomassa vegetal consiste no *design* de etapas de desmontagem sequencial da parede celular adaptadas a partir de processos desempenhados em plantas (Tavares *et al.*, 2015). O sucesso desta estratégia tem o potencial de representar um importante salto no que diz respeito à exequibilidade da hidrólise eficiente de biomassa lignocelulósica para bioenergia.

7. Viabilizando o pré-tratamento biológico em cana-de-açúcar

O pré-tratamento biológico é um conjunto de técnicas cujo principal objetivo é diminuir a acessibilidade da parede celular da biomassa empregada na produção de etanol 2G. Os procedimentos contidos nesta estratégia atuam em conjunto com técnicas de pré-tratamento químicas e físicas anteriores à hidrólise enzimática.

Em frutos foi demonstrado que o ataque à parede celular se inicia com a degradação da pectina e a desestabilização das pontes de hidrogênio que mantêm em interação xiloglucano e celulose. A primeira etapa mencionada é conduzida por pectinases e a segunda por expansinas. Diante da ocorrência da conservação destes mecanismos e desta ordem em outros sistemas (como por exemplo a formação de aerênquimas) ambos os processos foram categorizados nos módulos de separação e expansão celular da degradação de parede, que dependem de pectinases e expansinas (Grandis *et al.*, 2014). A atuação de ambos conjuntos de proteínas é chave para as etapas iniciais de degradação de parede e, no contexto biotecnológico, tem o potencial de predispor a parede celular à hidrólise (Tavares *et al.*, 2015).

O terceiro e quarto módulos de ataque à parede celular (Grandis *et al.*, 2014) consistem no ataque à hemiceluloses e celulose, respectivamente. A compreensão do mecanismo de mobilização do xiloglucano em *Hymenaea courbaril* L levou à identificação dos passos sequenciais requeridos para hidrólise deste polissacarídeo. Xiloglucano transglicosilase/hidrolase atuando em conjunto com β -galactosidase hidrolisam a supracitada hemicelulose, liberando oligossacarídeos. Estes por sua vez são alvo de α -xilosidases e β -glucosidases (Buckeridge *et al.*, 2000). Já o ataque à celulose em processos endógenos parece se dar em pontos específicos da molécula que, quando hidrolisados, resultam em aumento da cristalinidade e, portanto, também da recalcitrância (O'Donoghue *et al.*, 1994). A caracterização profunda de processos de degradação de parede celular leva inclusive à identificação de processos que, no contexto biotecnológico, devem ser evitados em razão da queda produzida na sacarificação. Nesse sentido, a descrição de vias regulatórias é indispensável, tornando possível a manipulação de determinados eventos.

A produção de plantas transgênicas de cana de açúcar apresentando biomassa com reduzida recalcitrância depende, em primeira instância, da seleção inteligente do gene alvo. Genes que atuam em processos chave de aumento da acessibilidade da parede celular a glicosil hidrolases são naturalmente fortes candidatos. Os genes que codificam para a maioria das enzimas em cana-de-açúcar e que participam nestes processos de degradação endógena de parede celular precisariam ser inicialmente sequenciados e caracterizados *in silico*. Por outro lado, a caracterização bioquímica de glicosil hidrolases de plantas pode também revelar aspectos funcionais surpreendentes que façam jus à maior complexidade estrutural de enzimas de plantas quando comparadas às de microrganismos. Almeja-se não só o uso das mesmas em coquetéis enzimáticos, complementando os já existentes compostos por enzimas de microrganismos, mas também enriquecer a literatura no que diz respeito ao padrão de expressão de genes hom(e)ólogos e quais fatores estruturais e/ou transcricionais estão envolvidos. Esta segunda finalidade é

especialmente interessante quando se leva em consideração a alta ploidia da cana-de-açúcar e o seu ainda desconhecido nível de complexidade subjacente à regulação da expressão gênica.

Entretanto, enquanto dados oriundos da formação do aerênquima em cana-de-açúcar e da estrutura e composição das paredes celulares desta planta encorajam o investimento em pesquisas em Fisiologia Vegetal, a alta complexidade genômica e a escassez de dados de sequências de DNA tornam o trabalho com esta planta desafiador, lento e muitas vezes limitado para a Biologia Molecular. As variedades de cana de açúcar atualmente empregadas na indústria sucroalcooleira no Brasil são plantas híbridas resultado de intenso melhoramento genético. São plantas poliplóides aneuplóides, com contribuição desigual dos genomas de parentais. *S. officinarum*, espécie com alto potencial de acúmulo de sacarose e genoma básico (1x) de 930 Mb, que foi repetidas vezes cruzada com *S. spontaneum*, gramínea com alta resistência a estresses bióticos e abióticos, cujo genoma básico (1x) tem 750 Mb. Como resultado, o genoma básico (1x) de cana-de-açúcar apresenta 1 Gb, preservando 15–25% dos cromossomos de *S. spontaneum* e 60–70% de *S. officinarum*. Além disso, os cultivares modernos apresentam 5–10% dos cromossomos recombinantes (D'Hont, 2005).

Os procedimentos de melhoramento genético pelos quais a cana-de-açúcar foi submetida sempre tiveram a finalidade de aumentar o conteúdo de açúcar, associado a resistência a doenças. Entretanto, foi observado que os cultivares mais recentemente produzidos apresentam genótipos cada vez menos diversificados. Como consequência, o ganho em produção de açúcar tem sido cada vez menor (Dal-Bianco *et al.*, 2012). Dada a relevância da cana de açúcar no cenário sucroalcooleiro brasileiro, aliada à necessidade de produção de biocombustíveis, faz-se imperativo o investimento em outras estratégias que não unicamente o melhoramento genético. Na ausência de um genoma sequenciado, a prospecção de genes de glicosil hidrolases e a identificação de proteínas que regulam sua expressão depende do sequenciamento das regiões alvo. A busca por fatores de transcrição e vias regulatórias subjacentes à degradação da parede celular poderá ser auxiliada pela identificação de sítios de ligação em promotores de genes que codificam para glicosil hidrolases, cujo papel é crítico em processos como é o caso da formação de aerênquima em raízes cana-de-açúcar. O sequenciamento destes genes alvo, a descrição de papéis regulatórios e a investigação filogenética têm o potencial de preencher uma grande lacuna no que diz respeito à dinâmica fisiológica atrelada à estrutura do genoma desta planta.

No contexto da ideologia subjacente ao pré-tratamento biológico proposto pelo nosso grupo (Grandis *et al.* 2014; Tavares *et al.*, 2015), as técnicas de aproveitamento da biomassa vegetal poderiam usar mecanismos de degradação endógena de parede celular como procedimento guia. Isso significa dizer que os processos celulares nos quais a parede é degradada precisam ser melhor compreendidos em suas instâncias morfo-fisiológicas e bioquímicas. Em cana-de-açúcar isso deve se dar de modo acoplado à dinâmica da expressão gênica que por sua vez está relacionada à organização do genoma.

OBJETIVOS

Objetivo geral

Identificar elementos chave na regulação do metabolismo de degradação da parede celular em cana-de-açúcar durante a formação do aerênquima

Objetivos específicos

- Determinar o perfil de formação de aerênquima lisígeno em resposta a estresse nutricional e 1-metilciclopropeno;
- Monitorar a concentração de auxina e etileno nas amostras de raiz tratadas com 1-metilciclopropeno;
- Analisar o perfil diferencial de expressão de genes de fatores de transcrição, enzimas de degradação e genes relacionados à morte celular programada em resposta às condições experimentais;
- Sequenciar genes de glicosil hidrolases e identificar elementos *cis* na região promotora que sejam possivelmente reconhecidos por fatores de transcrição também sequenciados;
- Caracterizar as sequências obtidas quanto à sintenia em relação às espécies aparentadas;
- Caracterizar *in silico* as proteínas preditas
- Avaliar a interação dos fatores de transcrição com os promotores de glicosil hidrolases em sistema heterólogo;
- Elaborar um modelo de regulação gênica da degradação de parede envolvendo proteínas reguladoras e seus alvos

A proposta das primeiras etapas do projeto residiu na realização de experimentos visando a perturbação da formação do aerênquima em raízes de cana-de-açúcar (Partes A e B; Fig 4) de modo investigar a regulação hormonal e identificar genes-alvo diferencialmente expressos para sequenciamento. Estas etapas foram conduzidas em colaboração com o Dr. Eduardo Purgatto (FCF/USP) e a Dra. Glaucia Mendes Souza (IQ/USP). Sendo assim, dentre genes de glicosil hidrolases e fatores de transcrição, os alvos selecionados de acordo com seu perfil transcricional foram sequenciados (Parte C; Fig. 4) em colaboração com o grupo da Dra. Marie-Anne van Sluys (IB/USP). As sequências foram analisadas quanto à sintenia e estrutura dos genes em relação à *Sorgum bicolor* (Parte D; Fig. 4) e em seguida foram selecionadas uma sequência de cada gene hom(e)ólogo de acordo com o mapeamento empregando o transcriptoma das raízes de cana-de-açúcar (Parte E; Fig. 4). A etapa final do projeto (Parte F; Fig. 4) envolveu a validação da interação entre fatores de transcrição que regulam a expressão de genes que codificam para glicosil hidrolases. Estes experimentos foram realizados durante o estágio sanduíche no *Laboratoire de Recherche en Sciences Végétales* em Toulouse – França, sob orientação da Dra. Jacqueline Grima-Pettenati. Um fluxograma contendo as estratégias básicas de cada uma destas partes e a maneira como elas se conectam neste projeto está disponível na Figura 4.

Figura 4. Fluxograma esquematizando as etapas (letras em círculos vermelhos) que compõem este trabalho e as principais técnicas empregadas para sua execução. TF: fator de transcrição; GH: glicosil hidrolase; 1-MCP: 1-metilciclopropano.

Cultivo de cana-de-açúcar e coleta de raiz

Em todos os experimentos foi empregada a variedade SP80-3280 de cana-de-açúcar coletada no sítio Victória, no município de Piracicaba-SP. Os colmos coletados foram seccionados em toletes de cerca de 10 cm contendo uma gema lateral e plantados em bandejas contendo vermiculita. Após 10-14 dias, as plantas foram transferidas para vasos contendo vermiculita com 40g de NPK (30:20:30). As plantas permaneceram nessas condições por 2 semanas até o início do tratamento com solução nutritiva ou inibidor de percepção de etileno (1-metilciclopropeno, 1-MCP), conforme descrito nas seções seguintes. No caso de plantas do tratamento com soluções nutritivas a rega se deu semanalmente com 100 mL de solução e no caso do experimento com 1-MCP o inibidor foi diluído em 5 mL de água aplicado no fundo do vaso a cada 48h.

Após aproximadamente 3 meses de cultivo, as raízes dos perfilhos (Fig. 5A) foram coletadas e os cinco primeiros centímetros foram cortados em segmentos de 1 cm cada (Fig. 5B), conforme previamente padronizado em nosso laboratório (Leite, 2013). Tais segmentos (S) serão referidos como S1 a S5 ao longo deste trabalho. Exceto em amostras destinadas a dosagens de etileno e microscopia de luz (descritas nos itens 5 e 3 desta metodologia, respectivamente), seções de raiz pertencentes ao mesmo segmento foram agrupadas em tubos de 50 mL contendo nitrogênio líquido. Os tubos contendo o material coletado foram armazenados a -80°C até o momento do uso.

Figura 5. Raízes de cana-de-açúcar. (A) Sistema radicular de secção de colmo de cana-de-açúcar (tolete) do qual emergem os perfilhos. As raízes usadas no presente trabalho foram as raízes dos perfilhos. Fonte: Smith, 2005. (B) Divisão das raízes de cana-de-açúcar que emergem do perfilho em 5 segmentos de um centímetro cada, a partir do meristema radicular (à esquerda).

PARTES A & B: Cultivo de cana-de-açúcar em estresse nutricional e na presença de 1-MCP**1. Cultivo de cana-de-açúcar irrigada com distintas soluções de nutrientes**

Duas semanas após o plantio em vasos, as plantas foram divididas em 10 grupos experimentais, com 10 plantas cada (Tabela 1). Por três meses, cada um dos grupos recebeu semanalmente 100 mL de solução nutritiva de Hoagland (Hoagland & Arnon, 1950, com modificações), cuja composição padrão usada em nosso laboratório é NH_4NO_3 10 mM, KCl 5 mM, KH_2PO_4 1 mM, $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 5 mM, Fe-EDTA 20 μM , H_3BO_3 25 μM , $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ 2 μM , ZnCl_2 2 μM , CuCl_2 0,5 μM e $\text{H}_2\text{MoO}_4 \cdot \text{H}_2\text{O}$ 0,5 μM . Os tratamentos receberam soluções nutritivas com distintas concentrações de nitrato de amônio (NH_4NO_3) (0, 2, 10, 20, 40 e 80 mM) e fosfato de potássio monobásico (KH_2PO_4) (0, 1, 3 e 9 mM). Um dos grupos (NPK) recebeu somente água. Para melhor distribuição da solução ao redor da planta, esta foi aplicada diluída em 500 mL de água destilada.

Tabela 1. Grupos de plantas em tratamento com diferentes composições de solução nutritiva.

Tratamento*	NH_4NO_3	KH_2PO_4
N2P1	2 mM	1 mM
N20P1	20 mM	1 mM
N40P1	40 mM	1 mM
N80P1	80 mM	1 mM
N20PØ	20 mM	0
N20P3	20 mM	3 mM
N20P9	20 mM	9 mM
N80P9	80 mM	9 mM
NØPØ**	0	0
NPK***	0	0

*N=concentração de nitrato de amônio, em mM; P=concentração de fosfato de potássio monobásico, em mM. **O tratamento NØPØ recebeu todos os nutrientes da solução de Hoagland, exceto nitrato de amônio e fosfato de potássio monobásico. *** O tratamento NPK recebeu somente NPK assim que as plantas passaram para os vasos, assim como ocorreu com os demais tratamentos.

As plantas foram mantidas em condições ambiente no Laboratório de Fisiologia Ecológica de Plantas (Lafieco, IB/USP) desde a transferência de bandejas para vasos, e durante o tratamento. Ao fim deste, as plantas foram agrupadas em 3 réplicas biológicas por tratamento, em *pools* de três plantas cada. As raízes foram coletadas conforme descrito anteriormente. Oito raízes por *pool* foram fixadas em etanol 70% para verificação do desenvolvimento diferencial do aerênquima entre os dois tratamentos. Três foram inicialmente cortadas à mão e analisadas por microscopia de luz. Após selecionados os dois tratamentos para os experimentos posteriores, outras cinco raízes de cada *pool* foram utilizadas no processo de parafinização para cálculo da área ocupada pelo aerênquima, conforme descrito no item 4.

2. Cultivo de cana-de-açúcar em 1-metilciclopropeno

Este experimento foi realizado com plantas cultivadas em câmaras de topo aberto (Fig. 6). Destas, metade foi cultivada em condições controle e a outra metade tratada com 1-MCP (AgroFresh®). Em dias alternados, o 1-MCP foi administrado às plantas na concentração final de 1 ppm. As câmaras foram mantidas com as portas fechadas e seladas no topo por 24 h após aplicação do inibidor na base do vaso, seguido de 5 mL de água destilada. O mesmo volume de água foi adicionado na base do vaso de plantas controle. Cada vaso foi mantido num envoltório plástico visando manter o 1-MCP junto à porção radicular da planta, visto que este reagente é liberado em sua forma gasosa uma vez que entra em contato com a água. Decorrido o tempo de tratamento, o sistema de exaustão das câmaras foi ligado e o 1-MCP remanescente liberado, diminuindo a contaminação da parte aérea. Ao final de 24 h da abertura das câmaras e das embalagens plásticas em torno dos vasos, o reagente foi administrado novamente e o procedimento se repetiu por 3 meses.

Figura 6. Grupos de plantas com vasos embalados cultivadas dentro de câmaras de topo aberto.

Após o fim do tratamento foram coletadas raízes de 3 réplicas biológicas de cada tratamento, sendo que cada réplica foi composta por um *pool* de 6 plantas. O material foi imediatamente congelado em nitrogênio líquido e mantido a -80°C . Para cada réplica também foram coletados (i) segmentos de cinco raízes para fixação em etanol 70% e (ii) segmentos de 8 raízes que foram pesados e armazenados em frascos âmbar herméticos para a dosagem de hormônios (He *et al.*, 1996a). Estes últimos permaneceram 24 h em temperatura

ambiente no escuro e, em seguida, uma alíquota de ar foi recolhida para dosagem dos níveis de etileno. O material fixado destinou-se à análise do padrão de formação do aerênquima por microscopia de luz.

3. Preparação das amostras para microscopia de luz e cálculo da área do aerênquima

Cortes à mão dos cinco segmentos de raízes das réplicas biológicas dos dez tratamentos oriundos com experimento de estresse nutricional foram avaliados visualmente quanto à tendência de formação do aerênquima, com coloração de Safrablau 1%. Outras cinco raízes dos dois tratamentos selecionados para experimentos posteriores foram submetidos ao procedimento de inclusão em parafina.

O material fixado em etanol 70% foi submetido à desidratação gradativa, passando por etanol 60% e em seguida por etanol 50%, ambos por no mínimo 3h em bomba de vácuo. Em seguida deu-se início à série butanólica, que consistiu em manutenção do material em soluções de etanol com concentrações crescentes de butanol (50, 70, 85, 95 e 100%), por 2h cada. Ao final desta série, as raízes foram mantidas em butanol 100% a 40°C durante a noite. Para inclusão, as raízes foram mantidas a 60°C em concentrações crescentes de parafina diluída em butanol (30%, 60% e 100%) por 3h cada, sendo a última (parafina 100%) trocada duas vezes e mantida a 60°C durante uma noite. No dia seguinte, o material foi emblocado em parafina.

Após endurecimento da parafina o material foi cortado em micrótomo (Leica RM2125) em secções de 10-16 µm, afixados em lâminas de vidro com adesivo de Bissing (Bissing, 1974) e deixados a 60°C por 24h para secagem. As lâminas contendo os cortes foram desparafinizadas a 60°C em acetato de butila por dois períodos de 1h cada, dando início à rehidratação com imersão em soluções de concentração decrescentes de etanol (100, 90, 80, 70, 60 e 50%) por 1 min cada. As lâminas foram mantidas imersas em solução azul de astra 1% por 1h (Kraus & Arduin, 1997). Em seguida, após rápida lavagem em etanol 50% para remover o excesso de corante, as lâminas foram coradas com safranina 1% (Kraus & Arduin, 1997) por 40 min. Após nova desidratação do material em concentrações crescentes de etanol (50-100%) por 1 min cada, as lâminas foram rapidamente lavadas em acetato de butila para montagem com bálsamo do Canadá.

Os cortes à mão e por microtomia foram analisados e fotografados em microscopia de luz Olympus modelo Bx51 com condensador Aplanat Achromat 1.4 e fotografados por câmera digital Olympus, marca Evolt E330. O cálculo da área referente aos espaços abertos no córtex pelo aerênquima foi realizado por meio do programa ImagePro (MediaCybernetics, v.6.3), procedimento que só se deu para amostras incluídas em parafina. Os valores obtidos em porcentagem da área do aerênquima em relação a área do córtex foram definidos em termos comparativos em relação à área do segmento 5 nos tratamentos NPK (do experimento de estresse nutricional) e controle (no experimento com 1-MCP). Neste segmento a área foi definida como equivalente a 1 e nos demais segmentos foi calculada a área proporcional.

4. Dosagem de AIA livre

A extração de AIA livre foi realizada de acordo com o protocolo elaborado por Ludwig-Muller e colaboradores (2008), com modificações. A 50 mg de material macerado em N₂ líquido e homogenizado foram adicionados 0,5 mL de solução isopropanol:ácido acético (95:5), ao qual foi posteriormente adicionado 0,5 µg de [¹³C₆]-AIA, o padrão de auxina. Os tubos foram agitados em vórtex por 1 minuto e mantidos a 4°C por 2h. Em seguida, o material foi centrifugado a 14.000 x g por 10 minutos a 4°C e o sobrenadante foi coletado e seco em fluxo de N₂ até atingir volume entre 20 e 50 µL. O pH foi ajustado para 2,5-3,5 com HCl 1N. Os tubos foram então agitados em vórtex por 1 minuto após adição de 500 µL de acetato de etila e centrifugados a 14.000 x g por 5 minutos a 4°C. Para extração exaustiva, o sedimento produzido nesta etapa foi utilizado para mais duas repetições da centrifugação, com o mesmo volume de acetato de etila. Os sobrenadantes de cada extração foram coletados e colocados em um novo tubo e secos em fluxo de N₂.

As amostras foram ressuspensas em 100 µL de metanol e transferidas para *vials* com insertos, ao quais foram adicionados 50 µL de trimetilmetilsilil-diazometano para derivatização da auxina. Os frascos foram imediatamente fechados e mantidos à temperatura ambiente por 30 minutos. Em seguida, as amostras foram completamente secas e ressuspensas em 20 µL de acetato de etila e aplicadas no GC-MS (Hewlett-Packard, modelo 6890) acoplado a um espectrômetro de massas (Hewlett-Packard, modelo 5973). A coluna utilizada para as separações foi a HP-1701 (30 m, D.I. 0,25 mm, 0,50 m de espessura do filme interno) tendo hélio como gás de arraste, com fluxo de 4 mL/min. A coluna foi mantida a 150°C por 3 minutos, seguida de rampa de temperatura de 5°C por minuto até 210°C e 15°C por minuto até 260°C. Foram monitorados os íons com relação massa/carga (*m/z*) entre 130 e 189 (correspondentes ao AIA endógeno) e 136 e 195 (correspondentes ao padrão interno de [¹³C₆]-AIA). As concentrações endógenas de AIA foram obtidas pela relação entre as áreas dos picos nos cromatogramas extraídos em *m/z* 130 e 136.

5. Dosagem de etileno

De posse dos frascos herméticos contendo segmentos de 8 raízes incubados por 24 h à temperatura ambiente e abrigados da luz, 1 mL de ar foi coletado com uma seringa e aplicado no GC-MS com detecção por ionização de chama. Para a injeção de 1 mL de ar, foi empregado o modo de injeção *pulsed splitless*, desenvolvido pela Hewlett-Packard. As condições de injeção foram: pressão de 20 psi por 2 minutos, fluxo de ventilação de 5 mL por minuto após 30 segundos de injeção e temperatura do injetor em 200°C. As demais condições cromatográficas empregadas foram: corrida isotérmica a 30°C empregando hélio como gás carregador em fluxo constante de 1 mL por minuto e temperatura do detector em 250°C. O fluxo de ar e hidrogênio no detector deu-se a 450 mL por minuto e 50 mL por minuto, respectivamente. A estimativa da quantidade de etileno produzida pelos segmentos de raiz foi feita em relação à injeção de um padrão de 0,1 nL/L de etileno em ar sintético (Air Liquid).

6. Escolha de sequências-alvo e desenho de oligonucleotídeos específicos

As sequências de glicosil hidrolases selecionadas para análise foram obtidas no banco de dados de ESTs de cana-de-açúcar (Vettore *et al.*, 2001; <http://sucest-fun.org/>). Exceto para os SAS SCACSB1036D01.g, SCJLFL1054H05.g e SCEPFL3083G08.g, foram avaliadas somente as sequências contidas no banco *Root Tip* (RT) cujo agrupamento de *clusters* foi previamente analisado por Lima e colaboradores (2001), apontando-as como referentes a enzimas modificadoras de parede celular. As sequências foram analisadas por meio de alinhamento local empregando sequência nucleotídica como consulta e bancos de proteínas como sujeito (<http://blast.ncbi.nlm.nih.gov/>; Altschul *et al.*, 1990). Dentre os melhores *hits* buscamos por relatos na literatura a respeito da confirmação da atividade glicolítica das proteínas homólogas obtidas no BLASTx.

Quanto à busca de sequências relacionados a proteínas ligantes a cálcio e fatores de transcrição e outras glicosil hidrolases foi realizada uma pesquisa por alinhamento local empregando como consulta genes apontados como diferencialmente expressos durante a formação de aerênquima em *Zea mays* (Rajhi *et al.*, 2011) e como sujeito o banco de sequências de cana-de-açúcar. Neste procedimento foram selecionados os três SAS supracitados, que não estão contidos no banco RT. Adicionalmente, realizamos o mesmo procedimento para um gene codificador da cisteína protease envolvido no processo de morte celular programada de *O. sativa* (Gothandam, 2011).

Além destes, oito normalizadores, gentilmente cedidos pela Dra. Glaucia Mendes Souza, foram testados (Tabela 2), foram testados quanto à expressão homogênea entre as amostras. A eficiência dos iniciadores foram calculadas a partir de curvas padrão empregando 4 diluições do *pool* de cDNAs e apenas aqueles com eficiência $\{[10^{(-1/\text{slope})} - 1] \times 100\}$ (Pfaffl, 2001) na faixa de 90-110% foram utilizados.

Tabela 2. *Sugarcane Assembled Sequences* (SAS) de cana-de-açúcar usados para desenho de oligonucleotídeos normalizadores

SAS	Sequência dos iniciadores ¹	Nome
SCCST2001G02.g	CCGGTCCTTTAAACCAACTCAGT; CCCTCTGGTGTACCTCCATTTG	Poliubiquitina
SCCCLR1069D05.g	CCAGTTCCATTGTCACAAACAAG TCTCCGGAATCCGTAGCAAA	Actina
SCCCLR1072A03.g	CCCTGTGCTGCTCACTGAAG GTCTCGAACATAATCTGGGTCATCT	Subunidade delta da ATP sintase mitocondrial
SCCCRZ1002H03.g	CCTCAGAGCCAATCCTTTTCC ATAACCGCATCAGGTCTCCAAG	Tubulina
SCBGLR1002D06.g	CAGGTCCTGCTGGTGAGGAT CACCTCCAGCATATGGACTATCAG	Enzima E2 de ubiquitinação
SCCCLR1048F12.g	CGTCCGCAGTCCCAAT TGAGAGGATCGCGAGGATTC	Ubiquitina
SCJFRZ2009G01.g	GCGAGTGCCTCACCTTTGAC; TCTTAGGTCCCCTCAGCAGAAC	Subunidade 60S do ribossomo

¹ Na direção 5' → 3', iniciador senso e antisenso, respectivamente

Todos iniciadores foram desenhados com a ferramenta Primer 3 (<http://frodo.wi.mit.edu/>), de acordo com os seguintes parâmetros: $58^{\circ}\text{C} \leq T_m$ (iniciador) $\leq 60^{\circ}\text{C}$; $30\% \leq \text{GC} \leq 80\%$; $18 \text{ pb} \leq \text{tamanho do iniciador} \leq 25 \text{ pb}$; $50 \text{ pb} \leq \text{tamanho do amplicon} \leq 150 \text{ pb}$; $0^{\circ}\text{C} \leq T_m \text{ do amplicon} \leq 85^{\circ}\text{C}$. Os iniciadores (Life Technologies®) foram recebidos na forma de um liofilizado, ressuspensos em Tris 10 mM EDTA 1 mM pH 7 para concentração final 100 μM e armazenados a -20°C até o momento do uso.

As sequências selecionadas por estas metodologias e usadas para desenho de iniciadores empregados no experimento de estresse nutricional descritos na Tabela 3. Já para as amostras do experimento com aplicação de 1-MCP os SAS foram selecionados para desenho de oligonucleotídeos (Tabela 4) de acordo com o padrão de expressão obtido por microarrajno (descrito no item 9). A nomenclatura dos genes empregados para desenho dos iniciadores é a mesma da anotação da proteína obtida como melhor *hit* em alinhamento por BLASTx empregando a sequência do SAS como consulta. Dentre os iniciadores listados estão aqueles usados também na triagem da biblioteca da variedade R570 (descrita no item 1 da Parte C da metodologia).

Tabela 3. *Sugarcane Assembled Sequences* (SAS) de cana-de-açúcar usados para desenho de oligonucleotídeos específicos

SAS	Sequência do iniciador	Melhor <i>hit</i>	Identidade e e-value
SCCCLB1004G05.g	tgcagcaggagatcgtgtc; tgagggagaactcggggatg	calcineurin subunit B; <i>Z. mays</i>	95% 4e^{-98}
SCCCLR2003G10.g	cctcccagccaagcatcac; gcggtcgcggtagattagatgatg	probable calcium-binding protein; <i>S. italica</i>	70%; 1e^{-25}
SCCCLB1023F09.g	tccgccccacttagattcc; ccccttctgctgtctctctc	cathepsin B-like cysteine proteinase 3; <i>Z. mays</i>	95% 0.0
SCCCL4002B07.g	ggctgtgcaacccttccaaa; gtggtgggatggacgaggatt	ethylene-responsive factor-like protein ; <i>Z. mays</i>	85% 2e^{-77}
SCAGLR2011B02.g	tggcgctgtccattccagt; catctccgcaacagaagcatac	DNA binding protein RAV AA58010.1; <i>Z. mays</i>	80% 9e^{-163}
SCRLLR1059G12.g	gttgtgacgggcggttcttg; gcggtgttgagatggaggga	MYB DNA-binding domain superfamily; <i>Z. mays</i>	81% 2e^{-154}
SCCCLR2003C07.g	gacagcttcagacgcacga; gtgccaatgccgttgttct	NAC domain-containing protein 68-like; <i>S. italica</i>	87%; 3e^{-139}
SCQSRT1034D03.g	ggaacgctgtacccttgga; ggccaagacgcgaactgtct	α -L-arabinofuranosidase; <i>Z. mays</i>	95% 0.0
SCEQRT1029E06.g	cctgtagcgttggggagag; agtaccatggaggccaag	lysosomal β -glucosidase-like isoform; <i>S. italica</i>	72% 7e^{-68}
SCCART1002G03.g	gccacggcatcacaatcaga; acgtaccagccgccttcaa	probable β -D-xylosidase 6-like; <i>O. brachyantha</i>	81% 0.0
SCVPRT2075H10.g	ccggggaggcagctacttag; gatctggggaagcacggcta	α -expansin 13 precursor; <i>Z. mays</i>	100% 1e^{-62}
SCAGRT2038D08.g	gcaccaagctcctccctgaa; gccgggttaccactccaga	putative pectinacetylesterase precursor; <i>O. sativa</i>	72% 2e^{-138}
SCEPFL3083G08.g	gttgccggtcatccagaaca; gcggcttgtcaggacatt	probable polygalacturonase-like; <i>S. italica</i>	91% 5e^{-174}
SCJLFL1054H05.g	tgctcatgcgaagtcgttc; gcgtgcgcctgtactcgt	endoglucanase 1 precursor; <i>Z. mays</i>	92% 1e^{-127}
SCACSB1036D01.g	tgggcaacaaggagatgcag; acgccgtcgatgtaccacac	xyloglucan endotransglucosylase/hydrolase <i>Z. mays</i>	90% 4e^{-87}

7. Extração de RNA de raízes de cana-de-açúcar e síntese de cDNA para qPCR e microarranjo

Os segmentos de raízes dos experimentos de estresse nutricional e 1-MCP foram macerados em N₂ líquido usando almofariz e pistilo. Cerca de 200 mg do pó obtido foi transferido para um tubo de 2 mL e homogenizados em 1,5 mL de Trizol (Life Technologies®) para a extração do RNA total, seguindo as instruções do fabricante. Para a PCR em tempo real o RNA total sedimentado foi ressuspensionado em 30 µL de água tratada com dietil pirocarbonato. As amostras foram quantificadas por espectrofotometria (NanoDrop ND-1000 UV–VIS Spectrophotometer Thermo Scientific®) e o controle de integridade analisado em gel de agarose 0,8% corado com brometo de etídeo. Cinco microgramas de RNA total foram tratados com *DNase I Amplification Grade* (Life Technologies®).

A verificação da contaminação com DNA genômico se deu por meio de amplificação com o iniciador GAPDH (gliceraldeído 3-fosfato desidrogenase). Neste caso, os iniciadores GAPDH-FWD (5' TTTGAATGGCAAGCTCACTG 3') e GAPDH-REV (5' GGTGGAAACCAAATCCTCCT 3') (Iskandar *et al.*, 2004) foram usados para teste de amplificação a partir de 1 µL de RNA total tratado com DNase, sendo água e 150 ng de DNA genômico (das linhagens SP80-3280 e R570) usados como controles negativo e positivo, respectivamente. O sistema de amplificação consistiu em 1U de GoTaq DNA Polimerase (Promega®), dNTPs (Invitrogen®) 0,2 mM, iniciadores (Invitrogen®) 0,2 µM e tampão da enzima 1x, em 25 µL de reação. O programa de amplificação iniciou-se com 5 minutos de incubação a 95°C, seguido de 30 ciclos de incubação a 95°C por 2 minutos, anelamento dos iniciadores a 62°C por 1 minuto e extensão a 72°C por 2 minutos. Em seguida deu-se a extensão final a 72°C por 5 minutos. O resultado foi analisado por eletroforese em gel de agarose corado com brometo de etídeo.

Confirmada a ausência de DNA contaminante, 2 µg de RNA foram submetidos a síntese de cDNA utilizando o sistema de síntese *Super Script™ First Strand* por RT-PCR (Life Technologies®) de acordo com as instruções do fabricante.

Para o microarranjo, 100 mg de raízes maceradas foram transferidos para um tubo de 2 mL e homogenizados em 0,5 mL de *Concert Plant RNA Reagent* (Life Technologies®) para a extração do RNA total, seguindo as instruções do fabricante. O RNA total sedimentado foi ressuspensionado em 10 µL de água tratada com DEPC. As amostras foram quantificadas por espectrofotometria (NanoDrop ND-1000 UV–VIS Spectrophotometer Thermo Scientific®) e o controle de integridade analisado em gel de agarose 0,8 % corado com brometo de etídeo. Cinco microgramas de RNA total foram tratados com *DNase I Amplification Grade* (Life Technologies®) e, em seguida, purificado em coluna utilizando o protocolo RNA *clean-up* do *RNeasy Mini kit* (Qiagen®), seguindo instruções do fabricante. O RNA eluído em 30 µL de água livre de nucleases foi analisado quanto à sua integridade em gel de agarose 0,8% corado com brometo de etídeo e quantificado por espectrofotometria. Para o microarranjo duas réplicas biológicas foram empregadas. A terceira réplica biológica foi usada nos experimentos de validação por qPCR.

8. Análise dos padrões de expressão gênica por PCR em tempo real

A amplificação por PCR foi realizada e monitorada no equipamento 7300 Real Time PCR System (Life Technologies®) usando o reagente SYBR Green PCR Master Mix (Life Technologies®) de acordo com o protocolo do fabricante. O programa de amplificação consistiu em desnaturação a 95°C durante 5 minutos, seguido por 40 ciclos de desnaturação a 95°C, hibridação do iniciador a 60°C e extensão a 72°C durante um minuto cada etapa. Os valores de CNRQ (*calibrated normalized relative quantity*) foram determinados pelo *software 2.0 qBasePLUS* e os iniciadores de referência foram selecionados usando a ferramenta *GeNormPLUS* (Biogazelle; Hellemans *et al.*, 2007). No caso das amostras do experimento de estresse ambiental os normalizadores selecionados foram poliubiquitina, 60S e ubiquitina. Já no experimento com aplicação com 1-MCP os normalizadores apontados pelo programa foram poliubiquitina, tubulina e enzima E2 de ubiquitinação (Tabela 2).

Os resultados são mostrados em expressão relativa ao valor de CNRQ obtido para o segmento 1 do tratamento N80P9 no experimento de estresse nutricional. No caso das amostras do experimento com 1-MCP os resultados são relativos ao valor de CNRQ obtido nas amostras controle de cada segmento. As reações de qPCR foram feitas em triplicatas e curvas de dissociação foram avaliadas quanto a presença de amplificações inespecíficas.

9. Hibridização em *oligoarray* e processamento da imagem

A preparação e hibridização das amostras foi executada de acordo com o protocolo *Two-Color Microarray-Based Gene Expression Analysis* (Quick Amp Labeling®) e os controles de *spike* foram preparados utilizando o *RNA Spike-In Kit, Two-Color* (Agilent Technologies®), seguindo instruções do fabricante. Os cRNAs amplificados marcados com Cy-5 e -3 foram obtidos a partir de 2 µg de RNA das amostras controle e tratadas usando o *Agilent's Quick Amp Labeling Kit* (Agilent Technologies®), que contém T7 RNA polimerase, que amplifica simultaneamente o material alvo e incorpora a citidina marcada com Cy-5 e Cy-3. O cRNA resultante foi purificado em coluna usando as *RNeasy mini spin columns* do *RNeasy Mini Kit* (Qiagen®) e quantificados por espectrofotometria (NanoDrop ND-1000 UV-VIS Spectrophotometer Thermo Scientific®). A hibridização foi executada de acordo com o *Gene Expression Hybridization Kit protocol* (Agilent Technologies®) e após 17h as lâminas foram lavadas nos tampões *Gene Expression Wash Buffer 1 e 2* (Agilent Technologies®) com Triton X-100 0,005% de acordo com o protocolo do fabricante. Em seguida as lâminas foram escaneadas (GenePix® 4000B scanner, Molecular Devices, Sunnyvale, CA, USA). Duas réplicas biológicas com troca de corantes foram usadas em cada hibridização. A normalização, processamento dos dados e análise estatística foram executados conforme descrito em Lembke e colaboradores (2012), sendo os resultados expressos como a razão dos valores de expressão das amostras tratadas com 1-MCP em relação ao mesmo valor das amostras controle (em log₂). A validação por qPCR se deu conforme metodologia descrita no item 8 e os iniciadores específicos encontram-se listados na Tabela 4.

Tabela 4. *Sugarcane Assembled Sequences* (SAS) selecionados para validação do microarranjo e respectivos iniciadores

SAS	Sequência do iniciador*	Anotação e número de acesso	E-value e identidade
SCVPLB1020G05.g	tggcaacactacggcactcc; gctgcgagcgaaccagagtaa	putative α -L-arabinofuranosidase; <i>Z. mays</i>	8e ⁻¹⁴⁷ 98%
SCEQHR1082B01.g	cctccagcggaatcgtcttg; gagttcgtgccatcttctcg	β -glucosidase isozyme 2 precursor; <i>O. sativa</i>	0.0 81%
SCVPRZ2042G03.g	acgccgagttctgctccag; gaaggtaaaccagttcttcaccctgtc	non-cyanogenic β -glucosidase; <i>Z. mays</i>	0.0 96%
SCCCCL7002A04.g	tggcgtttctcgggagtgt; tggcttggcacctacagcac	EREB-like protein; <i>O. sativa</i>	2e ⁻³⁴ 46%
SCVPRT2076E10.g	gccaaaggccaagctcaacttc; cgtcgtcgtcttgagtagtg	putative AP2/EREBP transcription factor superfamily protein; <i>Z. mays</i>	2e ⁻⁴⁸ 73%
SCJLLR1054A06.g	cgaaacgcaacgcaacgcata; gacaaggacgccaggttcca	1-aminocyclopropane-1- carboxylate oxidase; <i>Z. mays</i>	6e ⁻¹⁸⁰ 86%
SCAGLR2011D02.g	tcgaggacttctgcacaa; gtggaactcatgggtgtgtga	respiratory burst oxidase, putative, expressed; <i>O. sativa</i>	0.0 87%
SCEQRT1024C07.g	cgtcatcgccgagaaccagt; acgtgctgtcccgaagg	proton translocating pyrophosphatase; <i>O. sativa</i>	5e ⁻¹⁴⁵ 95%
SCSFSD1065B09.g	gcatgggacaacgccaaga; tccgatggtgtctccgatca	pyrophosphate-energized vacuolar membrane proton pump; <i>Z. mays</i>	5e ⁻⁸⁵ 98%
SCCCLR2001E08.g	tcccaccagtgtccagtgc; cctgctcctcgtgaagacg	non-symbiotic hemoglobin; <i>Z. mays</i>	1e ⁻¹⁰⁵ 93%
SCEPRT2043B01.g	atgggtggcgaggagatca; gtgtgtcgtgtgagtagaggctaca	β -expansin 3; <i>Z. mays</i>	2e ⁻³⁷ 69%
SCEQRT2026A06.g	ctgctgccgtcgttctgc; cgctatcatcgtgccctcgt	β -expansin EXPB4, putative; <i>R. communis</i>	5e ⁻⁴⁰ 68%
SCEQRT1025C04.g	cgccccacgagttctcatc; acctcaaccacgccatcacc	cysteine protease 1; <i>Z. mays</i>	5e ⁻⁶⁹ 91%

* Na direção 5' → 3', iniciador senso e antisenso, respectivamente

PARTE C: Seleção e sequenciamento de genes alvo

1. Triagem das sequências alvo na biblioteca de BACs de *Saccharum hybrid* (SHCRBa) da variedade R570

Nas triagens dos genes *ScERF1*, *ScEPG1*, *ScARA1* e *ScRAV1* (referentes aos SAS SCVPRT2076E10.g, SCEPFL3083G08.g, SCVPLB1020G05.g e SCAGLR2011B02.g, respectivamente) foi usado como *template* o genoma da linhagem R570 de cana-de-açúcar fragmentado e clonado em cromossomos artificiais de bactéria (BACs) por Tomkins e colaboradores (1999), consistindo em 103.296 clones armazenados em 269 placas. Para busca das sequências-alvo foram usados como sondas os oligonucleotídeos específicos desenhados a partir dos SAS mencionados (Tabelas 3 e 4) no *pool* 3D elaborado no CNR GV-INRA (Toulouse, França) em colaboração com o grupo da Dra. Marie-Anne van Sluys (GaTE, IB/USP).

O *pool* 3D consiste em *pools* de DNA plasmidial armazenados em tubos de 1,5 mL ou placas de 96 poços produzidos a partir de uma mistura definida de vários dos clones que compõem a biblioteca. Os *pools* armazenados em tubos constituem os 11 *superpools*, que contêm a mistura do DNA plasmidial

originalmente carregado por pouco mais de 9.000 clones. Cada *superpool* corresponde a um bloco que, por sua vez, consiste em uma placa de 96 poços contendo em cada poço misturas de DNA plasmidial de um conjunto menor de clones. Para cada um dos blocos foram realizadas misturas por séries de placas, linhas e colunas, conforme ilustrado na Figura 7. Após a busca nos blocos, as coordenadas dos clones positivos foram obtidas pela associação de uma linha, uma coluna e uma placa.

A	B	1	2	17	18	P9	P10	C+
C	D	3	4	19	20	P11	P12	C-
E	F	5	6	21	22	P13	P14	SP1
G	H	7	8	23	24	P15	P16	
I	J	9	10	P1	P2	P17	P18	
K	L	11	12	P3	P4	P19	P20	
M	N	13	14	P5	P6	P21	P22	
O	P	15	16	P7	P8	P23	P24	

Figura 7. Esquema de uma corrida de qPCR montada a partir do DNA contido nos blocos (neste caso o exemplo é o bloco 1) empregados na triagem. Cada um dos poços contém como molde o DNA correspondente a linhas (azul), colunas (laranja) ou placas (verde), originando a coordenada que leva à localização do clone a ser testado. Os *superpools* (SP) também são adicionados (rosa). Controle positivo (C+): DNA genômico da linhagem R570; controle negativo (C-): água ultra-pura.

Usando a Figura 7 como exemplo, caso houvesse *amplicons* com perfil de dissociação coincidente com o dos controles positivos (C+ e SP1) nos poços D, G, 6, 23, 18, P10 e P21, isso significaria que os clones carregando a sequência-alvo estão localizados nas placas 10 e 21, em pelo menos uma das posições seguintes: D6, D18, D23, G6, G18 e G23, indicando que todas devem ser testadas na qPCR de colônia.

A última etapa consistiu em uma qPCR de colônia. Para a preparação das culturas, os clones selecionados foram semeados em placas de Petri contendo LB ágar (peptona 1%, extrato de levedura 0,5%, NaCl 1%, ágar 1,5%, pH 7,2) suplementado com cloranfenicol na concentração final 50 µg/mL e mantidos a 37°C por 14-18 h. Em seguida, uma colônia isolada foi inoculada em 1,2 mL de meio LB líquido e cloranfenicol na concentração final 50 µg/mL a 37°C por 14 a 18h e 200 rpm. Decorrido o tempo de crescimento, 1 µL da cultura foi empregada como molde na qPCR de colônia. O resultado foi posteriormente validado por qPCR utilizando como molde a cultura crescida a partir de três colônias isoladas do clone identificado. Os clones positivos foram selecionados para extração de DNA plasmidial.

A montagem do sistema de reação empregou os seguintes reagentes nas dadas concentrações finais: SyBr Green PCR Master Mix (Life Technologies®) 1x, oligonucleotídeo senso e reverso 0,4 µM cada e 150 ng do DNA molde, em um volume final de 15 µL. As condições de amplificação seguiram o seguinte programa: 95°C por 5 minutos, 40 ciclos 95°C por 20 segundos, 60°C por 20 segundos e 72°C por 40 segundos, seguida da curva de dissociação (1 ciclo de 95°C por 15 segundos, 60°C por 30 segundos e 95°C por 15 segundos). Em todas as corridas de qPCR o critério para definição da presença do DNA plasmidial do clone positivo baseou-se na produção de um *amplicon* único no tamanho esperado num dado poço. Tal análise se deu a partir do perfil da curva de dissociação do poço em questão em

relação à curva obtida nos controles positivos (DNA genômico da variedade R570), pipetados na mesma placa.

2. Extração de DNA plasmidial em média escala

Uma colônia isolada obtida das placas confeccionadas para a qPCR de colônia foi inoculada em 3 mL de meio LB contendo cloranfenicol na concentração final de 50 µg/mL por 7 h a 37°C e 200 rpm, consistindo no pré-inóculo. Em seguida a 900 mL de meio LB com cloranfenicol na concentração final 50 µg/mL foram adicionados 900 µL do pré-inóculo crescido. As culturas foram mantidas a 37°C e 200 rpm por 14 h.

Ao final deste período, a extração se deu empregando o protocolo *Genopure Plasmid Midi kit* (Roche®), seguindo instruções do fabricante. Ao DNA plasmidial sedimentado ao final do protocolo foram adicionados 300 µL de água bidestilada e o material foi mantido sob agitação (200 rpm) por cerca de 14 h para completa ressuspensão. O DNA extraído foi quantificado por espectrofotometria (NanoDrop ND-1000 UV-VIS Spectrophotometer Thermo Scientific®) e analisado quanto à sua integridade por eletroforese em gel de agarose 0,8 %.

3. Preparo da biblioteca, sequenciamento na plataforma 454/Roche e predição de genes

As bibliotecas foram preparadas a partir de 100 ng de cada BAC extraído usando o *Rapid Library Preparation Method Manual* e os procedimentos para execução da PCR em emulsão se deram de acordo com o *emPCR Method Manual – Lib-L SV* (GS FLX Titanium Series, Roche®). O sequenciamento foi realizado em plataforma 454/Roche®. Após sequenciamento, a montagem das sequências foi realizada pelo Dr. João Paulo Kitajima (Hospital Albert Einstein), usando *Phrap* (www.phrap.org) conforme metodologia descrita por De Setta e colaboradores (2014). As sequências montadas dos BACs foram carregadas no *pipeline* para predição e anotação de genes em BACs de cana-de-açúcar desenvolvida pelo Dr. Alessandro de Mello Varani do Departamento de Tecnologia da Faculdade de Ciências Agrárias e Veterinárias – UNESP/Jaboticabal/SP.

4. Avaliação da predição automática e caracterização gênômica das sequências candidatas

A predição de todas dos genes alvo *ScARA1*, *ScEPG1*, *ScERF1* e *ScRAV1* foi submetida à avaliação comparativa empregando prováveis genes ortólogos em *Oryza sativa*, *Sorghum bicolor* e *Zea mays*. A busca pelo provável gene ortólogo no genoma de *Oryza sativa* v.7.0 (International Rice Genome Sequencing Project), *Sorghum bicolor* v.2.1 (Paterson *et al.*, 2009) e *Zea mays* v. 2.0 (Schnable *et al.*, 2009) se deu por meio de alinhamento local empregando como consulta as sequências do gene alvo de cada BAC usando a ferramenta *Phytozome* v.9.1 (<http://www.phytozome.net/>; Goodstein *et al.*, 2012) nos supracitados genomas não mascarados, usados como sujeitos. Para busca com a sequência do gene alvo foram considerados resultados com valores de *e-value* inferiores a $1e^{-50}$.

Após definição dos prováveis genes ortólogos, os genes alvo hom(e)ólogos foram analisados quanto à posição predita dos códons de iniciação e parada e número e posição de íntrons em relação à sequência dos ortólogos nas plantas supracitadas. Esta análise se deu por meio de alinhamento múltiplo por ClustalW (Thompson *et al.*, 1994) disponível no *software BioEdit Sequence Alignment Editor v.7.2.5* (<http://www.mbio.ncsu.edu/bioedit>; Hall *et al.*, 1999). Quando necessário, os alinhamentos foram otimizados manualmente.

Para análise filogenética as sequências codificantes de cana-de-açúcar e seu provável ortólogo de *S. bicolor* foram alinhadas utilizando a ferramenta MUSCLE para alinhamento múltiplo (Edgar, 2004) seguido reconstrução filogenética utilizando o método de *neighbor-joining* com distância de Tajima-Nei e 1000 repetições de *bootstrap* usando o *software* MEGA v.5.2.2 (Tamura *et al.*, 2011).

PARTE D: Análise de sintenia, filogenia e estrutura dos genes alvo sequenciados

1. Avaliação da colinearidade e sintenia dos BACs de cana-de-açúcar em relação a *S. bicolor*

Após identificação dos genes ortólogos aos genes alvo de cana-de-açúcar, foram realizadas buscas nas regiões 100 kb à jusante e à montante dos ortólogos visando identificar genes em *S. bicolor* com a mesma anotação dos genes presentes nos BACs de cana-de-açúcar, vizinhos ao gene alvo. A região do genoma de *S. bicolor* selecionada limitou-se àquela a contemplar o máximo de prováveis genes ortólogos em relação aos BACs de cana-de-açúcar. A representação de genes preditos dentro da região delimitada do genoma de *S. bicolor* e dos BACs de cana-de-açúcar seguiu dois critérios. Inicialmente foram realizados alinhamentos locais empregando como consulta a sequência codante de *S. bicolor* ou cana-de-açúcar e bancos de proteínas como sujeito pela ferramenta BLASTx (disponível em <http://blast.ncbi.nlm.nih.gov/>; Altschul *et al.*, 1990). A manutenção das sequências consultadas na representação dos BACs e na região selecionada do genoma de *S. bicolor* se deu mediante existência *hits* com alinhamento com *e-value* inferior a 10^{-15} . Após realizado este filtro, o segundo critério consistiu em eliminar da região representada dos BACs de cana-de-açúcar genes sem anotação, sem domínios preditos e presentes em só um dos BACs.

No caso de genes cuja colinearidade encontrava-se mantida entre *S. bicolor* e cana-de-açúcar mas a anotação encontrava-se distinta, foram realizados alinhamentos locais dois a dois pela ferramenta BLAST2Seq (disponível em <http://blast.ncbi.nlm.nih.gov/>; Altschul *et al.*, 1990) com configuração BLASTn. Ambos os genes foram considerados prováveis ortólogos em alinhamentos com cobertura e identidade superiores a 90% e 75%, respectivamente. Em casos como este, a presença de domínios preditos na sequência proteica dos prováveis ortólogos também foi avaliada por meio da ferramenta Pfam v.27.0 (<http://pfam.xfam.org/>; Finn *et al.*, 2014) visando a identificação de domínios em comum. No que concerne a genes preditos em ao menos um dos BACs de cana-de-açúcar e ausência de predição de genes hom(e)ólogos colineares ou não nos demais BACs, foi empregada metodologia similar à descrita acima. A sequência codante do gene cujos hom(e)ólogos estavam ausentes nos demais BACs foi empregada como consulta e a sequência integral dos demais BACs como sujeito para execução de

alinhamentos locais dois a dois. Os mesmos parâmetros de cobertura, similaridade e presença de domínios protéicos preditos em comum descritos anteriormente foram avaliados.

O mapeamento dos TEs contidos nos *loci* de interesse foi realizado somente para as sequências de cana-de-açúcar. A seleção dos TE para o mapeamento levou em consideração possíveis elementos identificados pelo *RepeatMasker* (<http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>) usando a base de dados Viridiplantae. Foram selecionados resultados com *score* superior a 250, com extensão igual ou superior a 3 kb e com descontinuidade tolerada de no máximo 300 pb.

As sequências codantes dos genes sintênicos foram concatenadas e empregadas em alinhamento múltiplo. A clusterização foi executada pela ferramenta MUSCLE (Edgar, 2004) seguida de reconstrução filogenética utilizando o método de *neighbor-joining* com distância de Tajima-Nei e 1000 repetições de *bootstrap* usando o *software* MEGA v.5.2.2 (Tamura *et al.*, 2011).

2. Análise da estrutura dos genes alvo

A análise da estrutura dos genes visou ilustrar regiões de éxons e íntrons bem como localização de elementos de transposição e sequências que codificam para os domínios protéicos preditos. Estes foram identificados pela NCBI *Conserved Domain* (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>; Marchler-Bauer *et al.*, 2015). A busca por TEs foi realizada com a base de dados de Viridiplantae no *RepeatMasker* (<http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>) e foram considerados resultados com *scores* superiores a 250. A representação de íntrons, éxons, domínios e elementos de transposição foi obtida pela ferramenta *Exon-Intron Graphic Maker* (<http://wormweb.org/exonintron>).

PARTE E: Seleção da sequência para clonagem e análise *in silico* dentre as versões do gene alvo

1. Mapeamento dos *reads* do banco de transcritos de raiz nas sequências codantes dos genes alvo

O LAFIECO dispõe de um banco de transcritos oriundo do sequenciamento do transcriptoma por RNAseq de três réplicas biológicas dos quatro segmentos mais apicais de raiz na plataforma ABI SOLiD, dando origem a 12 bibliotecas. Os *reads* componentes deste banco foram usados para mapear as sequências codantes dos genes alvo (referências) por três metodologias, executadas por Dr^a Marie-Anne van Sluys e Dr Jonas Weissmann Gaiarsa (GaTE Lab/IB/USP). A primeira empregou alinhamento manual por *Burrows-Wheeler Aligner* (<http://bio-bwa.sourceforge.net/>) com visualização pelo *software Integrative Genomics Viewer* (<https://www.broadinstitute.org/igv/v1.4>) e visou o mapeamento nas referências dos genes *ScEPG1* e *ScRAV1* (referentes à glicosil hidrolase endopoligalacturonase e ao fator de transcrição RAV). O segundo procedimento se deu pelo *software Bowtie 2.1.0* (<http://bowtie-bio.sourceforge.net/>) com as referências oriundas dos genes *ScARA1* e *ScERF1* (correspondentes à glicosil hidrolase arabinofuranosidase e ao fator de transcrição ERF). O procedimento de linhas de comando consistiu em

construir um índice das sequências codantes, seguido do mapeamento dos *reads* do transcriptoma contra os mencionados índices. Os ajustes padrão usados no *bowtie2* foram -D 15 -R 2 -N 0 -L 22 -i S,1,1.15

- D <int> *give up extending after <int> failed extends in a row* (15)
- R <int> *for reads with repetitive seeds, try <int> sets of seeds* (2)
- N <int> *max # mismatches in seed alignment; can be 0 or 1* (0)
- L <int> *length of seed substrings; must be >3, <32* (22)
- i <func> *interval between seed substrings with regards to read length* (S,1,1.15)

Na etapa de pós-alinhamento foi feita uma contagem de *reads* mapeados contra cada gene em cada tratamento de cada experimento usando as ferramentas *cut*, *sort* e *uniq* na linha de comando Linux.

A terceira metodologia usou um parâmetro distinto. Enquanto no segundo alinhamento um determinado *read* só poderia mapear uma só sequência codante, neste terceiro procedimento uma mesma *read* mapeou várias sequências codantes, desde que obedecidos os parâmetros padrão descritos acima. Além disso, o terceiro procedimento adicionou uma etapa posterior ao mapeamento por *bowtie2* na qual o número de *hits* obtidos no sequenciamento foi normalizado de acordo com o comprimento das sequências de referência e do número de *reads* nas 12 bibliotecas. O resultado é dado em FPKM (*Fragments Per Kilobase Of Exon Per Million Fragments Mapped*), dado pela equação $10^9 \cdot C / (N \cdot L)$, na qual C é o número de *reads* mapeadas na referência, N é o número de *reads* na biblioteca e L é o comprimento da referência, em pares de bases.

2. Análise *in silico* das proteínas preditas a partir dos genes-alvo

Visando analisar domínios preditos nas sequências traduzidas *in silico*, bem como de resíduos de aminoácidos críticos para a funcionalidade de cada proteína, foram realizadas buscas por alinhamento local pela ferramenta BLASTx (<http://blast.ncbi.nlm.nih.gov/>; Altschul *et al.*, 1990) empregando como consulta a sequência codificante do hom(e)ólogo do gene alvo selecionado no mapeamento com o transcriptoma. Dentre os *hits* obtidos selecionamos aqueles referentes a proteínas funcional e estruturalmente caracterizadas.

No caso das glicosil hidrolases procurou-se a família GH na qual tais proteínas estavam alocadas (<http://www.cazy.org/>; Lombard *et al.*, 2013) para levantamento de características gerais da família como padrão de dobramento e identidade da díade catalítica. Já para os fatores de transcrição, busca por BLASTx se restringiu a bancos de dados com estruturas tridimensionais de proteínas (<http://www.rcsb.org/pdb/>). Posteriormente foi feito um alinhamento múltiplo por ClustalW empregando as proteínas preditas de cana-de-açúcar, as sequências protéicas de cada provável ortólogo em *S. bicolor* e proteínas de mesma função oriundas de plantas e/ou microrganismos. Possíveis peptídios sinais nas sequências de cana-de-açúcar foram identificados pelo programa SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>; Petersen *et al.*, 2011) e sequências de localização nuclear foram

preditas pela ferramenta http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi; Kosugi *et al.*, 2009).

Tais procedimentos só foram executados com as sequências selecionadas pelo mapeamento, às quais foi adicionado o sufixo “a” (*ScERF1a*, *ScRAV1a*, *ScARA1a* e *ScEPG1a*).

PARTE F: Avaliação da interação dos fatores de transcrição RAV e ERF com os promotores das glicosil hidrolases

1. Identificação *in silico* de elementos em *cis* nos promotores de *ScEPG1a* e *ScARA1a*

Visando à identificação de prováveis elementos em *cis* nas sequências promotoras de *ScARA1a* e *ScEPG1a* sequências de 2 kb à montante do ATG predito para ambos genes foi analisada. Tal sequência foi utilizada na ferramenta PLACE (Database of Plant Cis-acting Regulatory DNA Elements; <http://www.dna.affrc.go.jp/PLACE/signalscan.html>; Higo *et al.*, 1999; Prestridge *et al.*, 1991) para mapeamento *in silico* de sequências potencialmente reconhecidas por fatores de transcrição de interesse, sobretudo os das sub-famílias RAV e ERF.

2. Método de choque térmico para obtenção de células competentes de *Escherichia coli* e transformação

Uma colônia de *E. coli* DH5 α crescida durante a noite a 37°C em placa contendo meio LB ágar foi inoculada em 5 mL de meio LB e mantida à mesma temperatura e sob agitação a 200 rpm de 14-18 h. Em seguida a 100 mL de meio LB foram inoculados 500 μ L da pré-cultura crescida. A cultura foi novamente mantida a 37°C e 200 rpm até atingir a faixa de DO₆₀₀ de 0,2 a 0,3.

A cultura foi mantida no gelo por 10 minutos e centrifugada a 2.060 x *g* por 10 minutos a 4°C. O *pellet* foi ressuspensionado em 10 mL de CaCl₂ 100 mM e em seguida mantido no gelo por 30 minutos. Em seguida a cultura foi novamente centrifugada nas mesmas condições anteriores e o *pellet* foi ressuspensionado em 2 mL de CaCl₂ 100 mM. Foi adicionado glicerol para uma concentração final de 20% e a mistura foi incubada no gelo por 2 h. Decorrido este tempo, as alíquotas de 200 μ L foram congeladas em N₂ líquido e mantidas a -80°C até o momento do uso.

Após adição do DNA plasmidial a ser usado na transformação conforme recomendado pelo fabricante dos kits de clonagem (item 6), as células foram incubadas no gelo por 30 minutos e em seguida a 42°C por 90 segundos. Posteriormente as células foram incubadas no gelo por 2 minutos e foram adicionados 300 μ L de meio SOC (triptona 2%, extrato de levedura 0,5%, NaCl 10 mM KCl 2,5 mM, MgCl₂ 10 mM, MgSO₄ 10 mM, pH 7). O sistema de transformação foi mantido a 37°C por 1h a 200 rpm e em seguida semeado em placas contendo meio LB ágar e antibiótico.

3. *Agrobacterium tumefaciens* eletrocompetentes e transformação

Uma colônia de *A. tumefaciens* GV3101::pMP90 foi inoculada em 2 mL de meio LB e mantida a 28°C e 200 rpm durante 16-18h para pré-cultura, utilizada a seguir para inocular 50 mL de meio LB. A

cultura foi crescida durante a noite até atingir OD₆₀₀ 0,5-0,7. Em seguida, a mesma foi centrifugada por 10 minutos a 1.150 x *g* a 4°C. O sedimento foi ressuspenso em 50 mL de glicerol 10% e novamente centrifugado nas mesmas condições para em seguida ser lavado em volumes decrescentes de glicerol até 500 µL. Aliquotas de 40 µL foram congeladas em N₂ líquido e mantidas a -80 ° C até o momento da utilização.

Para transformação, foram adicionados 150 ng de plasmídeo à alíquota de células. O sistema de transformação foi eletroporado a 12,5 kV / cm, 25 mF e 200 Ω. Em seguida 1 ml de meio LB foi adicionado às células, que foram incubadas a 28°C e 200 rpm durante 3 horas. Em seguida o material foi semeado em placas contendo LB ágar e antibiótico por 2-3 dias a 28°C.

4. Síntese de oligonucleotídeos para vetores para transformação de *A. tumefaciens*

As sequências dos genes alvo *ScRAV1a* e *ScERF1a* e de 2 kb da região à montante do ATG predito dos genes *ScARA1a* e *ScEPG1a* (nomeadas EPGpm e ARApM) foram empregadas para desenho dos iniciadores para construção dos vetores para ensaios de transativação. No caso específico de EPGpm, o par de iniciadores amplificou a região de -200 a -2200. Iniciadores *forward* apresentam uma cauda CACC para clonagem em vetores Gateway. Após a síntese, os iniciadores foram recebidos na forma de um pó liofilizado e reconstituídos em Tris 10 mM EDTA pH 8 e mantidos a -20°C até o momento do uso. Todos iniciadores usados neste experimento estão listados na Tabela 5.

Tabela 5. Iniciadores usados na construção de vetores para ensaio de transativação.

Iniciador	Sequência 5'→3'
RAV	CACCTAGCAGTGGTCCATTGGAAT; TCGGTCCGTAGCATTAGTCTAC
ERF	CACCTGTGTGCGTTTATACTGTGATTT; CGAGACACCGAGCAAAGAAA
EPGpm	CACCGTATGTCCACGTACTCGTGTCCG; TGATGAAACGGCAGCGCGCAT
ARApM	CACCCTCGCTAACAACACGCCG; TACGCTTCCACAGTTCCACT

5. Amplificação das sequências alvo

Ambos fatores de transcrição (sequências genômicas) e glicosil hidrolases (sequências promotoras) foram amplificados em reações empregando a *Phusion High-Fidelity DNA Polymerase* (Thermo Scientific®).

Visando a amplificação das sequências de 1,2 kb de *ScRAV1a* presente no BAC SHCRBa_021_A10 e *ScERF1a* do BAC SHCRBa_180_L14, de mesmo tamanho, 5-10 ng do *template* foram empregados em um mix de reação. Este mix apresentou os seguintes reagentes nas respectivas concentrações finais: dNTPs 0,2 mM (Promega), MgCl₂ 2,5 mM, DMSO 3%, 0,2 µM de cada iniciador e 1U de polimerase no tampão GC (para *ScRAV1a*) e HF (para *ScERF1a*) (Thermo Scientific®) 1x. As condições de amplificação consistiram em denaturação a 98°C por 35 segundos, seguido de 30 ciclos de desnaturação a 98°C, anelamento por 30 segundos e extensão por 40 segundos a 72°C. Para *ScRAV1a* temperatura de anelamento seguiu um procedimento *touchdown* iniciando a 61,5°C e decaindo 0,3°C a cada ciclo. Já para

ScERF1a o procedimento *touchdown* iniciou-se com temperatura de anelamento 66°C, decaindo 0,2°C a cada ciclo. Ambas as reações foram finalizadas com o último passo de extensão a 72°C por 10 minutos.

Em relação à amplificação dos fragmentos correspondentes à região à montante dos genes *ScARA1* e *ScEPG1* (ARAp_m e EPG_m), foram usados 10 ng de cada BAC (SHCRBa_111_H15 e 224_P16, respectivamente) em uma mistura de reação. A mistura foi montada com os seguintes reagentes nas concentrações finais: dNTPs 0,2 mM (Promega), MgCl₂ 1,5 mM (para ARAp_m) e 2,5 mM (para EPG_m), DMSO 3% (somente para ARAp_m), iniciadores a 0,4 µM (para EPG_m) ou 0,2 µM (para ARAp_m) e 1 U de polimerase em tampão HF 1x. As condições de amplificação consistiram em um passo de desnaturação inicial por 30 segundos a 98°C, 30 ciclos de desnaturação a 98°C por 10 segundos, anelamento a 70,5°C (para EPG_m) e extensão por 1 minuto a 72°C, seguido de um passo de extensão final a 72°C por 10 minutos. Para amplificação de ARAp_m, o passo de anelamento seguiu o procedimento de *touchdown* iniciando-se a 67°C e decaindo 0,3°C a cada ciclo.

6. Construção de vetores intermediários e finais

A fim de construir os vetores repórteres (com as sequências ARA1_{pm} e EPG1_{pm}) e ativadores ou efetores (com as sequências *ScRAV1a* e *ScERF1a*), os fragmentos correspondentes às sequências de interesse foram sub-clonados no vetor pENTR™/D-TOPO (Life Technologies®; Fig. 8), utilizando células competentes de *E. coli* seguindo as instruções do fabricante. De 5 a 10 colônias foram avaliadas quanto à presença do fragmento clonado por PCR empregando os mesmos procedimentos descritos acima. Os clones positivos foram inoculados em meio LB suplementado com kanamicina (na concentração final 50 µg/mL) e crescidos a 37°C a 200 rpm por 14-18h. O plasmídeo foi extraído utilizando *GeneJET Plasmid Miniprep Kit* (Thermo Science®), seguindo instruções do fabricante.

As construções foram verificadas por sequenciamento (descrito no item 7) e 150 ng de DNA plasmidial foram usados para transformar células de *E. coli* DH5α de modo a clonar os fragmentos em vetores pFAST-G02 (Shimada *et al.*, 2010) para construções efetoras e pKGWFS7 (Karimi *et al.*, 2002) para os repórteres usando *Gateway LR Clonase II* (Life Technologies®), seguindo as instruções do fabricante. A transformação de *E. coli*, bem como a verificação da presença do inserto e a extração de DNA plasmidial seguiram os procedimentos anteriormente descritos. Após checagem do material clonado por sequenciamento, 150 ng de DNA plasmidial foram adicionados a uma alíquota de células competentes de *Agrobacterium tumefaciens* para transformação seguindo protocolo já descrito (item 2).

Figura 8. Mapa do vetor pENTR/D-TOPO (Life Technologies®) empregado nas construções intermediárias. *attL1* e 2 representam as sequências para recombinação sítio-específica do vetor de entrada para um vetor de destino Gateway®. Além disso o vetor ainda apresenta o gene de resistência à kanamicina, origem de replicação para *E. coli* (pUC ori) e terminadores da transcrição T1 e T2.

7. Sequenciamento

As reações de sequenciamento foram realizadas no sequenciador ABI 3730 (Life Technologies®) com 150 ng de DNA molde, 0,5 pM de iniciadores e 3700 *big dye terminator mix* versão 3.1 1x. O programa de sequenciamento iniciou-se com desnaturação a 95°C durante 90 segundos seguido por 35 ciclos à mesma temperatura. A precipitação e a reação de sequenciamento se deram seguindo recomendação do fabricante (BigDye® Terminator Cycle Sequencing Kit v3.1, Life Technologies®).

8. Infiltração de folhas de *Nicotiana benthamiana*

Uma colônia isolada de cada um dos quatro clones de *A. tumefaciens* transformado foi inoculada em 2 mL de meio LB suplementado com gentamicina (na concentração final 20 µg/mL) para selecionar a linhagem bacteriana e espectinomicina (na concentração final 100 µg/mL) para selecionar o plasmídeo. A pré-cultura foi incubada a 28°C a 250 rpm durante 20-24h. A DO₆₀₀ da pré-cultura foi medida para avaliar o crescimento. Para a cultura foram realizadas diluições de 1:500 e 1:1000 da pré-cultura em LB. Após atingir OD₆₀₀ 0,5-0,8, a cultura foi centrifugada a 1.500 x g durante 10 minutos à temperatura ambiente.

O sedimento foi ressuspenso em um volume de tampão de infiltração (ácido N-2-morfolino-etanossulfônico 10 mM pH 5,6, MgCl₂ 10 mM e acetosiringona 100 µM) equivalente ao número de infiltrações feitas para cada construção (infiltrações / Σ volume de cada transformação). As culturas de *A. tumefaciens* abrigando as construções efetoras e repórteres foram misturadas a fim de obter OD₆₀₀=0,4 em cada cultura. Os frascos contendo *A. tumefaciens* foram mantidos no escuro por 3 h.

Uma quantidade mínima de 14 infiltrações (cobrindo um círculo de cerca de 13 mm de diâmetro) foram feitas na superfície abaxial de três folhas completamente expandidas em cada planta de

N. benthamiana. Três plantas foram infiltradas com cada combinação de construção repórter e efetora. As plantas tinham entre 4 e 5 semanas de idade, crescidas a 22-25°C, em fotoperíodo de 16-8h, intensidade luminosa de 12 $\mu\text{mol m}^{-2} \text{s}^{-1}$ e 80% de umidade. Após a infiltração, as plantas foram mantidas por dois dias no mesmo regime de temperatura e luz, mas com umidade reduzida para 70%. Cada planta foi coletada independentemente e três experimentos usando três plantas para cada construção testada foram realizados.

9. Extração de proteínas de discos foliares de *N. benthamiana* e ensaio de β -glucuronidase

Com o auxílio de um perfurador de 7 mm, 14 discos foliares foram coletados de cada planta e imediatamente congelados em N_2 líquido em tubos contendo esferas de vidro de 2 mm de diâmetro. Os tubos foram agitados em *TissueLyser* (Qiagen®) ainda sob refrigeração. O pó resultante foi ressuspensionado em 200 μL de tampão de extração (fosfato de sódio 50 mM pH 7, EDTA 10 mM, Triton X-100 0,1%, N-lauroilsarcosina 0,1% e β -mercaptoethanol 10 mM) e incubado por 1 h em gelo. A suspensão foi centrifugada por 10 minutos a 12.000 $\times g$ a 4°C. O sobrenadante foi coletado, dosado quanto à concentração de proteínas pelo método Bradford (Bio-Rad®) usando BSA como padrão e em seguida armazenado a -80°C por até duas semanas.

O ensaio enzimático da β -glucuronidase foi realizado em placas pretas de 96 poços e fundo chato (Greiner-Bio®) usando 10 μL de extrato protéico (na concentração estoque e nas diluições 1:5 e 1:10) e 90 μL de 4-metilumbeliferil β -D-galactopiranosídeo 10 mM diluído em tampão (mesma formulação do tampão de extração). A curva padrão foi preparada com 4-metilumbeliferil em concentrações de 0 a 800 μM . A reação enzimática se deu a 37°C por 2 h 30, com medidas realizadas a cada 5 minutos por um leitor de microplacas (Tristar LB 941, Berthold Technologies®) usando o *software* Microwin 2000. Os parâmetros de reação foram os seguintes: tempo de contagem 0,3 segundos, energia da lâmpada 5000, filtros de excitação e emissão configurados a 344 e 460 nm, respectivamente.

ANÁLISE ESTATÍSTICA

Para os resultados de expressão gênica ao longo de distintos segmentos, dosagem de hormônios e cálculo de área ocupada pelo aerênquima foi empregada uma análise ANOVA Two Way com teste Tukey *a posteriori* para comparar diferentes segmentos do mesmo tratamento. Para as medidas de atividade do gene repórter nos ensaios de transativação em folhas de *N. benthamiana*, a análise se deu por teste T, usando $p < 0,05$ para diferenças estatisticamente significativas. Especificamente nos dados do experimento de estresse nutricional o teste T também foi aplicado após identificação de expressão diferencial no mesmo segmento de diferentes tratamentos. Todas análises foram realizadas no programa GraphPad Prism v.6.

PARTES A & B: Cultivo de cana-de-açúcar em estresse nutricional e na presença de 1-MCP

1. Área ocupada pelo aerênquima no córtex da raiz

O padrão de formação do aerênquima nas raízes diante de distintas concentrações de fontes de N e P foi analisado visando selecionar duas amostras contrastantes para os experimentos seguintes. Um corte representativo de cada segmento pode ser observado na Figura 9. Em nenhum dos tratamentos o aerênquima encontra-se em formação nos dois segmentos mais apicais de raiz, corroborando o que foi demonstrado em experimentos anteriores com essa mesma variedade de cana-de-açúcar (Leite, 2013; Grandis, 2015). A formação do aerênquima se iniciou no segmento 3 ou 4, a depender do tratamento. Segundo o padrão observado nas raízes analisadas, dentre os tratamentos contendo 1 mM de KH_2PO_4 , as concentrações crescentes de NH_4NO_3 (2 a 80 mM) levaram a um discreto atraso na formação do aerênquima. Por outro lado, houve uma maior tendência de formação do aerênquima nos tratamentos com escassez de fontes de nitrogênio e fósforo (NPK e NØPØ), embora exista também uma sutil propensão de maior formação de aerênquima em tratamentos com reduzidas fontes de fósforo ou na ausência deste elemento (N20PØ, N2P1, N20P1). Em somente um dos tratamentos foram observados cortes em que no segmento 4 o aerênquima encontrava-se ainda ausente ou muito pouco desenvolvido (N80P9). Por fim, em todos os tratamentos o aerênquima estava presente no 5º segmento.

Por apresentarem padrões mais contrastantes de formação do aerênquima, raízes dos tratamentos N80P9 (com solução mais rica nutrientes) e NPK (sem nutrientes) foram selecionados para cálculo da área ocupada pelo aerênquima. Nas plantas cultivadas sem adição de solução nutritiva (barras brancas, Fig. 10) a formação do aerênquima não foi visível antes do segmento 3, a partir do qual a área ocupada pelos espaços de gás proporcional a S5 aumentou de 0,12 para 0,6 no segmento 4, atingindo o valor máximo (1) no segmento 5. Similarmente, nas plantas cultivadas na presença de nutrientes (barras pretas, Fig. 10), o segmento 3 apresentou área do córtex ocupada por aerênquima equivalente a 14% da área ocupada no S5 do tratamento NPK. Entretanto, a área ocupada pelo aerênquima não aumentou na transição para o segmento 4, atingindo no segmento 5 cerca de 70% da área observada no mesmo segmento do tratamento NPK. O atraso na formação do aerênquima observado em plantas crescidas em abundância de nutrientes foi estatisticamente significativo ($P < 0,001$).

RESULTADOS

Figura 9. Secções transversais dos cinco segmentos (S1-S5) de 1 cm de raiz de cana-de-açúcar irrigada com distintas soluções nutritivas. Coloração de Safranblau 1%. Escala 100 µm. Para informações quanto à nomenclatura dos tratamentos vide Tabela 1.

Figura 10. Área da seção transversal do córtex ocupada pelo aerênquima em plantas cultivadas em estresse nutricional. NPK: plantas sem solução nutritiva; N80P9: plantas que receberam solução nutritiva. Medidas foram feitas em cinco diferentes raízes oriundas de cada réplica biológica composta cada uma por *pool* de três plantas. Letras diferentes denotam diferenças estatisticamente significativas. A área do aerênquima é mostrada em proporção ao valor obtido no segmento 5 de plantas NPK.

2. Perfil transcricional diante de distintas disponibilidades de nutrientes

Para avaliar o efeito da aceleração da formação do aerênquima diante da escassez de nutrientes no âmbito transcricional, foi analisada a expressão de genes chave neste processo conforme apontado na literatura (Bouranis *et al.*, 2007; Rajhi *et al.*, 2011). Foi avaliado o padrão transcricional referente a genes que codificam para fatores de transcrição, proteínas do metabolismo de cálcio, uma cisteíno protease e glicosil hidrolases com o intuito de abranger aspectos da formação do aerênquima desde sua sinalização até o estabelecimento dos processos de morte celular programada e degradação de parede.

Embora os efeitos da formação do aerênquima só tenham sido visualizados por microscopia de luz a partir do segmento 3 (Figs. 9 e 10), os distintos padrões de expressão gênica ao longo dos segmentos já foram notados desde o primeiro segmento. No que diz respeito aos genes relacionados à morte celular programada, foi dado enfoque ao padrão transcricional de dois genes que codificam para sensores de cálcio e para uma cisteíno protease. Os níveis de transcritos referentes a uma calmodulina (*calcium binding protein CML 35-like*; SCCCLB1004G05.g; Fig. 11A) e calcineurina (SCCCLB1004G05.g Fig. 11B) foram menores no primeiro segmento, se elevando no segmento seguinte. Estes transcritos permaneceram estáveis de S3 a S5, embora as diferenças tenham sido estatisticamente significativas somente para a calmodulina ($p=0,001$). No caso específico deste transcrito os valores obtidos em S5 são o dobro do observado para S1. Já para a cisteíno protease (SCCCLB1023F09.g; Fig. 11C) foi observado um padrão transcricional bem distinto. O acúmulo mais proeminente destes transcritos foi observado no último segmento, atingindo níveis cerca de duas vezes maior em plantas em déficit nutricional comparado a plantas tratadas com nutrientes.

Figura 11. Perfil de expressão de genes relacionados ao metabolismo de cálcio e proteólise (A-C). O eixo y está normalizado para expressão relativa ao valor de CNRQ obtido para o segmento 1 do tratamento N80P9 $\pm$ desvio padrão entre amostras sem nutrientes (NPK) e tratadas com nutrientes (N80P9). As reações foram feitas em triplicatas usando três réplicas biológicas, resultado de um *pool* de três plantas cada. Letras diferentes denotam diferenças estatisticamente significativas entre segmentos de raiz (valor de p em preto) e asteriscos mostram diferenças estatisticamente significativas entre tratamentos no mesmo segmento (valor de p em vermelho).

O padrão transcricional de ambos fatores de transcrição das famílias MYB (SCRLLR1059G12.g) e NAC (SCCCLR2003C07.g) não apresentou diferenças estatisticamente significativas ao longo da formação do aerênquima. (Figs. 12A e B). Por outro lado, o fator de resposta a etileno 1 (SCCCCL4002B07.g) apresentou maior nível de transcritos em S1 do tratamento NPK, decaindo em direção no segmento seguinte e mantendo-se constante até S5 (Fig. 12C). Amostras de raiz coletadas de plantas cultivadas na escassez de nutrientes apresentaram menores níveis deste transcrito em S2, enquanto o oposto ocorreu em S4. Similarmente, o padrão de acúmulo do mRNA referente ao fator de transcrição RAV (SCAGLR2011B02.g) também foi distinto entre amostras dos dois tratamentos. No primeiro segmento, plantas cultivadas na escassez de nutrientes apresentaram níveis do mencionado transcrito quase três vezes superiores ao observado para plantas irrigadas com solução nutritiva (Fig. 12D). Nos demais segmentos o padrão de acúmulo deste mRNA permaneceu constante.

No que diz respeito às modificações na parede celular decorrentes da formação do aerênquima, foram avaliados os perfis transcricionais de genes relacionados aos processos de separação e expansão celular. Os transcritos relativos a uma expansina (SCVPRT2075H10.g; Fig. 13A) apresentaram maiores níveis no segmento 1 em ambos tratamentos, enquanto que aqueles referentes a XTH (SCBGLR1002B06.g; Fig. 13B) não apresentaram expressão diferencial ao longo da formação do aerênquima. Em plantas irrigadas semanalmente com solução nutricional os mRNAs correspondentes à expansina atingiram patamares mais elevados em S2 e S3, enquanto o oposto foi observado no segmento seguinte. Já os transcritos que codificam para uma endopoligalacturonase (SCEPFL3083G08.g; Fig. 13C) atingiram o ponto de maior acúmulo de S2 em diante e foram diferencialmente expressos nas amostras S2 a S4, apresentando maiores níveis em plantas em déficit nutricional. Os transcritos de uma pectina acetil esterase (SCAGRT2038D08.g; Fig. 13D) atingiram maiores níveis no tratamento NPK em S2, permanecendo estáveis nos segmentos seguintes.

RESULTADOS

Figura 12. Perfil de expressão de genes que codificam para fatores de transcrição das famílias MYB (A), NAC (B) e fatores de resposta a etileno (C e D). O eixo y está normalizado para expressão relativa ao valor de CNRQ obtido para o segmento 1 do tratamento N80P9 $\pm$ desvio padrão entre amostras sem solução nutritiva (NPK) e tratado com nutrientes (N80P9). As reações foram feitas em triplicatas usando três réplicas biológicas, resultado de um *pool* de três plantas cada. Letras diferentes denotam diferenças estatisticamente significativas entre segmentos de raiz (valor de p em preto) e asteriscos mostram diferenças estatisticamente significativas entre tratamentos no mesmo segmento (valor de p em vermelho).

Dentre os transcritos referentes a enzimas que atacam hemiceluloses, aqueles correspondentes a uma arabinofuranosidase (SCQSRT1034D03.g; Fig. 14A) mantiveram os níveis mais altos em todos segmentos exceto em S2, sendo diferencialmente expressos entre amostras NPK e N80P9 nos segmentos S2-S4. Já os transcritos referentes a uma β -xilosidase (SCCCRT1002G03.g; Fig. 14B) não apresentaram expressão diferencial ao longo da formação do aerênquima.

Dentre os mRNAs que codificam para enzimas de degradação de celulose, aqueles referentes a uma β -glucosidase (SCEQRT1029E06.g; Fig. 14C) não apresentaram diferenças estatisticamente significativas. Já nos níveis de transcritos correspondentes a endoglucanase (SCEQRT1029E06.g; Fig. 14D) atingiram maiores níveis em S1, S4 e S5, apresentando valores maiores para plantas tratadas com nutrientes.

RESULTADOS

Figura 13. Perfil de expressão de genes que codificam para proteínas relacionados à expansão (A e B) e separação celular (C e D). O eixo y está normalizado para expressão relativa ao valor de CNRQ obtido para o segmento 1 do tratamento N80P9 $\pm$ desvio padrão entre amostras sem nutrientes (NPK) e tratado com nutrientes (N80P9). As reações foram feitas em triplicatas usando três réplicas biológicas, resultado de um *pool* de três plantas cada. Letras diferentes denotam diferenças estatisticamente significativas entre segmentos de raiz (valor de p em preto) e asteriscos mostram diferenças estatisticamente significativas entre tratamentos no mesmo segmento (valor de p em vermelho).

Em razão do padrão de expressão coincidente com a formação de aerênquima e pela ocorrência de expressão diferencial entre os tratamentos, os genes referentes a uma endopoligalacturonase (SCEPFL3083G08.g; Fig. 13C) e fator de transcrição RAV (SCAGLR2011B02.g; Fig. 12D) foram selecionados para sequenciamento.

Figura 14. Perfil de expressão de genes que codificam para glicosil hidrolases relacionadas à degradação de hemiceluloses (A e B) e celulose (C e D). O eixo y está normalizado para expressão relativa ao valor de CNRQ obtido para o segmento 1 do tratamento N80P9 ± desvio padrão entre amostras sem nutrientes (NPK) e tratado com nutrientes (N80P9). As reações foram feitas em triplicatas usando três réplicas biológicas, resultado de um pool de três plantas cada. Letras diferentes denotam diferenças estatisticamente significativas entre segmentos de raiz (valor de p em preto) e asteriscos mostram diferenças estatisticamente significativas entre tratamentos no mesmo segmento (valor de p em vermelho).

3. Formação do aerênquima e dosagem hormonal diante da aplicação de 1-MCP

Visando bloquear a formação de aerênquima em toda extensão analisada das raízes, bem como avaliar o efeito indutor do etileno endógeno, um segundo cultivo de cana-de-açúcar foi realizado. A inibição da percepção ao etileno se deu por meio do emprego de 1-MCP. No que diz respeito ao cálculo da área ocupada pelo aerênquima no córtex das raízes, conforme observado em experimentos anteriores, o aerênquima não se formou nos dois primeiros segmentos a partir do ápice (Fig. 15). Este iniciou-se em S3, apresentando área proporcional (em relação a área ocupada em S5 do tratamento controle) de 0,04 e 0,08 em plantas controle e tratadas com 1-MCP, respectivamente. A extensão do aerênquima no quarto segmento de raiz analisado atingiu valores e área proporcional de 0,48 e 0,35 na região cortical em plantas tratadas ou não com 1-MCP, respectivamente, alcançando valores máximos de área em S5, equivalentes em ambos tratamentos. As diferenças entre os valores de área calculados para um mesmo segmento de tratamentos diferentes não foram estatisticamente significativas ($p=0,88$) (Fig. 15A). Sendo assim, nas condições testadas o 1-MCP demonstrou não apresentar efeito na formação do aerênquima em raízes de cana-de-açúcar (Fig. 15B).

Figura 15. Padrão de ocupação do aerênquima no córtex de raiz de cana-de-açúcar após tratamento com 1-MCP. (A) Área da seção transversal do córtex ocupada pelo aerênquima em plantas cultivadas em na presença de 1-MCP. Medidas foram feitas em cinco diferentes raízes oriundas de *pools* de seis plantas de cada uma das três réplicas biológicas. Letras diferentes denotam diferenças estatisticamente significativas entre segmentos de raízes. A área do aerênquima é mostrada em proporção ao valor obtido no segmento 5 de plantas controle $\pm$ desvio padrão. (B) Seções transversais dos cinco segmentos de raiz de cana-de-açúcar cultivada em 1-MCP. Corado com Safranblau 1%. Aumento 10x.

Por outro lado, o tratamento com 1-MCP afetou a concentração de etileno e AIA livre nos segmentos de raiz (Fig. 16). Nas plantas tratadas com 1-MCP os níveis de AIA livre apresentaram queda na transição S1-S2, enquanto plantas controle mostraram valores inalterados deste hormônio entre esses dois segmentos (Fig. 16A). No entanto, as concentrações de AIA livre em S1 de plantas tratadas com 1-MCP foram 1,5 superiores aos observados para plantas controle. No segmento de raiz seguinte o oposto foi observado e plantas não tratadas apresentaram maiores níveis de AIA livre, também superior na magnitude de 1,5 vezes comparado a plantas cultivadas em 1-MCP.

Já no que diz respeito ao etileno, em plantas controle a concentração deste hormônio permaneceu inalterada nos dois primeiros segmentos. Em contrapartida, em plantas tratadas com 1-MCP os níveis de etileno decaíram de 0,12 em S1 para 0,05 $\text{nL.g}^{-1}.\text{h}^{-1}$ em S2. A concentração de etileno nas raízes de plantas tratadas com 1-MCP foram 2 e 4,6 vezes menores que os observados para plantas controle nos segmentos 1 e 2, respectivamente (Fig. 16B). Do terceiro segmento em diante, tanto os níveis de auxina quanto os de etileno se mantiveram constantes entre os segmentos e tratamentos, apresentando tendência de equilíbrio entre as proporções das concentrações de um hormônio em relação ao outro (Fig. 16C).

Figura 16. Dosagem de hormônios em segmentos de raiz após tratamento com 1-MCP. (A) Perfil de concentrações de AIA livre, (B) etileno e (C) proporção entre ambos hormônios em amostras dos dois tratamentos ao longo da raiz. Acima das barras em C está a proporção de auxina em relação a etileno em cada tratamento e em cada segmento. Medidas foram feitas utilizando três repetições biológicas, como resultado de *pools* de raízes de seis plantas cada. Letras diferentes denotam diferenças estatisticamente significativas entre as amostras de raízes de ambos os tratamentos. Em A e B são mostradas as médias das concentrações $\pm$ desvio padrão.

4. Expressão diferencial de genes avaliada por microarranjo após tratamento com 1-MCP

O resultado do microarranjo mostrou 122 genes mais expressos e 43 menos expressos em plantas tratadas com 1-MCP em relação às plantas controle (Anexos 1-5). Genes relacionados ao transporte foram a categoria funcional de genes diferencialmente expressos mais representada em S1 e S5 (Fig. 17A e E). Pertencente a esta categoria está o gene que codifica para uma pirofosfatase vacuolar (H^+ -PPase, SCEQRT1024C07.g), diferencialmente expresso em S1. Também neste segmento um gene anotado como transportador da família ABC (SCQSHR1024F04.g), da mesma categoria funcional, apresentou-se mais expresso. Além disso também uma glicosil hidrolase da família 1 (SCEQHR1082B01.g) foi mais expressa no tratamento com 1-MCP (Anexo 1).

RESULTADOS

Figura 17. Categorias funcionais dos genes diferencialmente expressos após aplicação de 1-MCP nos segmentos 1 (A), 2 (B), 3 (C), 4 (D) e 5 (E). Foram excluídas as categorias "unknown", "others" e "uncategorized". Numerais indicam a porcentagem de genes de cada categoria diferencialmente expressa.

Dentre os genes diferencialmente expressos no segmento 2, se destacam os pertencentes à categoria "Metabolismo de carboidratos" (Fig. 17B), na qual estão contidos os genes referentes a arabinofuranosidase (SCVPLB1020G05.g), β -1,3-glucanase (SCEZLB1006G08.g) e outra GH1 (SCVPRZ2042G03.g), distinta da diferencialmente expressa em S1. Além destes, também destacam-se aqueles que codificam para duas proteinases (SCBGLR1119H08.g e SCEZHR1048C09.g), uma expansina (SCEPRT2043B01.g), uma pirofosfatase vacuolar (SCEQRT1024C07.g) e dois fatores de transcrição da família AP2/EREBP (SCCCCL7002A04.g e SCVPRT2076E10.g). Todos os genes mencionados apresentaram expressão aumentada após o tratamento com 1-MCP (Anexo 2).

O gene que codifica para uma pirofosfatase vacuolar (SCEQRT1024C07.g) mais expresso em S2 também o é em S3, assim como outros genes relativos a β -1,3-glucanase (SCJLRT1023E06.g),

RESULTADOS

superóxido dismutase (SCEQRT2101A06.g) e expansina (SCEQRT2026A06.g), distinta da mais expressa em S2 (Anexo 3). Já outro transcrito referente a uma pirofosfatase vacuolar (SCSFSD1065B09.g) encontra-se menos expresso após o tratamento com 1-MCP. Neste segmento a categoria “Transdução de sinal” é a predominantemente diferencialmente expressa (Fig. 17C).

No segmento 4, as categorias com maior número de genes diferencialmente expressos são e “Metabolismo redox” e “Parede celular” (Fig. 17D). Dentre os genes contidos nesta última categoria, observa-se a mesma expansina mais expressa em S3 (SCEQRT2026A06.g), uma fucosiltransferase (SCEQRT2100C04.g) e uma peroxidase (SCQGLR2032G02.g), todos com expressão aumentada diante do tratamento com 1-MCP. No que diz respeito ao metabolismo hormonal, o gene relativo à ACC oxidase (SCJLLR1054A06.g) encontra-se mais expresso diante da aplicação de 1-MCP. Dentre os genes que codificam para pirofosfatases vacuolares, um apresentou maior expressão (SCEQRT1024C07.g) após tratamento com 1-MCP, enquanto outro demonstrou resultado oposto (SCSFSD1065B09.g) (Anexo 4).

Por fim, conforme mencionado acima, no quinto segmento genes categorizados como relacionados a transporte foram predominantes dentre os diferencialmente expressos (Fig. 17E). Dentre eles, um gene relativo a uma pirofosfatase vacuolar (SCEQRT1024C07.g) apresentou expressão aumentada após o tratamento com 1-MCP, enquanto o gene anotado como transportador ABC (SCCCHR1001C06.g), teve o padrão oposto de expressão. Já os genes referentes à hemoglobina não-simbiótica (SCCCLR2001E08.g) e *respiratory burst oxidase homolog* (RBOH, SCAGLR2011D02.g) foram mais expressos após aplicação do inibidor da percepção de etileno (Anexo 5). Esta última codifica uma subunidade enzimática chave da NADPH oxidase, envolvida na produção de espécies reativas de oxigênio.

5. Validação do microarranjo por qPCR

Para a validação por qPCR foram selecionados predominantemente os genes relacionados à parede celular, estresse oxidativo e morte celular programada, além daqueles relacionados à biossíntese e sinalização hormonal. Os experimentos de qPCR validaram 89% dos padrões observados no microarranjo.

No que diz respeito aos genes referentes à parede celular, dentre aqueles que codificam para as GH1 mais expressas em S1 (SCEQHR1082B01.g; Anexo 1) e S2 (SCVPRZ2042G03.g; Anexo 2) de acordo com o microarranjo, somente a segunda teve seu padrão de acúmulo de transcritos confirmado pela qPCR (Tabela 6). Similarmente, os genes que codificam para uma arabinofuranosidase (SCVPLB1020G05.g; Tabela 6) e expansinas (SCEPRT2043B01.g e SCEQRT2026A06.g; Tabela 6) também foram mais expressos em plantas tratadas com 1-MCP em S2 e S2-S4, respectivamente. Especialmente no caso das expansinas, o aumento de expressão foi de 12 e 48 vezes em relação aos respectivos controles em S3 e S4, respectivamente.

RESULTADOS

Tabela 6. Padrão de expressão gênica observada por qPCR

	Categoria funcional	Genes diferencialmente expressos	Expressão relativa ¹	Valor de p	Coincidente com microarranjo?
S1	Metabolismo de carboidratos	SCEQHR1082B01.g; GH1	0,78 ± 0,0219	0,01	Não
	Transporte	SCEQRT1024C07.g; H ⁺ -PPase	5,78 ± 0,164	< 0,0001	Sim
S2	ND ²	SCCCCL7002A04.g; AP2/EREBP	1,78 ± 0,038	< 0,0001	Sim
	Metabolismo de carboidratos	SCVPLB1020G05.g; arabinofuranosidase	2,32 ± 0,050	< 0,0001	Sim
	Transporte	SCEQRT1024C07.g; H ⁺ -PPase	5,06 ± 0,042	< 0,0001	Sim
	Outra	SCEZHR1048C09.g; cisteíno protease	1,27 ± 0,097	0,0100	Sim
	Transdução de sinal	SCVPRT2076E10.g; AP2/EREBP	1,34 ± 0,089	0,005	Sim
	Metabolismo de parede celular	SCEPRT2043B01.g; expansina	3,02 ± 0,099	< 0,0001	Sim
	Metabolismo de carboidratos	SCVPRZ2042G03.g; GH1	2,06 ± 0,028	< 0,0001	Sim
	Metabolismo de parede celular	SCEQRT2026A06.g; expansina	12,98 ± 0,078	< 0,0001	Sim
S3	Transporte	SCEQRT1024C07.g; H ⁺ -PPase	5,56 ± 0,081	< 0,0001	Sim
	Transporte	SCSFSD1065B09.g; H ⁺ -PPase	0,65 ± 0,027	< 0,0001	Sim
	Biossíntese de hormônios	SCJLLR1054A06.g; ACO	3,01 ± 0,172	< 0,0001	Sim
S4	Transporte	SCEQRT1024C07.g; H ⁺ -PPase	9,03 ± 0,154	< 0,0001	Sim
	Transporte	SCSFSD1065B09.g; H ⁺ -PPase	0,89 ± 0,065	< 0,0001	Não
	Metabolismo de parede celular	SCEQRT2026A06.g; expansina	48,16 ± 0,300	< 0,0001	Sim
	ND	SCCCLR2001E08.g; Hb	3,91 ± 0,048	< 0,0001	Sim
S5	ND	SCAGLR2011D02.g; RBOH	7,39 ± 0,055	< 0,0001	Sim
	Transporte	SCEQRT1024C07.g; H ⁺ -PPase	4,98 ± 0,454	< 0,0001	Sim

¹ Valor de expressão relativa ao valor de CNRQ obtido para o mesmo gene no mesmo segmento do tratamento controle

² Não determinado

Dentre os genes envolvidos na morte celular programada, os transcritos correspondentes à *respiratory burst oxidase homolog* (RBOH; SCAGLR2011D02.g; Tabela 6) e hemoglobina não-simbiótica (Hb; SCCCLR2001E08.g; Tabela 6) foram mais abundantes no quinto segmento de plantas tratadas com 1-MCP. Juntamente ao acúmulo de transcritos referentes a uma cisteíno protease (SCEZHR1048C09.g) em S2 após tratamento com 1-MCP (Tabela 6), os dados de qPCR confirmam o microarranjo no que diz

RESULTADOS

respeito ao aumento global da expressão de genes relacionados a morte celular programada e estresse oxidativo diante da aplicação de 1-MCP.

Em relação a genes relacionados à sinalização e biossíntese de hormônios, em S2 os transcritos referentes a dois genes correspondentes a fatores de resposta a etileno (SCCCCL7002A04 e SCVPRT2076E10.g; Tabela 6) apresentaram maior acúmulo após tratamento com 1-MCP. O mesmo foi observado para o mRNA da ACC oxidase (ACO; SCJLLR1054A06.g), mais abundante no quarto segmento diante da aplicação de 1-MCP (Tabela 6).

O único gene diferencialmente expresso entre amostras controle e aquelas tratadas com 1-MCP em todos segmentos codifica para uma H⁺-PPase. As sondas desenhadas para os transcritos referentes a essas bombas (SCEQRT1024C07.g e SCSFSD1065B09.g) mostraram maior expressão deste gene no microarranjo (Anexos 1-5), resultados estes confirmados por qPCR exceto para o transcrito SCSFSD1065B09.g em S4 (Tabela 6).

Dentre os fatores de transcrição diferencialmente expressos no conjunto de amostras tratadas com 1-MCP, encontram-se dois da família ERF (Tabela 6). Em ambos casos os genes referentes aos fatores de transcrição são mais expressos em 1-MCP, assim como o gene da arabinofuranosidase o é (Tabela 6). Proteínas 45-57% idênticas à sequência protéica predita a partir do SAS EREBP1 (SCCCCL7002A04.g; Tabela 6) demonstraram estar envolvidas em processos relacionados à resistência a estresses bióticos e abióticos em *O. sativa* (Cao *et al* 2006) e *S. lycopersicum* (Wang *et al.*, 2007). Em contrapartida, proteínas 74-78% similares à proteína predita com base na sequência do SAS SCVPRT2076E10.g (Tabela 6) estão envolvidas no amadurecimento de *Actinidia deliciosa* (Yin *et al* 2010) e *Malus domestica* (Tacken *et al* 2010). Neste último caso, foi sugerida a regulação da expressão de um gene que codifica para uma endopoligalacturonase.

Visando avaliar se o perfil transcricional coincidente entre fatores de transcrição e glicosil hidrolases possivelmente reflete a regulação transcricional do primeiro sobre o segundo, foram escolhidos para o sequenciamento o gene que codifica para um fator de resposta a etileno referente ao SAS SCVPRT2076E10.g (Tabela 6) e para uma arabinofuranosidase (SCVPLB1020G05.g; Tabela 6).

PARTE C: Seleção e sequenciamento de genes alvo

1. Características gerais dos BACs sequenciados

Uma vez que os genes que codificam para os fatores de transcrição RAV e ERF e para as glicosil hidrolases endopoligalacturonase e arabinofuranosidase foram selecionados a partir dos experimentos de disponibilidade de nutrientes e aplicação de 1-MCP, os mesmos foram triados em uma biblioteca de BACs para sequenciamento das versões *full length*. A triagem das sequências-alvo por meio da utilização do *pool* 3D construído para a linhagem de cana-de-açúcar R570 permitiu o rápido isolamento de 38 clones carregando fragmentos de DNA aos quais se alinharam os pares de iniciadores desenhados para endopoligalacturonase - SCAGLR2011B02.g (8), fator de transcrição RAV - SCEPFL3083G08.g (9), arabinofuranosidase - SCVPLB1020G05.g (10) e fator de transcrição de resposta ao etileno -

RESULTADOS

SCVPRT2076E10.g (11). Os números entre parênteses indicam a quantidade de BACs selecionados com cada par de iniciadores desenhados para amplificação de genes. Os clones apontados como positivos na triagem foram validados e extraídos. Após a extração, a integridade do material foi confirmada em gel de agarose e sua quantificação mostrou que havia DNA plasmidial com qualidade e quantidade suficiente para construção da biblioteca (dados não mostrados).

Após os procedimentos de construção da biblioteca, sequenciamento na plataforma 454 e montagem das *contigs*, os BACs foram analisados quanto à presença e posição dos genes alvo e respectivas sequências promotoras, bem como para a ocorrência de elementos de transposição e outros genes preditos. O levantamento destas características nos conjuntos de BACs encontra-se descrito na Tabela 7.

Em todos BACs foram identificadas sequências codantes, sendo SHCRBa_109_O23 aquele com maior número de genes preditos (26), enquanto SHCRBa_056_A24 e 198_M03 contém o menor número de genes observado, 5 no total. Já de acordo com a porcentagem de genes, SHCRBa_023_N06 e SHCRBa_028_C15 ocupam as posições de BACs com menor (6,8%) e maior (35,6%) proporção de sequências gênicas, respectivamente (Tabela 7).

Em média um gene foi predito a cada 6.284 pb. O BAC SHCRBa_056_A24 apresentou menor frequência de genes preditos (1 a cada 17 kb), enquanto o oposto foi observado para SHCRBa_224_P16 (1 gene a cada 727 pb). O material sequenciado apresenta um conteúdo médio em torno de 30% de elementos de transposição, cerca do dobro da porcentagem média de genes preditos. Quanto aos genes alvo, todos estão completos nos BACs sequenciados que cobriram também 2 kb à montante do ATG, possibilitando o estudo da provável região promotora. Uma exceção é o gene referente a arabinofuranosidase presente no BAC SHCRBa_054_G11, no qual cerca de 1970 pb à montante do ATG estão cobertos pelo sequenciamento (Tabela 7).

A partir deste ponto, o conjunto de diferentes versões de cada sequência alvo foram nomeados *ScEPG1* e *ScARA1* para as glicosil hidrolases *endopoligalacturonase* e α -L-*arabinosidase* e *ScRAV1* e *ScERF1* para os fatores de transcrição *relative to ABI3/VP1* e *ethylene response factor* de *Saccharum hybrid*, respectivamente. A mesma nomenclatura em caracteres não itálicos se refere às proteínas preditas a partir das sequências codantes. As sequências alvo e os respectivos *loci* foram então analisados quanto à sintonia e colinearidade em relação ao provável ortólogo em *S. bicolor*, bem como à conservação em comparação com os demais BACs hom(e)ólogos.

Tabela 7. Características gerais dos 38 BACs sequenciados da variedade R570 de cana-de-açúcar

	BAC	Extensão (pb)	Genes preditos	% Genes	% TEs	Posição do alvo
ScEPG1	SHCRBa_023_N06	90.632	11	06,89	46,51	2700 - 5746
	SHCRBa_028_C15	82.420	16	35,64	26,38	59401 - 62180
	SHCRBa_036_E16	76.139	18	26,41	25,82	39569 - 42579
	SHCRBa_069_B14	102.231	22	15,45	15,73	88199 - 91197
	SHCRBa_095_L16	94.039	20	16,48	14,02	79807 - 83002
	SHCRBa_191_M21	124.054	17	18,83	42,16	29069 - 32115
	SHCRBa_224_P16	145.38	20	21,81	32,08	52945 - 56404
	SHCRBa_237_A10	99.495	10	21,65	58,18	48814 - 50312
ScRAV1	SHCRBa_021_A10	120.641	12	15,75	21,47	85163 - 86377
	SHCRBa_023_O12	164.282	11	12,67	25,00	34326 - 35540
	SHCRBa_042_C10	122.405	10	16,42	26,74	22094 - 23308
	SHCRBa_052_D15	127.452	12	23,32	18,76	4844 - 6018
	SHCRBa_056_A24	88.009	5	08,03	38,43	39189 - 40403
	SHCRBa_170_P07	118.045	8	16,29	19,98	6891 - 8087
	SHCRBa_174_O12	101.798	9	13,39	41,89	68297 - 68517
	SHCRBa_198_M03	82.381	5	09,27	36,20	41693 - 42907
ScARA1	SHCRBa_199_K01	128.611	11	23,15	25,05	48404 - 49609
	SHCRBa_048_E22	113.909	17	13,31	25,96	83838 - 89008
	SHCRBa_054_G11	102.985	15	12,18	27,40	95054 - 101018
	SHCRBa_109_O23	129.393	26	13,57	21,41	69325 - 74465
	SHCRBa_111_H15	104.864	13	13,59	39,35	20114 - 25283
	SHCRBa_124_A06	192.948	26	22,85	40,19	41290 - 57083
	SHCRBa_125_A09	101.988	13	19,80	12,96	67198 - 72378
	SHCRBa_133_H06	89.222	19	13,96	25,92	30039 - 35180
ScERF1	SHCRBa_143_N21	138.807	13	10,25	52,26	20113 - 25282
	SHCRBa_191_E06	99.815	9	12,12	51,23	2629 - 7798
	SHCRBa_220_C01	126.767	14	10,78	21,52	63184 - 69148
	SHCRBa_056_P16	89.984	12	21,03	10,27	61048 - 61812
	SHCRBa_062_K05	126.792	18	16,28	05,44	60146 - 60904
	SHCRBa_084_H10	162.354	19	16,78	27,17	94367 - 95125
	SHCRBa_095_D17	124.738	14	19,87	36,51	105240 - 106028
	SHCRBa_168_I11	112.459	16	23,47	29,68	26415 - 27203
ScERF1	SHCRBa_188_L14	100.743	14	08,04	07,41	8779 - 9537
	SHCRBa_180_G11	103.821	12	21,84	34,29	52514 - 53273
	SHCRBa_191_C20	119.673	9	11,39	56,11	72594 - 73353
	SHCRBa_193_M18	74.690	12	21,75	30,60	65174 - 65962
	SHCRBa_216_I12	100.394	9	15,72	42,50	56295 - 57053
	SHCRBa_247_B02	132.071	18	22,60	14,03	16578 - 17342

PARTE D: Análise de sintenia, filogenia e estrutura dos genes alvo sequenciados**1. Análises de colinearidade e sintenia dos BACs hom(e)ólogos em relação a *S. bicolor***

Em razão da alta similaridade entre as sequências codantes obtidas (Anexos 6-9), bem como a presença de versões de genes aparentemente não funcionais, aspectos relativos à conservação dos *loci* gênicos e também da estrutura gênica foram avaliados usando *S. bicolor* como base de comparação. Além disso tal análise visou identificar se tais sequências tratavam-se de genes hom(e)ólogos e quão variantes entre si e *S. bicolor* elas se apresentavam.

RESULTADOS

Após identificados os genes ortólogos em *S. bicolor* e eliminados os genes preditos que não atendiam os critérios estabelecidos (vide metodologia, item 1 Parte D), os *loci* correspondentes foram empregados como referência nas análises de sintonia. Sendo assim a numeração dos genes de cada BAC de cana-de-açúcar obedeceu a ordem representada em *S. bicolor*. A ilustração dos *loci* dos quatro genes alvo se deu por meio da representação esquemática de genes como setas (pretas e brancas), e regiões intergênicas em linhas pretas. Além disso, encontram-se representados os elementos de transposição de DNA e RNA em caixas pontilhadas e cinzas, respectivamente (Figs. 18-21).

De modo geral, é possível observar em todas as figuras a colinearidade de genes de cana-de-açúcar entre si, demonstrando tratar-se de genes hom(e)ólogos. Além disso, a colinearidade destes em relação aos ortólogos em *S. bicolor* também é mantida. Entretanto, a perda de colinearidade entre as duas espécies e mesmo entre os BACs hom(e)ólogos de cana-de-açúcar deve-se a presença uma diversidade de TEs característica de cada BAC. Além disso, são evidentes numerosos casos em que os genes ortólogos não foram identificados na região selecionada do genoma de *S. bicolor* (genes indicados por letras nas Figuras 18-21). Mesmo assim, considera-se que as regiões representadas de cana-de-açúcar em relação às de *S. bicolor* são sintênicas. Isso se deve ao fato de que na grande maioria dos casos foi encontrado um ortólogo em cana-de-açúcar para os genes preditos nos *loci* representados de *S. bicolor*. Exceções e outras observações pontuais relevantes serão discutidas caso a caso abaixo, de acordo com o gene alvo.

• **ScRAV1**

Observa-se que as nove versões do gene *ScRAV1* e seu hom(e)ólogo em *S. bicolor* encontram-se vizinhas à uma sequência anotada como codificadora de uma lipídeo fosfato fosfatase (8; Fig. 18) à jusante na posição *tail-to-tail* (em orientações opostas, com a região 5' do gene *ScRAV1* vizinha à região 5' do gene referente à lipídeo fosfato fosfatase), exceto nos BACs em que a mencionada região não foi contemplada no sequenciamento. Entretanto, à montante do gene alvo em cana-de-açúcar nos BACs SHCRBa_021_A10, SHCRBa_023_O12, SHCRBa_042_C10, SHCRBa_174_O12, SHCRBa_198_M03 e SHCRBa_199_K01 a sequência anotada como UDP glucosil transferase foi identificada em cana-de-açúcar (A; Fig. 18) com valores de identidade de pelo menos 99% entre si em termos de sequência codante. O ortólogo não foi encontrado na região representada do cromossomo 3 de *S. bicolor*, e sim no cromossomo 6 (Sobic.006G124100). Esta é de 73 a 85% idêntica às sequências codantes deste gene nos BACs de cana-de-açúcar. Em *Z. mays* (GRMZM2G131928, Chr03) e *O. sativa* (LOC_Os07g32620), genes com sequência codante com considerável identidade (68-77% e 66-77%, respectivamente) também foram encontrados quando as seis versões da UDP glucosil transferase de cana-de-açúcar foram usadas como consulta. Somente o gene proveniente de *Z. mays* encontra-se nas proximidades da região onde foi identificado o ortólogo de *ScRAV1* nesta planta (dados não mostrados). A proteína predita a partir deste gene no mencionado BAC apresenta um domínio predito da família 1 de glicosil transferases (GT1; PLN00164) e o mesmo domínio foi observado nas proteínas preditas de *S. bicolor* e *O. sativa* dos genes acima mencionados.

Figura 18. Relações físicas e filogenéticas dos BACs contendo o gene ScRAV1 comparados à sequência na qual se localiza o ortólogo em *S. bicolor* (Chr3: 60.193.612- 60.305.674). Os hom(e)ólogos de ScRAV1 estão representados por setas brancas e os demais genes por setas pretas. Regiões sombreadas em verde ilustram a colinearidade para o gene alvo e em azul para os demais genes. Sombreados em laranja denotam genes colineares em relação a *S. bicolor*, porém de orientação oposta. Transposons contendo *long terminal repeats* (LTR) estão representados por retângulos em cinza e transposons de DNA estão representados por retângulos pontilhados. Somente elementos de transposição com sequência maior que 3 kb são mostrados. A reconstrução filogenética por *neighbor-joining* foi inferida pelo método Tajima-Nei com 1000 repetições de *bootstrap*, cujos valores estão maiores que 70 estão representados na árvore. Retângulos azuis mostram os agrupamentos de BACs de acordo com os valores de *bootstrap*. Numeração e anotação de genes: (1) Sobc.003G265900, *serine/threonine protein kinase*; (2) Sobc.003G265800, *similar to Os12g0102600 protein* (domínio PPR repeat); (3) Sobc.003G265700, *serine/threonine protein kinase*; (4) Sobc.003G265600, *non-specific serine/threonine protein kinase*; (5) Sobc.003G265500, *peptidyl-tRNA hydrolase*; (6) Sobc.003G265400, *similar to Os01g0693800 protein* (domínio pyridoxal-phosphate dependent enzyme); (continuação da legenda no rodapé da página seguinte).

RESULTADOS

Dentre os demais genes cujo ortólogo não foi identificado na região representada do genoma de *S. bicolor*, encontra-se a sequência anotada como gene hipotético (B; Fig. 18). Este encontra-se presente nos BACs SHCRBa_052_D16 e 174_O12, com sequência codificante 95% idênticas entre si, não apresentando domínios protéicos preditos. Por fim, a sequência anotada como fasciclina (C; Fig. 18) no BAC SHCRBa_174_O12 apresenta o domínio *fasciclin superfamily* (pfam02469) e é 94% idêntica ao gene Sobic.003G250700. Porém este posiciona-se a cerca de 1300 kb à montante do gene ortólogo a *ScRAV1*, logo fora da região ilustrada na Figura 18. Situação similar foi observada para a sequência anotada como ATPase (D; Fig. 18) no mesmo BAC. Esta é 92% idêntica ao gene Sobic.003G250900, localizado em região vizinha ao gene Sobic.003G250700 supracitado, estando, portanto, também fora da região representada do cromossomo 3 de *S. bicolor*.

Foram observados alguns casos em que a colinearidade foi perdida entre genes de cana-de-açúcar e o provável ortólogo em *S. bicolor* em razão das orientações serem opostas (sombreados em laranja; Fig. 18). O gene 4, anotado como quinase de serina-treonina encontra-se na mesma orientação quando comparadas às sequências presentes em *S. bicolor* e no BAC SHCRBa_052_D15. Entretanto, ainda que este mesmo gene esteja também nos BACs SHCRBa_021_A10 e 023_O12, sua orientação é oposta. Já no BAC SHCRBa_174_O12 a orientação do gene 10, anotado como RHG-1 *like* encontra-se invertida em relação aos genes hom(e)ólogos e ao ortólogo. Também o gene 2, cuja anotação é *similar to* Os12g0102600 *protein* em *S. bicolor*, apresenta orientação oposta nesta planta em relação ao que é observado nos BACs SHCRBa_52_D15 e 023_O12.

Por fim, o gene cuja anotação funcional é *lysosomal acid lipase-related* (11; Fig. 18) encontra-se duplicado no BAC SHCRBa_042_C10. A sequência codificante de ambas cópias são altamente idênticas (95-97%) ao ortólogo em *S. bicolor* (Sobic.003G264800). Em outros genomas como o de *O. sativa* e *Z. mays*, o ortólogo foi identificado, mas também não se encontra duplicado nestas plantas.

Devido ao fato que apenas o gene alvo estava presente em todos os BACs e na sequência representada do cromossomo 3 *S. bicolor*, apenas sua sequência codante foi utilizada para a reconstrução filogenética. O resultado mostra que os BACs podem ser divididos em dois grupos (retângulos azuis; Fig. 18), de acordo com o maior valor de *bootstrap* obtido na filogenia (99%).

• *ScEPG1*

Todos os genes identificados na região representada de *S. bicolor* foram identificados nos BACs de cana-de-açúcar. Nos genes que flanqueiam o gene alvo a colinearidade é bastante conservada principalmente à jusante de *ScEPG1* (Fig. 19). Em todos BACs de cana-de-açúcar a sequência anotada como proteína ribossomal L5P está presente, assim como em *S. bicolor* (Sobic.001G524900) (7; Fig. 19).

Continuação da legenda da Figura 18: (7) Sobic.003G265300, *similar to putative lectin-like receptor kinase*; (8) Sobic.003G265100, *hypothetical gene* (anotação funcional *lipid phosphate phosphatase*); (9) Sobic.003G265000, *weakly similar to putative uncharacterized protein* P0034C09.35; (10) Sobic.003G264900, *protein RHG1a-like* (11) Sobic.003G264800 *similar to* Os01g0692600 *protein* (anotação funcional *lysosomal acid lipase-related*); (gene alvo) Sobic.003G265200, *factor RAV-like*. Indicação e anotação de genes de cana-de-açúcar cujos ortólogos não foram identificados na região ilustrada do cromossomo 3 de *S. bicolor*: (A) UDP glucosil transferase; (B) *hypothetical gene*; (C) *fascilin*; (D) ATPase

RESULTADOS

Entretanto, somente em quatro dos oito BACs de cana-de-açúcar (SHCRBa_023_N06, 069_B14, 095_L16, 191_M21) o gene anotado como gene hipotético (6; Fig. 19) é aquele imediatamente à montante do gene alvo, embora com orientação invertida em relação a *S. bicolor*. Dentre os demais BACs em que a região à montante de *ScEPG1* foi contemplada, o gene hipotético A encontra-se na região vizinha ao gene alvo nos BACs SHCRBa_036_E16 e 224_P16. Por fim, no BAC SHCRBa_237_A10 o gene pectato liase (D; Fig. 19) é a sequência imediatamente à montante de *ScEPG1*.

As sequências identificadas pelo número 6 (Fig. 19) e anotadas como galactose oxidase em cana-de-açúcar são de 91 a 99% idênticas entre si, das quais somente a sequência oriunda do BAC SHCRBa_224_P16 não apresenta o domínio F-box. Tais domínios são comumente presentes em proteínas envolvidas em processos de degradação via proteassomo (Skowrya *et al.*, 1997; Ho *et al.*, 2006). Estas sequências codantes são de 89 a 100% idênticas ao gene Sobic.001G525100 de *S. bicolor* presente na região representada do cromossomo 1. Entretanto, em *S. bicolor* tal sequência está anotada como gene hipotético e não tem predição de domínios, além de apresentar orientação oposta às sequências de cana-de-açúcar (sombreados em laranja; Fig. 19). Além disso, a sequência codificante do provável ortólogo da galactose oxidase de *S. bicolor* tem somente 345 pb, enquanto as de cana-de-açúcar têm entre 1193 e 1239 pb. No entanto, o alinhamento das sequências da galactose oxidase de cana-de-açúcar com a região do provável ortólogo em *S. bicolor* demonstrou que a sequência completa deste gene está presente no mencionado trecho em *S. bicolor*, mas a predição presente na base de dados do genoma possivelmente contém erros.

Dentre os genes cujos ortólogos não foram identificados na região supracitada do cromossomo 1 de *S. bicolor*, estão os genes A-D (Fig. 19). Os genes hipotéticos identificados por A só estão presentes nos BACs SHCRBa_036_E16 e 224_P16 à montante do gene alvo conforme mencionado acima, não existindo predição de domínios protéicos. Já a sequência anotada como fator de elongação S-II (B; Fig. 19) no BAC SHCRBa_028_C15 apresenta o domínio *TFIIS helical bundle-like domain* (pfam08711) e a respectiva sequência codante é 92% idêntica à do gene Sobic.005G049000. Este localiza-se no cromossomo 5 e, portanto, fora da região representada na Figura 19. Por fim, a sequência anotada como proteína de ligação ao oxisterol (C; Fig. 19) nos BACs SHCRBa_028_C15 e 224_P16, não apresenta domínios protéicos preditos e a sequência codificante correspondente é 93% idêntica ao gene Sobic.001G421700. Este gene posiciona-se a cerca de 1200 kb à montante da região representada do cromossomo 1 de *S. bicolor* na Figura 19.

Figura 19. Relações físicas e filogenéticas dos BACs contendo o gene *ScEPG1a* comparados à sequência na qual se localiza o ortólogo em *S. bicolor* (Chr1: 71.679.279 - 71.777.845). Os hom(e)ólogos de *ScEPG1* estão representados por setas brancas e os demais genes por setas pretas. Regiões sombreadas em verde ilustram a colinearidade para o gene alvo e em azul para os demais genes. Sombreados em laranja denotam genes sintênicos em relação à *S. bicolor*, porém de orientação oposta. Transposons contendo *long terminal repeats* (LTR) estão representados por retângulos em cinza. Somente elementos de transposição com sequência maior que 3 kb são mostrados. A reconstrução filogenética por *neighbor-joining* foi inferida pelo método Tajima-Nei com 1000 repetições de *bootstrap*, cujos valores estão maiores que 70 estão representados na árvore (continuação da legenda no rodapé da página seguinte).

RESULTADOS

O gene identificado pela letra D está anotado como uma pectato liase, enzima responsável pela clivagem eliminativa de pectina de-esterificada, gerando oligossacarídeos com resíduos de galacturonosil insaturados na extremidade não redutora (Carpita & Gibeaut, 1993). Este está presente nos BACs SHCRBa_224_P16 e 237_A10 e apresentam sequência codificante 94 e 89% idênticas ao gene ortólogo a *ScEPG1* em *S. bicolor*. Já os resultados de alinhamento dois a dois empregando sequências codificantes, revelou que tais genes de ambos os BACs são de 94 a 100% (SHCRBa_224_P16) e 89 a 100% idênticas (SHCRBa_237_A10) aos genes alvo *ScEPG1*. A duplicação deste gene não foi observada *S. bicolor*, *Z. mays* ou *O. sativa*. O mesmo domínio GH28 presente nas versões de *ScEPG1* não foi observado em nenhum dos genes D.

A colinearidade do gene que codifica para uma proteína ribossomal (7; Fig. 19) à jusante do gene alvo possibilitou que a sequência codificante do mesmo fosse concatenada à sequência codificante de *ScEPG1*, assim como foi realizado para os respectivos ortólogos em *S. bicolor*. De acordo com os altos de valores *bootstrap* (95 e 91%), na reconstrução filogenética ficam evidentes dois agrupamentos, um com os BACs SHCRBa_023_N06 e 191_M21 um segundo grupo com os demais BACs.

- **ScERF1**

A análise da região contendo o gene ortólogo de *ScERF1* em *S. bicolor* revelou a presença de 12 outros genes cujos ortólogos também foram observados BACs de cana-de-açúcar. Nos BACs em que a região à jusante do gene alvo foi contemplada pelo sequenciamento observa-se a presença do gene *weakly similar to putative uncharacterized protein*, colinear entre *S. bicolor* e na maioria dos BACs de cana-de-açúcar (8; Fig. 20). No caso dos BACs SCHRBa_084_H10 e 180_G11 a colinearidade é interrompida pela presença dos genes B ou B e C, respectivamente.

No que concerne ao *locus* ocupado à montante de *ScERF1*, o gene que codifica para uma proteína da superfamília de transportadores transmembrana *major facilitator* superfamily (A; Fig. 20) não apresenta ortólogos na região representada de *S. bicolor*. Este foi identificado em todos BACs, com exceção de SHCRBa_188_L14, apresentando o domínio *nodulin-like* (pfam06813). O alinhamento local no

Continuação da legenda da Figura 19: Retângulos azuis mostram o agrupamento de BACs de acordo com os valores de *bootstrap*. Numeração e anotação de genes: (1) Sobic.001G525600, *similar to fringe protein, putative, expressed*; (2) Sobic.001G525500, *similar to protein kinase domain containing protein, expressed* (3) Sobic.001G525400; *similar to protein kinase, putative, expressed*; (4) Sobic.001G525300, *weakly similar to expressed protein*; (5) Sobic.001G525200, *hypothetical gene*; (6) Sobic.001G525100, *hypothetical gene*; (7) Sobic.001G524900, *similar to 50S ribosomal protein L5, chloroplast precursor*; (8) Sobic.001G524800, *hypothetical gene*; (9) Sobic.001G524700, *hypothetical gene*; (10) Sobic.001G524600, *hypothetical gene*; (11) Sobic.001G524500, *similar to HAD-superfamily hydrolase, subfamily IA, variant 3 containing protein, expressed*; (12) Sobic.001G524400, *similar to D111/G-patch domain-containing protein, putative, expressed*; (13) Sobic.001G524300, *similar to expressed protein (diacylglycerol kinase catalytic domain)* (14) Sobic.001G524200, *similar to protein kinase domain containing protein, expressed*; (15) Sobic.001G524100, *similar to expressed protein* (domínio *Det1* complexing ubiquitin ligase); (16) Sobic.001G524000; *hypothetical gene*; (gene alvo) Sobic.001G525000, *similar to polygalacturonase, putative, expressed*. Indicação e anotação de genes de cana-de-açúcar cujos ortólogos não foram identificados na região ilustrada do cromossomo 1 de *S. bicolor*: (A) *hypothetical gene*; (B); *transcription elongation factor (TFIIS) family protein*; (C) *oxysterol-binding_family_protein*; (D) *pectate lyase*.

RESULTADOS

genoma de *S. bicolor* revelou o gene Sobic.010G061800 como o mais similar à sequência da *major facilitator* de cana-de-açúcar, com valores de identidade de 77 a 81%, porém este não se encontra na região ilustrada do cromossomo 2 de *S. bicolor*.

Os genes hipotéticos 10 (Sobic.002G261000) e 11 (Sobic.002G261100) estão possivelmente duplicados em *S. bicolor*, enquanto em cana-de-açúcar somente uma destas sequências foi identificada (10/11; Fig. 20). No que diz respeito ao gene B, anotado como polimerase de DNA e RNA e presente nos BACs SHCRBa_180_G11 e 084_H10, apresenta alta similaridade com algumas regiões do cromossomo 1 de *S. bicolor*, mas em nenhuma delas há predição de genes. O mesmo resultado foi observado quando a busca se deu nos genomas de *O. sativa* e *Z. mays*. Já a sequência codificante do gene anotado como ribonuclease nos BACs SHCRBa_056_P16 e 180_G11 (C; Fig. 20) é 85% similar à do gene Sobic.002G046600. Embora localize-se no cromossomo 2, encontra-se a cerca de 60.000 kb da região representada na Figura 20. Por fim, o gene D que codifica para uma fosfatase de serina-treonina, só ocorre no BAC SHCRBa_247_B02 e apresenta alta similaridade com sequências contidas no cromossomo 1 de *S. bicolor*, porém nenhuma em região na qual há predição de genes.

A perda de colinearidade e sintenia na região vizinha ao gene alvo fez com que somente a sequência codante do fator de resposta a etileno em cana-de-açúcar e *S. bicolor* fossem empregadas na reconstrução filogenética. Os genes hom(e)ólogos oriundos dos BACs SHCRBa_056P16, 191_C20 e 247_B02 ocupam o grupo I, enquanto os demais BACs formam o agrupamentos II (retângulos azuis; Fig.20).

- **ScARA1**

A região representada do genoma de *S. bicolor* na qual se localiza o provável gene ortólogo de ScARA1 apresenta outros seis genes preditos (Fig. 21). Destes, todos foram identificados nos BACs homeólogos de cana-de-açúcar, porém com perda de colinearidade. Uma importante exceção é a ocorrência do gene anotado como *similar to putative uncharacterized protein* com domínio invertase/pectina metil esterase (4; Fig. 21), identificado em todos os BACs à jusante do gene alvo, em orientação *tail-to-tail*. Em alguns BACs à jusante do gene 4 observa-se a sequência anotada como Co-chaperona_GrpE (A; Fig. 21). Este está presente em sete dos dez BACs de cana-de-açúcar e o ortólogo não foi identificado na região representada do cromossomo 5 de *S. bicolor*. Embora existam domínios na sequência da proteína predita (F-box; CL0271), a região mais similar à sequência codificante deste gene no genoma de *S. bicolor* localiza-se no cromossomo 10, porém não há predição de genes. Além disso, resultados de alinhamento local em bancos de proteínas revelou alta similaridade somente com proteínas hipotéticas e preditas de gramíneas. Por sua vez, a sequência codante do gene anotado como hipotético (B; Fig. 21) é 77 a 100% idêntica à da chaperona. Entretanto, optou-se por identificá-los como genes distintos em razão da ausência do domínio F-box na proteína predita a partir do gene B. Este gene está presente em sete dos dez BACs da Figura 21, de uma a três cópias.

Figura 20. Relações físicas e filogenéticas dos BACs contendo o gene *ScERF1* comparados à sequência na qual se localiza o ortólogo em *S. bicolor* (Chr2: 64.480.316 – 64.600.167). Os hom(e)ólogos de *ScERF1* estão representados por setas brancas e os demais genes por setas pretas. Regiões sombreadas em verde ilustram a colinearidade para o gene alvo e em azul para os demais genes. Transposons com *long terminal repeats* (LTR) estão representados por retângulos em cinza e transposons de DNA estão representados por retângulos pontilhados. Somente elementos de transposição com sequência maior que 3 kb são mostrados. A reconstrução filogenética por *neighbor-joining* foi inferida pelo método Tajima-Nei com 1000 repetições de *bootstrap*, cujos valores estão maiores que 70 estão representados na árvore. Retângulos azuis mostram o agrupamento de BACs de acordo com os valores de *bootstrap*. Numeração e anotação de genes: (1) Sobic.002G259900, *hypothetical gene*; (2) Sobic.002G260000, *hypothetical gene*; (3) Sobic.002G260200, *similar to Os09g0510000 protein*; (4) Sobic.002G260400, *similar to similarity to unknown protein*; (5) Sobic.002G260500, *hypothetical gene*; (6) Sobic.002G260600, *similar to Os09g0510500 protein* (domínio *helix-loop-helix DNA-binding domain*); (7) Sobic.002G260700, *similar to Os09g0510700 protein* (domínio *putative GTPase activating protein for Arf*); (8) Sobic.002G260900, *similar to putative uncharacterized protein* (domínio *protein kinase*); (9) Sobic.002G261000, *weakly similar to putative uncharacterized protein* (continuação da legenda no rodapé da página seguinte).

RESULTADOS

Como mencionado acima, para a maioria dos genes a colinearidade foi perdida entre as sequências de cana-de-açúcar e *S. bicolor*. Um exemplo é o gene anotado como *predicted protein* (com domínio GATA Zn finger) (3; Fig. 21) que está presente seis dos dez BACs carregando o gene *ScARA1*. Enquanto em *S. bicolor* o gene 3 se localiza a cerca de 15 kb à montante do gene alvo, em cana-de-açúcar o provável gene ortólogo situa-se de 40 a 100 kb à montante de *ScARA1*. Além disso, em cana-de-açúcar a região compreendida entre GATA Zn finger e arabinofuranosidase é ocupada por um ou mais genes cujos ortólogos não existem em *S. bicolor* (genes D, E, G, H e I). Parece existir pelo menos um padrão mais evidente de ocupação destes loci. No caso dos BACs SHCRBa_111_H15, 143_N21 e 191_E06 os genes anotados como *tetratricopeptide repeat* (D), transducina (E) e quinase de diacilglicerol (H; Fig. 21) são observados. Já nos demais BACs de cana-de-açúcar, a colinearidade é fraca e dificultada pelo não sequenciamento desta região.

Foram encontrados numerosos casos de genes cujos ortólogos não foram identificados na região representada do cromossomo 5 de *S. bicolor*. Entretanto, alguns ainda têm sentença mantida, como é o caso dos genes *tetratricopeptide repeat* (D), nos BACs SHCRBa_109_O23, 111_H15 e 191_E06, e transducina (E), nos BACs SHCRBa_111_H15, 143_N21 e 191_E06 (Fig. 21). Entretanto, em ambos casos a região com maior identidade às sequências codantes de ambos genes está de 2000 a 5000 kb da sequência representada do cromossomo 5 de *S. bicolor*. Por outro lado, as sequências codantes dos genes anotados como hipotético (G), nos BACs SHCRBa_109_O23 e 125_A09 e quinase de diacilglicerol (H), nos BACs SHCRBa_109_O23, 111_H15; 133_H06, 143_N21, 191_E06 (Fig. 21) são altamente idênticas a sequências localizadas no cromossomo 8 de *S. bicolor*. No caso de G na região idêntica em questão no genoma de *S. bicolor* não há predição de genes. Já o gene Sobic.008G077600 cuja sequência codificante é 92% idêntica à do gene H também codifica uma quinase de diacilglicerol, apresentando o mesmo domínio protéico predito (DAGK_acc; pfam00609).

Por sua vez, o gene anotado como celulase (J; Fig. 21) está presente nos BACs SHCRBa_220_C01 e 054_G11. Não apresentando predição de domínios protéicos, a sequência codante é 81% idêntica à de Sobic.006G122300.1, do cromossomo 6 de *S. bicolor*. Por fim, os genes hipotéticos (C, no BAC SHCRBa_124_A06 e F, nos BACs SHCRBa_109_O02 e 133_H06) e exoribonuclease (I, nos BACs SHCRBa_109_O23, 125_A09, 133_H03 e 220_C01) (Fig. 21) não apresentam domínios protéicos preditos e a região com maior identidade a estas sequências no genoma de *S. bicolor* não apresenta genes preditos.

A concatenação do gene alvo com gene 4 possibilitou a reconstrução filogenética que se organiza em quatro grupos (Fig. 21). Embora estes dois genes representem somente uma pequena extensão dos BACs hom(e)ólogos, o agrupamento observado na filogenia reflete a colinearidade das sequências de cana -de-açúcar.

Figura 21. Relações físicas e filogenéticas dos BACs contendo o gene *ScARA1* comparados à sequência na qual se localiza o ortólogo de *S. bicolor* (Chr5: 2.030.596 - 2.083.021). Os hom(e)ólogos de *ScARA1* estão apresentadas por setas brancas e os demais genes por setas pretas. Regiões sombreadas em verde ilustram a colinearidade para o gene alvo e em azul para os demais genes. Transposons contendo *long terminal repeats* (LTR) estão representados por retângulos em cinza. Somente elementos de transposição com sequência maior que 3 kb são mostrados. A reconstrução filogenética por *neighbor-joining* foi inferida pelo método Tajima-Nei com 1000 repetições de *bootstrap*, cujos valores estão maiores que 70 estão representados na árvore. Numeração e anotação de genes: (1) Sobic.005G022100, *similar to transporter, putative, expressed* (domínio *triose-phosphate transporter family*); (2) Sobic.005G022200, *similar to putative 5'-3' exoribonuclease*; (3) Sobic.005G022400, *predicted protein* (domínio *GATA zinc finger*); (4) Sobic.005G022600, *similar to putative uncharacterized protein P0702G08.13* (domínio *plant invertase/pectin methylesterase inhibitor*); (5) Sobic.005G022700, *hypothetical gene*; (6) Sobic.005G022800, *similar to putative uncharacterized protein* (domínio *50S ribosome-binding GTPase*); (gene alvo) Sobic.005G022500, *similar to α -L-arabinofuranosidase C-terminus family protein, expressed* (continuação da legenda no rodapé da página seguinte).

2. Caracterização da estrutura gênica dos genes alvo hom(e)ólogos em relação a *S. bicolor*

Embora as sequências codantes dos genes hom(e)ólogos de cana apresentassem alta identidade entre si e em relação à *S. bicolor*, os tamanhos dos referidos genes apresentaram variações. Em razão disso, a quantidade e o tamanho de éxons e íntrons foi avaliado em ambas espécies. Além disso, diante do elevado conteúdo médio de elementos de transposição (Tabela 7) presentes nos BACs sequenciados, esse procedimento visou mapear a localização destes elementos.

Os genes hom(e)ólogos *ScERF1* apresentam estrutura relativamente simples, com a presença de um íntron de posição e extensão consideravelmente conservada entre as sequências hom(e)ólogas e também em relação a *S. bicolor* (Fig. 22A). A sequência genômica apresenta de 759 a 789 pb, e a codificante entre 642 e 705 pb. O provável ortólogo em *S. bicolor* se estende por 789 pb e a respectiva sequência codante apresenta 683 pb. Em todos hom(e)ólogos há predição do domínio protéico AP2, contando com cerca de 60 resíduos de aminoácidos responsáveis por ligação ao DNA (em azul na Figura 22A).

Todas as versões de *ScRAV1* sequenciadas, bem como o provável ortólogo em *S. bicolor* estendem-se por cerca de 1,2 kb. Dentre os quatro genes alvo sequenciados, *ScRAV1* é o único a não apresentar íntrons, com exceção da versão presente no BAC SHCRBa_052_D15 (Fig. 22B). Este homeólogo apresenta uma possível inserção a partir do 502º nucleotídeo, consistindo um fragmento de 25 pb que não apresenta identidade em relação aos demais hom(e)ólogos. Possivelmente trata-se de uma versão não funcional do gene *ScRAV1*. Todas sequências apresentam os dois domínios preditos de ligação ao DNA, AP2 (pfam00847) e B3 (pfam02362) (em azul e verde, respectivamente, na Figura 22B) característicos desta sub-família de fatores de transcrição. O primeiro se estende por cerca de 60 aminoácidos, enquanto o segundo apresenta em torno de 90. Exceto pela sequência de *ScRAV1* presente no BAC SHCRBa_052D15, todos homeólogos de *ScRAV1* e *ScERF1* de cana-de-açúcar apresentam mesma estrutura genômica em relação aos respectivos genes ortólogos de *S. bicolor* (Fig. 22).

No caso das estruturas dos genes que codificam para as glicosil hidrolases arabinofuranosidade e endopoligalacturonase, 1970 pb da sequência à montante do códon de início da tradução foram empregados em alinhamento múltiplo para reconstrução filogenética. Para *ScARA1*, de acordo com altos valores de *bootstrap* (98 a 100%) os BACs SHCRBa_109_O23 e 133_H06 e encontram-se em um dos grupos, divergindo dos demais (Fig. 23). Tais sequências promotoras foram mapeadas no que diz respeito à presença, número e posição de elementos *cis* para domínios de ligação ao DNA AP2 (RAV1AAT; Kagaya *et al.*, 1999), B3 (RAV1BAT; Kagaya *et al.*, 1999) e ERF (GCCCORE; Chakravarthy *et al.*, 2003) (setas laranjas, amarelas e azuis, respectivamente; Figs. 23B). Em alguns casos, os elementos RAV1AAT e GCCCORE estavam posicionados *in tandem* (setas verdes; Figs. 23B).

Continuação da legenda da Figura 21. Indicação e anotação de genes de cana-de-açúcar cujos ortólogos não foram identificados na região ilustrada do cromossomo 5 de *S. bicolor*: (A) *Co-chaperone_GrpE_family_protein*; (B) hypothetical gene; (C) hypothetical gene; (D) *Tetratricopeptide repeat (TPR)*; (E) *Transducin*; (F) *hypothetical gene*; (G) *hypothetical gene*; (H) *Diacylglycerol kinase*; (I) *exoribonuclease 2*; (J) *Cellulase*. *Três cópias identificadas do gene C.

Figura 22. Representação esquemática da estrutura gênica das versões do gene que codifica para (A) o fator de resposta a etileno 1, *ScERF1* e (B) para o fator de transcrição *relative to ABI3/VP1 1*, *ScRAV1* em cana-de-açúcar e nos respectivos ortólogos de *S. bicolor* (Sobic.003G265200, *Predicted protein* e Sobic.002G260800, *RAV-like factor*). Retângulos e linhas pretas indicam éxons e introns, respectivamente. Regiões em azul e verde denotam os domínios AP2 e B3 presentes nas proteínas preditas, respectivamente. A barra inclinada em vermelho indica a posição de uma possível inserção responsável por uma sequência provavelmente não funcional ("X" em vermelho). Para construção desta representação foi empregada a ferramenta *Exon-Intron Graphic Maker*. As barras de escala correspondem a 100 pb.

O padrão de mapeamento de ambos os sítios coincide com a divisão em dois agrupamentos (retângulos azuis; Fig. 23A). Nos BACs que ocupam o agrupamento II os sítios RAV1AAT são de 1,6 a 10 vezes mais numerosos que os GCCCORE, enquanto que para os do agrupamento I o número de GCCCORE é duas vezes superior ao de RAV1AAT (Fig. 23B).

Quanto à estrutura do gene, as versões de *ScARA1* são bem mais complexas que a dos fatores de transcrição supracitados. Na maioria dos casos, a estrutura gênica (número, tamanho e posição de introns e éxons) foi mantida em relação ao provável ortólogo de *S. bicolor*, que apresenta 17 éxons (Fig. 23C). Exceções são os hom(e)ólogos dos BACs SHCRBa_054_G11 e 133-H06, com 18 éxons e 220_C01 com 19. Os genes hom(e)ólogos apresentam sequência genômica que se estende de 5.142 a 5.965 pb, equiparável com os 5.570 pb observados para o provável ortólogo em *S. bicolor*.

Figura 23. Representação esquemática da estrutura gênica e promotora das versões do gene que codificam para a arabinofuranosidase 1 de cana-de-açúcar, *ScARA1*, e em *S. bicolor* (Sobic.005G022500 similar to α -L-arabinofuranosidase C-terminus family protein, expressed) (A) Reconstrução filogenética obtida a partir do alinhamento múltiplo das sequências de 1970 pb kb à montante do sítio de início da tradução (provável região promotora). A reconstrução filogenética por *neighbor-joining* foi inferida pelo método Tajima-Nei com 1000 repetições de *bootstrap*, cujos valores estão maiores que 70 estão representados na árvore. (B) Região de 1970 pb à montante do códon de iniciação da tradução (indicado por seta vermelha em L) na qual elementos RAV1BAT (setas amarelas) RAV1AAT (setas laranjas) e GCCCORE (setas azuis) foram mapeados (continuação da legenda no rodapé da página seguinte).

RESULTADOS

Embora apresente os mesmos 17 éxons presentes em *S. bicolor*, o gene que codifica para a arabinofuranosidase presente no BAC SHCRBa_124_A06 é uma outra exceção no que diz respeito à extensão dos íntrons. Dois grandes elementos de transposição encontram-se inseridos no 5° e 9° íntrons neste gene. Ambos são elementos contendo LTR, um Gypsy de 4.778 pb e um Copia, de 5.486 pb (Fig. 23C). Como consequência, este hom(e)ólogo tem 15.794 pb. Um outro TE mais curto, do tipo LINE, também está presente no hom(e)ólogo proveniente do BACs SHCRBa_220_C01.

As sequências codificantes por sua vez apresentam de 1.896 a 2.346 pb para cana-de-açúcar e 2.016 pb em *S. bicolor* (Fig. 23C). Em todas as versões existem domínios protéicos preditos correspondentes às famílias 4 e 9 de módulos de ligação a carboidrato (CBM4_9; pfam02018) e o domínio C-terminal de arabinofuranosidases (α-L-AF-C; pfam06964).

Para as regiões promotoras de *ScEPG1*, a reconstrução filogenética levou à divisão houve divisão em três grupos, com os BACs SHCRBa_023_N06 e 191_M21 ocupando o grupo I (Fig. 24A) e por sua vez apresentando 3 GCCCORE e 4 RAV1AAT. Os genes *ScEPG1* dos BACs SHCRBa_224_P16 e 028_C15 dispõem de 5 elementos RAV1AAT enquanto no agrupamento III este número é reduzido para 3 ou 2. Neste último grupo a região promotora de *ScEPG1* do BAC SHCRBa_237A10 destoa das demais (Fig. 24B).

Por fim, *ScEPG1* apresenta sequências genômicas de 2,999 kb a 3,456 kb, enquanto as sequências codantes se estendem por 1,077 a 1,830 kb. Em *S. bicolor*, a sequência genômica e codante tem 3,023 e 1392 kb, respectivamente. A estrutura genômica de *ScEPG1* é variável em um dos oito hom(e)ólogos (Fig. 24C). Enquanto as demais sequências de *ScEPG1* e o provável ortólogo se apresentam em 6 éxons, aquela oriunda do BAC SHCRBa_224_P16 organiza-se em 8. Tal padrão de *splicing* distinto provavelmente se deve à presença de um TE do tipo LINE de cerca de 500 pb localizado próximo à extremidade 3'.

Em todas as versões de *ScEPG1* há a predição do domínio protéico de aproximadamente 240 aminoácidos característico de glicosil hidrolases da família 28 (GH28; pfam00295). No hom(e)ólogo de *ScEPG1* proveniente do BAC SHCRBa_237_A10 a região do gene que codifica para este domínio predito encontra-se mais deslocado no sentido da extremidade 5' do gene. Além disso, neste gene há presença de um TE de 96 pb da família Gypsy no 5° éxon (Fig. 24C).

Continuação da legenda da Figura 23. Setas verdes indicam localização de GCCbox e RAV1AAT *in tandem*. Setas empilhadas indicam a ocorrência de sítios sobrepostos. (C) Estrutura dos hom(e)ólogos *ScARA1*. Retângulos e linhas pretas indicam éxons e íntrons, respectivamente. Regiões em vermelho denotam o domínio α-L-AF-C presente na proteína predita, regiões em amarelo marcam a região do módulo de ligação a carboidrato. Barras duplas em vermelho indicam interrupção na escala no 5° e 9° íntrons de *ScARA1*. Asteriscos e barras verdes indicam a presença de elementos de transposição. Para construção desta representação foi empregada a ferramenta Exon-Intron Graphic Maker. As barras de escala em C correspondem a 100 pb nas sequências de éxons, íntrons e promotores.

Figura 24. Representação esquemática da estrutura gênica e promotora das versões do gene que codificam para a endopoligalacturonase 1 de cana-de-açúcar, *ScEPG1* e em *S. bicolor* (Sobic.001G525000 similar to polygalacturonase, putative) (A) Reconstrução filogenética obtida a partir do alinhamento múltiplo das sequências de 2 kb à montante do sítio de início da tradução (provável região promotora). A reconstrução filogenética por *neighbor-joining* foi inferida pelo método Tajima-Nei com 1000 repetições de *bootstrap*, cujos valores estão maiores que 70 estão representados na árvore. (B) Região à montante do códon de iniciação da tradução (indicado por seta vermelha em L) na qual os elementos em *cis* RAV1AAT (setas laranjas) e GCCCORE (setas azuis) foram mapeados. Setas verdes indicam localização de GCCCORE e RAV1AAT *in tandem*. Setas empilhadas indicam a ocorrência de sítios sobrepostos. (C) Estrutura dos genes *ScEPG1*. Retângulos e linhas pretas indicam éxons e íntrons, respectivamente. Regiões em vermelho denotam o domínio GH28 presente na proteína predita e barras em verde indicam presença de elementos de transposição. Para construção desta representação foi empregada a ferramenta *Exon-Intron Graphic Maker*. As barras de escala em C correspondem a 100 pb nas sequências de éxons, íntrons e promotores.

PARTE E: Seleção da sequência para clonagem e análise *in silico* dentre as versões do gene alvo

1. Análise do perfil transcricional das cópias de *ScEPG1*, *ScRAV1*, *ScARA1* e *ScERF1*

Visando avaliar os padrões de expressão dos homeólogos de cada um dos quatro genes alvo, foi realizado o mapeamento por alinhamento dos *reads* do transcriptoma (RNAseq) de raiz sobre as sequências codificantes. A primeira (por *Burrows-Wheeler*) e a segunda (alinhamento por *Bowtie*) metodologias empregadas levaram em conta a quantidade de *reads* que mapearam as sequências codificantes de cada gene alvo (Figs. 25B, 26B, 27B e 28B). Uma terceira análise, realizada posteriormente também por *Bowtie*, normaliza o valor de *reads* pelo tamanho de cada uma das bibliotecas de transcriptoma de raiz (Anexo 10A) e pelo comprimento das sequências codificantes (Anexo 10B) (Figs. 25A, 26A, 27A e 28A).

Os resultados da primeira e da segunda análises mostraram que nem todos os hom(e)ólogos foram expressos nos tecidos analisados e condições experimentais estabelecidas (Figs. 25B, 26B, 27B e 28B). Para o gene *ScEPG1*, dentre as 8 sequências codificantes utilizadas nesta análise, somente metade foi coberta por *reads* durante o mapeamento, sendo aquela advinda do BAC SHCRBa_224_P16 a mais mapeada, saltando de 307 e 389 *hits* nos segmentos 1 e 2 e atingindo seu pico com 668 *hits* no S3 e, em

RESULTADOS

seguida, caindo para 610 *hits* no S4 (Fig. 25B). De acordo com o mapeamento, os demais hom(e)ólogos expressos atingiram patamares de *hits* pouco expressivos.

Figura 25. Mapeamento de *reads* do transcriptoma de raiz nas sequências codantes de *ScEPG1*. (A) Mapeamento em termos de valores médios de FPKM $\pm$ erro padrão para cada segmento de raiz. (B) Mapeamento em termos de valores médios de *hits* para cada segmento de raiz.

Os *reads* do transcriptoma foram mapeados em quatro das nove sequências codantes referentes às versões do gene *ScRAV1*, sendo o hom(e)ólogo contido no BAC SHCRBa_021_A10 preferencialmente expresso (Fig. 26B). Esta apresenta um perfil crescente de acordo com as etapas iniciais da formação do aerênquima, também com pico em S3, quando apresenta 67 *hits*.

Figura 26. Mapeamento de *reads* do transcriptoma de raiz nas sequências codificantes de *ScRAV1*. (A) Mapeamento em termos de valores médios de FPKM $\pm$ erro padrão para cada segmento de raiz. (B) Mapeamento em termos de valores médios de *hits* para cada segmento de raiz.

Por sua vez, o gene codificador da arabinofuranosidase, *ScARA1*, teve oito de suas dez versões expressas (Fig. 27B), sendo o hom(e)ólogo presente no BAC SHCRBa_111_H15 o mais expresso. A quantidade de *hits* atinge em S3 níveis 7 e 2,5 vezes superiores aos observados para S1 e S2, respectivamente, decaindo em S4.

Por fim, o mapeamento empregando sequências codificantes correspondentes aos 11 hom(e)ólogos de *ScERF1* revelou que dentre as seis sequências expressas, aquela contida no BAC SHCRBa_188_L14 se destaca (Fig. 28B). Esta atinge valores em torno de duas vezes maiores que a segunda sequência codante com maior número de *hits* no mapeamento. No caso de *ScERF1* do BAC

RESULTADOS

SHCRBa_188_L14, não observamos o mesmo padrão de queda na transição entre os segmentos 3 e 4, como relatado para as sequências codantes mais mapeadas dos demais genes alvo.

Figura 27. Mapeamento de *reads* do transcriptoma de raiz nas sequências codificantes de *ScARA1*. (A) Mapeamento em termos de valores médios de FPKM $\pm$ erro padrão para cada segmento de raiz. (B) Mapeamento em termos de valores médios de *hits* para cada segmento de raiz.

Por outro lado, a partir do mapeamento com normalização não é possível sugerir uma sequência mais expressa, apresentando todas de modo geral os aproximados de FPKM (Figs. 25A, 26A, 27A e 28A), exceto pela sequência de *ScEPG1* do BAC SHCRBa_237A10, que não foi mapeada (Fig. 25A). De acordo com esta metodologia não é possível selecionar uma sequência mais expressa de cada gene alvo para os experimentos posteriores. Tal resultado foi obtido posteriormente ao mapeamento presente nas Figuras 25B, 26B, 27B e 28B, num momento em que todos os experimentos de ensaios de transativação e todas as análises de proteínas preditas (descritos a seguir) já haviam sido concluídos.

Figura 28. Mapeamento de *reads* do transcriptoma de raiz nas sequências codificantes de *ScERF1*. (A) Mapeamento em termos de valores médios de FPKM $\pm$ erro padrão para cada segmento de raiz. (B) Mapeamento em termos de valores médios de *hits* para cada segmento de raiz.

Levando em conta as duas primeiras análises (Figs 25B, 26B, 27B e 28B), as sequências dos genes que codificam para arabinofuranosidase, ERF, RAV e endopoligalacturonase cujas versões mapeadas por um maior número de *reads* encontram-se nos BACs SHCRBa_111_H15, 188_L14, 021_A10 e 224_P16, respectivamente. Portanto, essas foram as sequências escolhidas para os procedimentos

RESULTADOS

posteriores (ensaios de transativação e análises de proteínas preditas). Estas foram nomeadas *ScARA1a*, *ScERF1a*, *ScRAV1a* e *ScEPG1a*, nos quais o sufixo “a” indica a sequência mais expressa de acordo com o mapeamento empregando o RNAseq sem normalização (Figs. 23B, 24B, 25B e 26B).

2. Caracterização das proteínas preditas a partir da sequência selecionada de cada gene-alvo

A partir das sequências codificantes de *ScARA1a*, *ScERF1a*, *ScEPG1a* e *ScRAV1a* foram preditas as estruturas primárias das proteínas correspondentes, a fim de identificar o grau de conservação das enzimas de cana-de-açúcar em relação àquelas já estrutural e funcionalmente caracterizadas. Além disso, esta análise visou avaliar se os genes sequenciados possivelmente codificam proteínas funcionais. Em cada caso, as proteínas preditas foram comparadas a proteínas similares disponíveis em bancos de dados no que diz respeito à estrutura tridimensional, aminoácidos chave, domínios estruturais, função e conservação de sequência. A avaliação destas sequências evidenciou aspectos estruturais relevantes, dos quais ressaltamos a presença de domínios proteicos e sequências de direcionamento preditos na Tabela 8.

Tabela 8. Características gerais proteínas preditas a partir do hom(e)ólogo selecionado dentre os genes alvo

BAC	Gene alvo	Gene	Proteína	Domínios ^a	Localização subcelular	Massa molecular ^b
SHCRBa_021_A10	ScRAV1a	1215 pb	404 aa	AP2: 80-129 B3: 218-316	Nuclear	41,9
SHCRBa_224_P16	ScEPG1a	2625 pb	448 aa	GH28: 138-429	Extracelular	48,0
SHCRBa_111_H15	ScARA1a	5169 pb	658 aa	α -L-AF_C: 456-641	Extracelular	72,5
SHCRBa_188_L14	ScERF1a	1078 pb	213 aa	AP2: 83-134	Nuclear	22,6

^a Posições dos domínios relativas às sequências das proteínas

^b Em kDa, predita pela ferramenta *Compute pI/Mw*, disponível em http://web.expasy.org/compute_pi/

• *ScEPG1a*

A proteína predita a partir da sequência codante de *ScEPG1a* apresenta 409 aminoácidos e massa molecular predita de 48 kDa (Tabela 8). Além da identificação de um peptídeo-sinal de 29 aminoácidos (sombreado em verde; Fig. 29) e da localização extracelular prevista, esta proteína apresenta um domínio predito característico de glicosil hidrolases da família GH28 (sombreado em vermelho; Fig. 29). Dentre as endopoligalacturonases com estrutura tridimensional elucidada nesta família, constam somente enzimas de fungos e bactérias. Um exemplo é a endopoligalacturonase EndoPG-I produzida por *Stereum purpureum* que se dobra em estrutura do tipo β prisma (Shimizu *et al.*, 2002), assim como endopoligalacturonases bacterianas (van Santen *et al.*, 1999), de outros fungos (Pickersgill *et al.*, 1998; Cho *et al.*, 2001) e demais enzimas que atuam sobre pectinas (Petersen *et al.*, 1997; Scavetta *et al.*, 1999).

RESULTADOS

```

1  atggcaccgggaagatctcttcgccctgtggcggtggcgctgtggcggttcgctgctg
   M A P G R S L R P V A V A L W A A S L L 20
61  ctggcggtgctgctgcgcgcgcagggcgctgctccgggggctggatgccggaagcac
   L A A S C A R A G A S S G G P G C R K H 40
121  gtggcgaaggtcacggagtacggcggtgggggacgggaggacgctcaacacggcggtg
   V A K V T E Y G A V G D G R T L N T A A 60
181  ttgccagggcggtggcggtgacgtgcggcgcgcgcgacggcggtcggtggtg
   F A R A V A D L S R R A R D G G A A L V 80
241  gtgccggtgggaagtggctcacggggccttcaacctcacaagctgcttcacgtctac
   V P P G K W L T G P F N L T S C F T L Y 100
301  ctcgacgagggcgccgagatcctcgcgtcccaggacatgaaccattggccccctcatagct
   L D E G A E I L A S Q D M N H W P L I A 120
361  cccctgcccgtttacgggagaggaagggacgagcctggccaaggtacatcaatttcatt
   P L P S Y G R G R D E P G P R Y I N F I 140
421  ggaggatccaatctcactgacgtcatcatcacaggtaaaaatggaacaatcaacgggcag
   G G S N L T D V I I T G K N G T I N G Q 160
481  gggcaagtctggtgggacaagttccatgccaaaggagctcaagtcaccccgcgccacctc
   G Q V W W D K F H A K E L K S T R G H L 180
541  ctggagctcctctactctgataacatcatcatctcgaatgtcacctccatcaacgcgcca
   L E L L Y S D N I I I S N V T S I N A P 200
601  tactggaacctccaccctacctattgcaccaatgtgacctcagtggtcgtcaccattctc
   Y W N L H P T Y C T N V T I S G V T I L 220
661  gcgcccgtgaattcgctaacaccgatggaattgaccagattcttcctcgcgtgtcaag
   A P L N S P N T D G I D P D S S S H V K 240
721  atcaggagctgctacatcgtctcggcgacgactgctgcgcgtgaagagcggtgggac
   I E D C Y I V S G D C V A V K S G W D 260
781  gactacggcatcaggttcaacatgccagcagcacatcgtcatcaggaggtgacctgc
   E Y G I R F N M P S Q H I V I R R L T C 280
841  atctccccacgagcgccatgatcgcgtgggcagcgagatgtccggcgccatccgcgac
   I S P T S A M I A L G S E M S G G I R D 300
901  gtgcgcgccgaggacaacgtcgccatcaacacggagtcggccgtcaggtggatgttctgg
   V R A E D N V A I N T E S A V R W M F W 320
961  atgaccggcaactacgggcagcaccccgacaacacgtccaacccaacgccttgcccgag
   M T G N Y G Q H P D N T S N P N A L P E 340
1021  gtcaccggcatcaactacagcgacgtgttcgctgagaacgtcaccggcatcaactacagc
   V T G I N Y S D V F A E N V T G I N Y S 360
1081  gacgtgttcgctgagaacgtgacctggccggcaggatggagggtcatcccaacgacccc
   D V F A E N V T M A G R M E G I P N D P 380
1141  tacaccgggatctgcatatccaacgtcactgcccgcctcgcgccgaacgccaggagctg
   Y T G I C I S N V T A R L A P N A Q E L 400
1201  cagtggaaactgcaccaacgtcaaggggtcacctccgacgtctcgcccaagccgtgcca
   Q W N C T N V K G V T S D V S P K P C P 420
1261  gagctcggcgcgaggaggcaagccgtgcgcttcccgggtggaagagctcgtcattggcca
   E L G A E G K P C A F P V E E L V I G P 440
1321  ccggagctgccaagtgtagctactga
      P E L P K C S Y - 448

```

Figura 29. Sequência nucleotídica (letras minúsculas) e de proteína predita (letras maiúsculas) a partir da sequência *ScEPG1a*, contida no BAC SHCRBa_224_P16 de cana-de-açúcar, que codifica para uma endopoligalacturonase. O peptídeo sinal predito encontra-se sombreado em verde e o domínio GH28 em vermelho. Resíduos em retângulos verdes, amarelos e azuis são aqueles críticos para a funcionalidade de endopoligalacturonases (Kester *et al.* 1996; Armand *et al.* 2000) e que se encontram conservados (amarelo) ou não (azuis) na sequência predita de cana. Resíduos em retângulos verdes indicam a provável díade catalítica (Pickersgill *et al.*, 1998). Resíduos em retângulos laranja e vermelho possivelmente são os responsáveis por conferir não-processividade e insensibilidade a proteínas inibidoras de poligalacturonase (*polygalacturonase inhibiting protein*, PGIP) à *ScEPG1* (van Pouderooyen *et al.*, 2003; Federici *et al.*, 2001).

RESULTADOS

A estrutura tridimensional da endopoligalacturonase II de *Aspergillus nidulans* revelou que os resíduos Asp180, Asp201, Asp202, His223, Arg256 e Lys258 são bastante conservados nestas enzimas (Pickersgill *et al.*, 1998). Destes seis resíduos, cinco estão presentes na proteína predita a partir de ScEPG1a (Asp229, Asp261, Asp262, His292 e Arg316, retângulos amarelos e verdes; Figs. 29 e 30). Além disso, a região onde supostamente se localiza a díade catalítica na enzima de cana-de-açúcar é bastante conservada entre cana-de-açúcar e enzimas de microrganismos (retângulos verdes; Fig. 30). O resíduo Lys258 da endopoligalacturonase II de *A. niger* não tem equivalente na proteína predita a partir de ScEPG1a, tampouco foram identificados outros resíduos positivamente carregados nas proximidades da região onde este resíduo seria esperado na endopoligalacturonase de cana-de-açúcar. Na posição correspondente (318) é possível identificar um resíduo de metionina (retângulo azul, Figs. 29 e 30).

Figura 30. Alinhamento múltiplo do *core* catalítico de endopoligalacturonases de *Aspergillus niger* e de gramíneas. ScEPG1a, proteína predita a partir do gene homônimo de cana-de-açúcar; Sbicolor, proteína predita a partir do gene Sobic.001G525000.1 de *S. bicolor*; Aniger_Pgal, poligalacturonase I (CAA41693.1) de *A. niger*; Aniger_Pgall, poligalacturonase II (CAA41694.1) de *A. niger*. Retângulos amarelos e verdes denotam resíduos conservados em relação à Pgall de *A. niger* e críticos para a funcionalidade da enzima do fungo. Resíduos em retângulos verdes sinalizam a provável diade catalítica de ScEPG1a enquanto retângulos azuis denotam resíduo de aminoácido não conservado, ambos de acordo com a enzima de *A. niger* (Pickersgill *et al.*, 1998). Retângulo em laranja marca o resíduo Arg96 que define a processividade de Pgal (van Pouderoyen *et al.*, 2003) e em vermelho os dois resíduos de triptofano nas proteínas de planta que possivelmente conferem insensibilidade à PGIP (Federici *et al.*, 2001). À direita do alinhamento são mostradas as posições do último resíduo de aminoácido de cada proteína. Abaixo da sequência, resíduos idênticos (*), conservados (:) e semiconservados (.) são marcados conforme gerado pela ferramenta ClustalW.

A presença de um resíduo de serina na posição 124 (retângulo laranja Figs. 30) alinhado com o resíduo Arg96 da endopoligalacturonase I de *A. niger* sugere que a proteína predita de cana-de-açúcar seja não-processiva. Embora as endopoligalacturonases I e II de *A. niger* sejam 60% idênticas, atribui-se a não-processividade da endopoligalacturonase II à presença de um resíduo de serina na posição 96, enquanto na enzima processiva está presente um resíduo de arginina (van Pouderooyen *et al.*, 2003).

Uma característica fundamental na distinção entre endopoligalacturonases de plantas e de microrganismos é a presença de dois resíduos de triptofano (retângulos vermelhos, Figs. 29 e 30), identificados nas enzimas de cana-de-açúcar de *S. bicolor*. Tais resíduos possivelmente conferem insensibilidade à proteína inibidora de poligalacturonases (*polygalacturonase inhibitor protein*; PGIP; Federici *et al.*, 2001).

Os dois fatores de transcrição têm localização subcelular predita no núcleo (Tabela 8), dado condizente com a presença de sinais de localização nuclear (sombreados em vermelho; Figs. 31 e 32). Ambos os fatores de transcrição apresentam um domínio de ligação ao DNA do tipo AP2 (sombreados em azul; Figs. 31 e 32). Neste domínio foram identificados os motivos YRG e RAYD (sublinhados em azul; Figs. 31 e 32), que se estendem por cerca de 20 e 40 aminoácidos, respectivamente, cuja função é relevante a atividade de ligação ao DNA em fatores de transcrição da família AP2/EREBP (Okamuro *et al.*, 1997).

Figura 31. Sequência nucleotídica (letras minúsculas) e de proteína predita (letras maiúsculas) a partir da sequência *ScERF1a*, contida no BAC SHCRBa_188_L14 de cana-de-açúcar, que codifica para um fator de resposta a etileno. Aminoácidos sublinhados em azul correspondem a assinaturas do domínio AP2, sombreado em azul. Retângulos em azul sinalizam resíduos de aminoácidos responsáveis pela ligação ao DNA em AtERF1 de *A. thaliana* (Allen *et al.*, 1998). A provável sequência de localização nuclear está sombreada em vermelho. Numerais à esquerda e à direita indicam a posição de nucleotídeos e resíduos de aminoácidos, respectivamente.

RESULTADOS

```

1  atggacagcgccagcagcctcgtggacgacaccagcagcggcggtgggtggcgcgctcc
   M D S A S S L V D D T S S G G G G G A S 20
61  acggacaagctaagggctttggccgtcgccgcccgcctcgggcccccgctggagcgc
   T D K L R A L A V A A A S G P P L E R 40
121 atgggcagcggcgccagcgcggtcgtggacgcggccgagccgggcgccgaggcggactcg
   M G S G A S A V V D A A E P G A E A D S 60
181 ggcggtcccgccggtggcgcccgggcgcggtgggctgggcggggaagctgccgtcgtcc
   G G S A G G A A G A V G V G G K L P S S 80
241 aggtacaagggcggtggcgccagcccaatgggcggtggggcgcgagatctacgagcgc
   R Y K G V V P Q P N G R W G A Q I Y E R 100
301 caccagcgctgtgtggctcggcacgttcgcgggcgaggccgacgcggcgcgcgcttacgac
   H Q R V W L G T F A G E A D A A R A Y D 120
361 gtcgcggcgcagcgggtccggggcgcgacgcggtcaccaacttcgcccgtcgcggac
   V A A Q R F R G R D A V T N F R P L A D 140
421 accgacccggagcgcgcgagctccggttctcgcgtcccgtccaaggccgaggtc
   T D P D A A A E L R F L A S R S K A E V 160
481 gtcgacatgctccgcaagcacactacttcgacgagctcgcgcagaacaagcgcgccttc
   V D M L R K H T Y F D E L A Q N K R A F 180
541 gccgcccgcgcgctccgcgccacggccaccgcctcgtcgccaaccaacaccaacagc
   A A A A A S A A T A T A S S P T N T N S 200
601 aaccatttcctccctcgcgtcgccctccccgcgacggcgcgaggacacctcttcgacaag
   N H S S L A S P S P A T A R E H L F D K 220
661 acggtgacccccagcgcggtgggcaagctgaaccggctggtgatcccgaagcagcagcc
   T V T P S D V G K L N R L V I P K Q H A 240
721 gagaagcacttcccgcgtgcagctcccgtctgccggcgagatcaaggcggtgctcctc
   E K H F P L Q L P S A G G E I K G V L L 260
781 aacctggaggacgcccgggcaaggtgtggcggttcgctactcctactggaacagcagc
   N L E D A A G K V W R F R Y S Y W N S S 280
841 cagagctacgtgctcaccaagggctggagccgcttcgtcaaggagaaaggcctccaggcc
   Q S Y V L T K G W S R F V K E K G L Q A 300
901 ggcgacgtcgtcggcttctaccgctcgtccgcgccagcgcggcgccgacagcagcaag
   G D V V G F Y R S S A A S A G A D S S K 320
961 ctcttcacgactgcaagctgcggcccaacaccgcgaccgcctcgacgacgacaagcccc
   L F I D C K L R P N T A T A S T T T S P 340
1021 gtggggtcgtcgccggctccggcgccggtggcggaaggccgtgcgcctcttcggcgtcgac
   V G S S P A P A P V A K A V R L F G V D 360
1081 ctgctgacggcgccggccaccgcccggcgccagcggaggccatggccgcccgggtgc
   L L T A P A T A A A A P A E A M A A G C 380
1141 aagagagccagggacttggcctcgccccgcaggcggttcaagaagcagctcgtggag
   K R A R D L A S P P Q A A F K K Q L V E 400
1201 ctggcactagtgtag
      L A L V - 420

```

Figura 32. Sequência nucleotídica (letras minúsculas) e de proteína predita (letras maiúsculas) a partir do gene hom(e)ólogo ScRAV1a contido no BAC SHCRBa_021_A10 de cana-de-açúcar. Aminoácidos sublinhados em azul correspondem a assinaturas do domínio AP2, sombreado em azul. Regiões sombreadas em verde sinalizam ao domínio B3. Aminoácidos envolvidos em retângulos azuis e verdes correspondem aos responsáveis por contactar o DNA dos domínios AP2 (Allen *et al.*, 1998) e B3 (Yamasaki *et al.*, 2004), respectivamente. Resíduos sublinhados em verde representam o motivo repressor de B3 (BRD, *B3 repressor domain*) (Ikeda & Ohme-Takagi, 2009) e sombreado em vermelho a região do provável sinal de localização nuclear. Numerais à esquerda e à direita indicam a posição de nucleotídeos e resíduos de aminoácidos, respectivamente.

RESULTADOS

Os resíduos de Arg170 e Trp172 responsáveis por contactar o DNA em AtERF1 (Allen *et al.*, 1998) encontram-se presentes nos fatores de transcrição da supracitada família cujos genes foram sequenciados no presente trabalho (retângulos em azul; Figs. 31 e 32). Ambos os nucleotídeos são conservados nas regiões formadoras de uma α -hélice e três folhas β no domínio AP2 da proteína de *A. thaliana* (Fig. 33).

Figura 33. Alinhamento múltiplo da região correspondente ao domínio AP2 em proteínas de cana-de-açúcar (preditas a partir de ScERF1a e ScRAV1a), *A. thaliana* (At) e *N. tabacum* (Nt). As setas amarelas representam regiões de formação de estruturas secundárias do tipo folhas β e a hélice amarela, α -hélice. As barras horizontais pretas sinalizam os elementos YRG e RAYD do domínio e as setas vermelhas apontam resíduos responsáveis por contactar o DNA, conforme descrito por Allen e colaboradores (1998) para AtERF1. Proteínas empregadas no alinhamento e intervalo contemplado no alinhamento (sobrescrito): ScRAV1¹⁸⁰⁻¹³⁵ e ScERF1¹⁸⁹⁻¹⁴⁰ (sequências preditas a partir dos genes ScRAV1a e ScERF1a, presentes nos BACs SHCRBa_021_A10 e SHCRBa_188_L14 de cana-de-açúcar, respectivamente, sequenciados neste trabalho); NtERF1¹⁰³⁻¹⁶¹ BAA07321.1; NtERF4⁹⁷⁻¹⁵⁵ BAA07324.1; AtERF2¹¹⁵⁻¹⁷³ BAA32419.1; AtERF1¹⁴⁴⁻²⁰² BAA32418.1; AtERF5¹⁵⁴⁻²¹² BAA32422.1; NtERF2²⁵⁻⁸⁰ BAA07323.1; AtERF3²⁶⁻⁸³ BAA32420.1; AtERF4²³⁻⁸⁰ BAA32421.1; NtERF2²⁵⁻⁸² BAA07322.1. Acima do alinhamento são mostradas as posições dos aminoácidos, a partir do primeiro. Abaixo da sequência, resíduos idênticos (*), conservados (:) e semiconservados (.) são marcados conforme gerado pela ferramenta ClustalW.

O alinhamento entre a região do domínio AP2 das proteínas ScRAV1a e ScERF1a juntamente com outros fatores de transcrição da superfamília AP2/EREBP de *A. thaliana* e *N. tabacum* sugere similaridades estruturais entre as proteínas de cana-de-açúcar e a proteína empregada na cristalização, AtRAV1 (Fig. 33).

No que diz respeito ao domínio B3 de ScRAV1a, este apresenta uma assinatura típica denominada domínio repressor de B3 (BRD, *B3 repressor domain*) (Ikeda & Ohme-Takagi, 2009), sublinhado em verde na Figura 32. Já quanto aos aspectos estruturais, foi realizada uma análise comparativa entre a proteína de cana-de-açúcar e as proteínas da sub-família RAV de *A. thaliana*, dentre elas AtRAV1 (At1g13260.1), cuja estrutura tridimensional do mencionado domínio foi elucidada (Yamasaki *et al.*, 2004) (Fig. 34). A estrutura tridimensional desta região da proteína mostrou-se bem mais complexa que o observado para AP2, contando com 7 folhas β e duas α -hélices. A estrutura tridimensional da proteína de *A. thaliana* apresenta um total de 17 resíduos de aminoácidos, sobretudo Arg e Lys, distribuídos na superfície das regiões que contactam o DNA. Estes são, em sua maioria, conservados no domínio B3 da proteína predita a partir de ScRAV1a (Fig. 34).

Figura 34. Alinhamento múltiplo da região correspondente ao domínio B3 em proteínas de cana-de-açúcar (ScRAV1a) e *A. thaliana* (At). As setas amarelas representam regiões de formação de estruturas secundárias do tipo folhas β e a hélice amarela, α-hélices. Círculos indicam resíduos responsáveis por formar bolsos positivamente carregados (vermelhos) e contactar o DNA (azuis), conforme descrito por Yamasaki e colaboradores (2004) para AtRAV1 (At1g13260.1). As posições dos aminoácidos empregados no alinhamento estão indicadas nas laterais. Abaixo da sequência, resíduos idênticos (*), conservados (:) e semiconservados (.) são marcados conforme gerado pela ferramenta ClustalW.

• ScARA1a

Em razão da alta identidade da arabinofuranosidase de cana-de-açúcar em relação a membros da GH51, ScARA1a é provavelmente alocada nesta família. A maioria das arabinofuranosidases da família GH51 têm extensa sequência codificante (Lombard *et al.*, 2013), sendo as proteínas codificadas também grandes, de cerca de 50 a 70 kDa (Lee *et al.*, 2001; Rosli *et al.*, 2009). A estrutura de arabinofuranosidases desta família é bastante complexa, conforme observado para as três arabinofuranosidases de fungos contidas nesta família e com estrutura tridimensional já descrita (Paes, *et al.*, 2008; Souza *et al.*, 2011; Im *et al.*, 2012).

Três prováveis domínios protéicos foram identificados na proteína predita a partir do gene ScARA1a. Um deles é o domínio AbfA (COG3534; e-value $3,80e^{-40}$), característico de α-L-arabinofuranosidases que se estende do 238º resíduo de aminoácido até a extremidade C-terminal da proteína predita (sublinhado em vermelho; Fig. 35), sobrepondo-se com o domínio α-L-AF_C (pfam06964; e-value $1,29e^{-59}$) (sombreado em vermelho; Fig. 35). Este, por sua vez, consiste em aproximadamente 200 resíduos de aminoácidos presentes tanto em arabinofuranosidases (EC 3.2.1.55) bacterianas quanto de eucariotos. Tal domínio é considerado assinatura de proteínas capazes de hidrolisar ligações arabinofuranosídicas a partir do terminal redutor em polissacarídeos contendo arabinose (Gilead & Shoram, 1995).

Por fim, o terceiro domínio consiste em um módulo de ligação a carboidrato (CBM, *carbohydrate binding module*), o CBM_4_9 *super family* (cl19911; e-value $2,73e^{-04}$; Fig. 35). O CBM4 é um módulo de aproximadamente 150 resíduos, cuja ligação já foi demonstrada a xilanos, β-1,3-glucanos, β-glucano de ligação mixta, β-1,6 glucano e celulose amorfa. Por outro lado, o CBM9 apresenta cerca de 170 resíduos e só foi descrito até o momento em xilanas tendo em um caso sido demonstrada a ligação à celulose (Lombard *et al.*, 2013).

RESULTADOS

Figura 35. Sequência nucleotídica (letras minúsculas) e de proteína predita (letras maiúsculas) a partir da sequência *ScARA1a*, contida no BAC SHCRBa_111_H15 de cana-de-açúcar, que codifica uma provável arabinofuranosidase. Trechos sombreados em verde, amarelo e vermelho indicam as regiões correspondentes ao provável do peptídeo sinal, módulo de ligação ao carboidrato e domínio α -L-AF_C, respectivamente. Aminoácidos sublinhados em vermelho correspondem ao domínio AbfA. Retângulos amarelos e verdes envolvem resíduos conservados em relação a enzimas do clã GH-A (Sakon *et al.*, 1996), sendo os verdes prováveis componentes da díade catalítica de acordo com a caracterização da estrutura tridimensional da arabinofuranosidase AbfA de *Geobacillus stearothermophilus* T-6 (Hövel *et al.*, 2003). Resíduos em retângulos azuis são conservados no clã GH-A, porém não em *ScARA1a* (Sakon *et al.*, 1996).

RESULTADOS

No que diz respeito a aspectos estruturais, a arabinofuranosidase caracterizada em melhor resolução de uma da família GH51 é a de *Geobacillus stearothermophilus*, AbfA, de 57 kDa (Hövel *et al.*, 2003). Esta consiste em um homohexâmero, sendo que cada monômero se organiza em dois domínios: um barril α/β e um sanduíche de folhas β estabilizada por pontes de hidrogênio, com topologia *jelly roll*. O primeiro destes domínios, característico da família GH51 (Lombard *et al.*, 2013), se estende dos resíduos 20 a 383 da proteína bacteriana, o que corresponde aos resíduos 52 a 521 de ScARA1a (Fig. 36).

Alguns resíduos de aminoácido são conservados em proteínas do clã GH-A, no qual está contida a família GH51 (Lombard *et al.*, 2013). Dentre eles, destacam-se os resíduos Glu29, Arg69, Asn74, Asn174, Glu175, His244, Tyr246, Glu294, Gln351 e Trp344 da arabinofuranosidase AbfA de *G. stearothermophilus* (Hövel *et al.*, 2003), a maioria de identidade também conservada em endoglucanases (Sakon *et al.*, 1996) e galactanases (Ryttersgaard *et al.*, 2002). A conservação da identidade de sete destes resíduos (Glu61, Arg256, Asn378, Glu379, His421, Glu457 e Trp461) foi observada na proteína predita a partir do gene ScARA1a (retângulos amarelos; Figs. 35 e 36). O resíduo Asn74 de AbfA encontra-se alinhado a Cys261 na arabinofuranosidase de cana-de-açúcar, enquanto que a região protéica na qual o resíduo Gln351 está presente na enzima bacteriana não existe na sequência de cana-de-açúcar e nem na arabinofuranosidase de *S. bicolor* (retângulos azuis; Fig. 36).

RESULTADOS

ScARA1a	EEINHAGAGGIWAEIVSNRGFEAGGLHTPSNIDPWSIIGDSSIYVETDRTSCFSRNIIA	120
Sbicolor	EEINHAGAGGIWAEIVSNRGFEAGGLHTPSNIDPWSIIGDSSIYVETDRTSCFSRNIIA	119
Txylanilyticus	EHLGRCIYEGLVVG-----	41
Gstearothermophilus	EHLGRAVYGGIYE-----	41
Tmaritima	EHLGRCIYGGIYE-----	38
* : . . . * : :		
ScARA1a	LRMEVLCDCNPTGGVGIYNSGFWMNIEDGKAYNLVMYVKSPETTDLTIVSLTSSDGLQNL	180
Sbicolor	LRMEVLCDCNPTGGVGIYNSGFWMNIEDGKAYNLVMYVKSPETTDLTIVSLTSSDGLQNL	179
Txylanilyticus	-----EDSP-----	45
Gstearothermophilus	-----PGHP-----	45
Tmaritima	-----EGSP-----	42
. *		
ScARA1a	ASATITVSGESNWTKEVQNLVAKGTNRTSRLQITTNKKGVVWFDQVSLMPEDTYKGHGFR	240
Sbicolor	ASATITVSGESNWTKEVQNLVAKGTNRTSRLQITTNKKGVVWFDQVSLMPEDTYKGHGFR	239
Txylanilyticus	-----IPNT-----NGIR	53
Gstearothermophilus	-----QADE-----NGFR	53
Tmaritima	-----LSDE-----RGFR	50
: : . . . * : *		
ScARA1a	TELISMILDLKPRFLRFPGCCFVEGEWLRNAFWRESIGPWEERP GHFGDVW-HYWTDDG	299
Sbicolor	TELISMILDLKPRFLRFPGCCFVEGEWLRNAFWRESIGPWEERP GHFGDVW-HYWTDDG	298
Txylanilyticus	NDVLEALKQMKIPVLRWPGCCFADE-----YHWKDGVPREKRKRNVNTHWGGVIENNH	107
Gstearothermophilus	QDVIELVKELQVPIIRYPGCFVSG-----YNWEDGVGPKEQRPRLDLAW-KSVETNE	107
Tmaritima	KDVLEAVKRIKVPNLRWPGCFVSN-----YHWEDGIGPKDQRPVRFDLAW-QQEETNR	103
: : . . . * : * : * : * : * : *		
ScARA1a	LGYEFLQLAEDLDAEPIWVFNNGVSHNDEVDTAIAPFVNDVLSIEFARGSANSTWGS	359
Sbicolor	LGYEFLQLAEDLDAEPIWVFNNGVSHNDEVDTAIAPFVKDVLDSIEFARGSANSTWGS	358
Txylanilyticus	FGTHEFMMLCELLGCEPYISGNVSGST-----VQEMSEWVEYITFDGESPMAN	155
Gstearothermophilus	IGLNEFMDWAKMVGAEVNMVNLGTRG-----IDAARNLVEYCNHPSGSSYSD	155
Tmaritima	FGTDEFIEYCREIGAEPYISINMGTT-----LDEALHWLEYCNGKNTYYAQ	151
* * : * : * : * : * : * : *		
ScARA1a	VRAAMGHPEFPVKYVAIGNEDCGK-----KFYRGNLYKFYNAIRQA-YPDIQMIS-	409
Sbicolor	VRAAMGHPEFPVKYVAIGNEDCGK-----KFYRGNLYKFYNAIRQA-YPDIQMIS-	408
Txylanilyticus	WRENGREKPKWRIKWGVNENWGC---GGNMRAEYADLYRQFTYLRNYGDNKLHKIAC	213
Gstearothermophilus	LRIAHGYKPHKIKTWCLGNEMDGPWQIGHKTAVEYGRICEAAKVMKWV-DPTIELVVC	215
Tmaritima	LRRKYGHPEPYNVKFWGIGNEMYGEWQVGHMTADEYARAAKEYTKWMKV-FDPTIKAIIV	211
* * : * : * : * : * : * : *		
ScARA1a	NCDG-----SSTQLDHPADLYDF-----HYY-----TDSKT	435
Sbicolor	NCDG-----SSTPLDHPADLYDF-----HYY-----TDSKT	434
Txylanilyticus	GANTAD---YHWTEVLMKQAAPFMHGLSLHYTVPGPWEKKGPATGFTTDEWVTLKKA	269
Gstearothermophilus	GSSNRNMPTFAEWATVLDHTYDHDYISLHCY-----YGNRNDNTANYLALS	263
Tmaritima	GCDD-----PIWNLRLVLEAGDVIDFISYHFY-----TG--SDDYYETVSTV	251
. . . : . . . * : *		
ScARA1a	LF-----NMKSTFDRTSR--GPKAFVSEYAVWRGDAG-----R	467
Sbicolor	LF-----NMKSTFDRTSR--GPKAFVSEYAVWRGDAG-----R	466
Txylanilyticus	LFMDRLVTKHSAIMDVYDPDKR---IDLIVDEWGTWYDV-----EP-GTNPGFLYQQ	318
Gstearothermophilus	LEMDDFIRSVVAIADYVKAKRKSKTIHLSFDEWNVWYHSNEADKLEPWTAPPLLEDI	323
Tmaritima	YLLKERLIGVKKLIDMVDTARKR--GVKIALDENWVYRVSD-NKLEEP-----	298
: . . . * : * : *		
ScARA1a	GSLLASLAEEAFLTGLEKNSDIVMASY-----APLFVNNNDQTNPD-AIVFNSWQ	518
Sbicolor	GSLLASLAEEAFLTGLEKNSDIVMASY-----APLFVNNNDQTNPD-AIVFNSWQ	517
Txylanilyticus	NSIRDALVAGATLHIFHRHCDVRMANIAQLVNVLQSVILTEGERMLLTPT-YHVFNMFK	377
Gstearothermophilus	YNFEDALLVGMLITLMKHADRVKIACLAQLNVI-APIMTEKNGPAWKQTIYYPFMHAS	382
Tmaritima	YDLKDGIFACGVLVLLQKMSDIVPLANLAQLVNAL-GAIHTEKDGLILTPV-YKAFELIV	356
: : . . . * : . . . * : *		

Figura 36. Alinhamento múltiplo do core catalítico de arabinofuranosidases de bactérias e de gramíneas. ScARA1a, proteína predita a partir do gene homônimo; Sbicolor, proteína predita a partir do gene Sobic.005G022500.1 de *S. bicolor*; Gstearothermophilus, AAD45520.2 de *Geobacillus stearothermophilus*; Txylanilyticus, CAA76421.2 de *Thermobacillus xylanilyticus*; Tmaritima, AAD35369.1 de *Thermotoga maritima* MSB8. Retângulos amarelos e verdes denotam aqueles conservados entre a sequência de cana-de-açúcar e a de *G. stearothermophilus*, sendo que aqueles em verde indicam a díade catalítica da enzima de *G. stearothermophilus*. Resíduos em retângulos azuis indicam aqueles não conservados em cana-de-açúcar e com função relevante na enzima de *G. stearothermophilus* (Hövel *et al.*, 2003). À direita do alinhamento são mostradas as posições do último resíduo de aminoácido de cada proteína. Abaixo da sequência, resíduos idênticos (*), conservados (:) e semiconservados (.) são marcados conforme gerado pela ferramenta ClustalW.

PARTE F: Avaliação da interação dos fatores de transcrição RAV e ERF com os promotores das glicosil hidrolases

1. Mapeamento de elementos em *cis* nos promotores de *ScEPG1a* e *ScARA1a*

O mapeamento de 2 kb à montante do ATG de *ScEPG1a* e *ScARA1a* mostrou a presença de vários prováveis sítios de ligação a fatores de transcrição tais como para ERFs (GCCGCC, GCCCORE; Chakravarthy *et al.*, 2003), RAVs (CAACA, RAV1AAT; Kagaya *et al.*, 1999) e MYBs (MYB1AT; CANNTG e WAACCA; Abe *et al* 2003). Além destes, também foram identificados nos promotores alguns elementos de resposta à giberelina (*giberelic acid response elements*, GAREs, TAACAAR; Ogawa *et al.*, 2003) e elementos de resposta à auxina (*auxin response elements*, AuxREs, KGTCCCAT; Ballas *et al.*, 1993). Na região analisada à montante do ATG de *ScARA1a*, cinco elementos RAV1AAT e três GCCCORE foram identificados, além de 3 MYB1AT (Fig. 37). Por outro lado, no promotor de *ScEPG1a* oito elementos RAV1AAT, 2 GAREs, 2 AuxREs e 5 MYB1AT foram observados (Fig. 38).

-210 CTATCGGCGGTGCAGAGCTCGCTAA **CAACA** GCCTCGCTCGCCGGCAGAGTGTGCGTGAGCTCAATAATGACGCGTAGATGTC **GGCGGC** GCAGAGCTC

-1912 AGCGGGGCGATTGAGGCTGCAAGAGGTGCGCGAGCTGCGGCATGGGGTGAGGAAGGCGATGGGTGGGCGGTTTGGACGGAGGTTGCACGGAGAA

-1814 GGGCGCTCAGAGGCTGGAAGTAGCG **ACAACCG** TAGAAGCTGGGCG **GGCGGC** CATGGCGCTCGATGCCTATGCTCTCGACAGAGCAAGACGAAGAG

-1716 GAGCGAGGGCAAGCAATGGGATAGGGAGAGAGGGCGCAAGCTCGGTCTCCACACTTGCGCGAACCCGAGGCGCAGAGTGGGGGCACGAAAGCAGCAT

-1614 GGCGACCATGCGTTGGGCATATTCTGTCCGACATCTTCTCTCCCTGTTTGCTTGAATGGAGTGACGGAGACAGTGAGAGACGAGATGGAATCAGAAAG

-1512 GATATTGATGTGGTGGGTAAATACGTGGACACAATAAGACCACGTAGCTGGATGCGGGACAGATGCCCGCAGTGGAGCATTATCGTACAAACAAGCATCT

-1404 ATGTTATACTATTGTGAGGCGTCACATTCTCTATAGTTTGATCTAATATGTATATGATTTTTTGCTTATCCGGTTTGGTAGCATATCTGTATACCATGTTTCG

-1300 ATCAGAAACACTTCTCCACGACGACGATTGCTTCGATGATGATTAGACTATCTC **CAACAACAA** TATCCAAATATAAGACACATTTGACGTTGGGTAGCAT

-1197 TACAGACAGAGGGTTCAATACCCATTTCGACGTCTTCTCTAA **CAACA** TGACCCAAAAGATAATCCTTTCTGTAATGAGTTTTTCAGGAGAGATGATCCCATATT

-1092 TGGGTTGTGCAAGACAACCTAAATCGGTCTCTCATATAAATACTA **TGTTG** GAGGCCGATCTTAGATCTCCTGCTATCTATTTTGAGTTAGGTGCCACGTG

-988 GCCACCTTTTGGAGAGTCTTATAATGTAATAATAGAAGTAGTATCATTAACGCTTGTGAACAAATATAATACACTCCGTCAACACGATGGGTACACGCGG

-893 GGCGGGGTGGGCCCCACGTGTCCGTGGCGCGGGCGGGCGCAGAGACAGCCACC **GCCGCC** AGAAATGGCCCGCTCCGGCAAGCA **ACTGCGGCC** GC

-793 GCCGCGCGCGACCTGGGCGCTGGCCGTCCCGCTCGCTCGCGCGCGTGACCCGATCAGAATCCGACGCGCGCCCCGCACACAGGCTTGCGTGGCTTC

-692 CGTAGCCCGCAGCCGACGGCACACTTCCCTGCCTTCAAAGCCACTGCTAACCGCTCTCTGCTTCTCTACTTAAGCGCGAGCGAGCAGCGCTCCCGGG

-589 CTCTCTCTATGCCCTGCCCTGCATCGCAGCCGGGAGCTCGCATCTTTGGATGTTGGATGCATGGTGAGCAACCTCTGAAGACCCGCTTTCTGTTTC

-484 CTTCGTCTTTGTATCGGATTTCGGGCTCATGACTTGTTTGATGGAATGCCAGCTCAGGAGAATTCGCTAAATGCTTAATCTTTCGCTAGTTGTGCTTTGT

-382 CTGGTGAATAATTACACGTTTTAAATAAAGGCAGCTCTTCTTTTCTATTCCAAGTGCCTGCACAGCTCTTGCATGACGAAA **CAACA** AGCACCATTTTAA

-276 ATTTCACTGCGACTGATGATGAGTACTGGTGAGCTTATAAAGCTATAATGTCCAGCATCTATAGTTGTTATTGCATATGTGCGTAATATGAGTGGTGTGC

-170 GGTTTTCTAATCTGATGAACGCGTCAATACTCAGTTGCCATCTAGCTTGTGCTTATCAGTTTCCACCTAACTCTAATGTCTTATGATTATATCGCCTTCTCTC

-63 TTATGTGGTCGTAGACACACCGGCTCAACTAGAGTTGACAAGTGGAAGTGTGGAAGCGTATG

Figura 37. Mapeamento de possíveis sítios para ligação de fatores de transcrição das famílias MYB (vermelho), RAV (verde) e ERF (roxo) no promotor de *ScARA1a*. O sítio de início da tradução está em negrito e sublinhado.

Em razão da relevância da sinalização hormonal na formação do aerênquima (Fig. 16C) foi dado foco à avaliação do papel dos fatores de resposta a etileno identificados (ScERF1a e ScRAV1a) na regulação da degradação de polissacarídeos de parede celular. Em um sistema heterólogo (*N. benthamiana*) foram realizadas as transformações transientes para avaliar as interações entre os fatores de transcrição sequenciados e os promotores dos genes referentes às duas glicosil hidrolases expressas em raiz de cana-de-açúcar que foram escolhidas para aprofundamento.

RESULTADOS

-2187 GTTAATGAGTGATGCACACACTATTTAAAGGTGTATGTACTGATACTCAAAAAGTATGTCCACGTACTCGTGCCGTATGTGTACGCACTAACAATA**TAAAGT**
-2084 ATGTGTGCTTGTCTTCTGAAGCAGTATTTGTTCACTATTTTATTA**TAAAGT**ATTGGTCCATACTATTA**TGTTG**GACTAGTATTTGTTCACTATTTT**TAA**
-1981 **CAAA**GACCATGTGTGCGATACTTCGTACAAACATGCTAATGAGCGTT**GCTAACCA**AACTTAAACAAGTATGAATACATACCAATCCCCGCAAGAAGTATC
-1877 TATCCAT**ACTTTA**AAAAAAGTACACCTGTTCCGCGAGTATGTTCTCGCGCTTATAAACCGGTTGTGGCTGTTTATTATGAGAGAAAAAATCTGTACCAC
-1778 GGCTAAGCCG**TGTTG**ATAAGAGCAGCCGAACAGGGTGCAGGTACATACACAAACGTATACTCCCTCCGTCTAGGATATAAGGCG**TAAACCA**CTTTTATT
-1676 CATGTCCCATGATATAAGGCGTCTCTCACAGACATTAATGCAAGTAGTCTAGTGTGAGAGAGAGAATTAATGTGTTCTTGGTCTTTGAACCAGAGGT
-1576 GGTACGTCTTATATCTAGGACGGAGGGAGTAAATATATAAAAAAGGTAAGTGGCCATGCAAGCCCTAGGACGGAGGGAGTAAATATAAAAA
-1474 AAGGTAAGTGGCCATGCAAGCCCTGAGCAACGAATCCCGTACTAGCGCCTCATTTTAAATTCCTGTAGATACACGCACGCTCCACATTGTATGCAAA
-1373 **CCGGTTA**GATCCATCACCGGC**CCGTTA**ACTTAGTTTCGCGCGGAATGGCTCCCTCGCGTGCATATACACGCAGATTATGACATCAAAATAATATATT
-1271 GAAAATATAATTAATAAGAACCTAATGATACCACTGTTCTGCTAAATGGCCGTTTAGTTCTGTTAAACACCGTTAAACGGTAACAGAAAGTCAGCCG**TGT**
-1169 **TG**ATTAGGCCCTAAACCGTGAATTAAC**TGTTG**GGCCGTTTAAACCGTTTAAACCGTTCAATCTCGTCTAAACGGTCTAAACAAATCAGATTAAACGTC
-1067 ATTAATGAGCTAAACAACCATTTAATCCATAATTAAGTGAATTAAGGACTACGTCTATCACAAT**CAACA**TGAAACTAATTAACATTGAATTATATGAAG**T**
-966 **GTTG**TGGGATTAGATTCTT**CAACA**AGATTATTTGGTAGAATTTGATTTTACATGTGTAAATATATATGGTTTCTAGCAAATTTATTTTACCTACATA
-863 AATATGTATATTTATTATTATTGTACAAATCGTTTAGGCCGTTTAACTCTGTTTAAACATCATTTAAATGGTCTAAACGCTAAACGAAGGATGACCGACTGTT
-758 TACCGTTTACCGTTTAGGAAACATTGAATGATACATAGTTGGTATCATAAATCGTATTATCTTATCATATAAATTTGGTCAAACTTGAGATGGTTGACTCTCCA
-653 AGATTCTTGAATGCTTATAATTAGGATGGAGGGAGTATATGAGAGAAGTATTTGGACTTTTAAATTTAAT**TAACAAA**ATACTCTTGCTCACTCACA
-547 ACGAATCCCGATCGGTTCTCACACCTCCACCCCATACACGGCAACTCTCAAATATTTTCACTCGACTGCGCAGGCAAAAGAGGCAAGATATCGTCA
-444 GTTCATCACACTGAAGGCAACCAAAATCCCAAAACATCTCACAGGATGGCTTGCACTCCTCACCTTCTCACGCGCGACTCTCCAGTTCCCTCCGCTCCCT
-341 CTGTCACTTCCCTTGCAAGGCAACTCATCACGAGGCCGGGATCTCCGCGACATCAATCCGTGCGGTATGGACCTCCGTG**CAACA**GGTCCATAGGCCTGT
-241 GAGCCAGCTGCCGTCGCGGTCCACGGTGAGCTCGCGCGTGGACCGGGCC**ATG**CGCGCTGCGGTTTCATCAGCAGTTCAGCACTTGCCACGTTTGTG
-141 GCCGTGGGATTATGGCAGCAGGGGCACATCCGTAATCTCTCCGCGGTCTCGGCCTCTCTCGGCTCGCTATCCACCGCTTCCGCAATGGCCACCTAAT
-36 TCGCTCGGCGTCGGGGGAGAAGAGAAGCGCGCGCG**ATG**

Figura 38. Mapeamento de possíveis elementos de resposta a giberelina (laranja) e possíveis sítios para ligação de fatores de transcrição das famílias MYB (vermelho), RAV (verde) ARF (azul) e ERF (roxo) no promotor de *ScEPG1a*. A região clonada para os ensaios de transativação correspondeu aquela à montante do ATG em negrito na posição -183 até -2187. Posição +1 correspondente ao ATG em negrito e sublinhado.

2. Construção dos vetores para transformação de *A. tumefaciens*

Para avaliar a interação entre os fatores de transcrição *ScERF1a* e *ScRAV1a* com os promotores de *ScEPG1a* e *ScARA1a* (EPGpm e ARApM, respectivamente), as quatro sequências foram amplificadas empregando os respectivos BACs como *template*. A amplificação dos genes *ScRAV1a* e *ScERF1a* geraram *amplicons* de 1,2 kb cada (Fig. 39A), os quais foram ligados ao vetor pENTR™/D-TOPO gerando as construções intermediárias para transformação de *E. coli*. Em seguida, os insertos foram transferidos para os vetores finais gerando as construções RAVpFAST e ERFpFAST (Anexos 11 e 12). Por sua vez, os *amplicons* de 2 kb cada produzidos a partir da região promotora de *ScEPG1a* e *ScARA1a* (Fig. 39B) geraram os vetores finais EPGpm-pKGWFS7 e ARApM-pKGWFS7 (Anexos 13 e 14).

Figura 39. Produção dos *amplicons* para construção dos vetores intermediários e finais dos ensaios de transativação. (A) Amplificação de ScRAV1a (seta azul) e ScERF1 (seta verde) e (B) sequências promotoras de ScARA1 (retângulo vinho) e ScEPG1 (retângulo laranja) em géis de agarose 0,8% corados com brometo de etídeo. O BAC usado como *template* em cada amplificação é indicado para cada *amplicon* (círculos pretos) produzido por iniciadores específicos (> e <). Setas pretas, verdes e azuis representam sequências genômicas, retângulos brancos mostram sequências intrônicas e seta preta em L indica o ponto de início da tradução. Setas amarelas apontam as bandas do marcador na faixa de tamanho dos *amplicons*. Marcadores moleculares: (A) 1 kb DNA ladder (Promega); (B) FastRuler™ Middle Range DNA Ladder (Fermentas).

3. Regulação transcricional mediada por ScRAV1a e ScERF1a em folhas de *N. benthamiana*

Após a transfecção de combinações de construções efetoras e repórteres em folhas de *N. benthamiana*, a atividade de do reporter (GUS) foi avaliada. Visto que o extrato protéico produzido a partir de folhas transformadas com o vetor vazio (pFASTØ) exibiram níveis mensuráveis de atividade GUS após a co-transfecção com ambas construções repórteres, foi possível comparar o nível de atividade controle (vetor vazio) em relação às folhas transformadas com as construções efetoras (RAVpFAST e ERFpFAST).

No que diz respeito à co-transfecção empregando ARAp_m, nenhuma das construções efetoras resultou em valores de atividade GUS com diferenças estatisticamente distintas da obtida para o controle com o vetor vazio (Fig. 40). O mesmo foi observado após a medida de atividade GUS em extratos protéicos produzidos a partir de plantas co-transfectadas com a construção efetora ERFpFAST (Fig. 41A, barra verde) e a repórter EPG_m- pKGWFS7 em relação à infiltração empregando a construção efetora pFASTØ. Os níveis da atividade GUS foram similares na transformação com o vetor vazio e com ERFpFAST (Fig. 41B, barra verde). Por outro lado, a co-transfecção com a construção RAVpFAST (Fig. 41A, barra azul) resultou em uma redução média e estatisticamente significativa de 50% na ativação do promotor de ScEPG1a (Fig. 41B, barra azul). O mesmo padrão foi reproduzido em três experimentos independentes.

RESULTADOS

Figura 40. Efeitos de ScRAV1a (RAVpFAST; azul) e ScERF1a (ERFpFAST; verde) sobre o promotor de ScARA1a (ARApM-GUS; vinho) e ScEPG1a (EPGpm-GUS; laranja) em folhas de *N. benthamiana*. (A e C) mapas esquemáticos de construções repórter e efetoras. CaMV35S, promotor 35S; GUS, sequência codificante da β -glucuronidase; +1, códon de início da tradução; RAV1AAT, elementos para ligação de RAV; GCCCCORE, elementos para ligação de ERF; ter, terminador da nopalina sintase; R-AP2, domínio AP2 de ligação a DNA de ScRAV1a; R-B3, domínio B3 de ligação a DNA de ScRAV1a; E-AP2, domínio AP2 de ligação a DNA de ScERF1a. (B e D) Medidas de atividade de GUS a partir dos extratos protéicos produzidos após co-transfecção em folhas de *N. benthamiana*. Os dados representam as médias e desvio-padrão de três experimentos independentes, cada um contendo três repetições biológicas. A atividade de GUS é expressa em termos de razão da inclinação da curva dos valores de atividade GUS relativos às plantas controle (pFAST, vetor vazio de folhas co-transfectadas com construção repórter ARApM-GUS em B e EPGpm-GUS em D). Para o cálculo da inclinação, apenas a fase exponencial da cinética de atividade foi levada em conta. Asterisco indica diferença estatisticamente significativa em relação à EPGpm Ø de acordo com teste T de Student ($p < 0,05$).

1. A formação constitutiva do aerênquima em cana-de-açúcar

O padrão observado de formação do aerênquima em cana-de-açúcar na ausência de nutrientes e 1-MCP (tratamento NPK na Figs 9-10 e tratamento controle na Fig. 15) é consistente com o gradiente observado em estudos prévios empregando o mesmo modelo (Leite, 2013; Grandis, 2015). No entanto, o desenvolvimento do aerênquima diante da aplicação de 1-MCP ou na ausência de nutrientes não havia sido avaliado nesta planta, embora a formação constitutiva do mesmo já tivesse sido reportada mediante alagamento (Karim & Tetsushi, 2007; Begum *et al.*, 2013). Visando melhor caracterizar a resposta do aerênquima constitutivo as plantas foram cultivadas em estresse nutricional e os resultados obtidos corroboraram em parte aqueles existentes para *Z. mays* (He *et al.*, 1992) e *Phaseolus vulgaris* (Fan *et al.*, 2003) cultivados em condições similares no que diz respeito à indisponibilidade de nutrientes. Especialmente em *Z. mays*, indivíduos cultivados na ausência de fontes de nitrogênio apresentaram níveis de formação do aerênquima equiparáveis com os de plantas em hipóxia (He *et al.*, 1994), indicando que ambas condições culminam na formação de aerênquima.

Em cana-de-açúcar, a tendência de formação do aerênquima aparentou ser tão maior quanto menor a disponibilidade de fontes de nitrogênio ou fósforo testadas no experimento (Fig. 9). Entretanto, ao contrário do que se encontra descrito na literatura para *Z. mays* (He *et al.*, 1992) e *P. vulgaris* (Fan *et al.*, 2003), em cana-de-açúcar a presença de nutrientes (Figs. 9 e 10) não resultou em inibição total do desenvolvimento do aerênquima. Alternativamente, este foi claramente retardado. Especificamente no quarto segmento de raiz, a área ocupada pelo aerênquima no córtex radicular de plantas em déficit nutricional foi cerca de sete vezes maior do que o observado para plantas semanalmente irrigadas com solução nutritiva (Fig. 10). Por outro lado, o emprego de 1-MCP visando bloquear a percepção de etileno não alterou a área dos espaços de gás nas seções transversais do córtex das raízes de cana-de-açúcar (Fig. 15). Em plantas com formação de aerênquima induzida como *Z. mays* e *P. vulgaris* o aerênquima não se desenvolveu (Rajhi *et al.*, 2011) ou o fez ocupando menor área no córtex (Fan *et al.*, 2003) diante da aplicação deste inibidor em concentrações iguais ou superiores à usada em cana-de-açúcar. Em adição a dados anteriores (Karim & Tetsushi, 2007; Begum *et al.* 2013; Leite, 2013), ambos os tratamentos executados em cana-de-açúcar forneceram um conjunto maior de evidências que melhor sustentam a afirmação que em cana-de-açúcar a formação do aerênquima é um processo constitutivo.

Embora a ocorrência de formação constitutiva do aerênquima seja comum em plantas aquáticas ou de ambiente alagado (Evans, 2003), o relato do mesmo em plantas de terra firme não é inédito na literatura científica. Em raízes de plantas como *H. vulgare* (Garthwaite *et al.* 2003), *S. bicolor* (Promkhambut *et al.*, 2011) e *Musa* spp (Aguilar *et al.* 1998) a formação de aerênquima foi observada em condições experimentais que não envolviam alagamento ou hipóxia. Entretanto, embora para a formação do aerênquima constitutivo não seja requerido um indutor externo, o mesmo pode ser intensificado quando as plantas são cultivadas em condições que favorecem este processo. Em *Musa* spp. o aerênquima constitutivo se inicia a 100 mm da ponta da raiz em plantas alagadas, enquanto em plantas controle o mesmo ocorre a 300 mm do ápice radicular (Aguilar *et al.* 1998). Resultados similares foram obtidos para

variedades de *O. sativa* com aerênquima constitutivo. A formação de aerênquima na presença de ACC se inicia a 15 mm mais próximo da ponta da raiz quando comparado a plantas do grupo controle (Yukiyoshi & Karahara, 2014). Por outro lado, em cana-de-açúcar os resultados obtidos neste e em trabalhos anteriores demonstraram que a formação de aerênquima sempre se inicia a 3 cm do ápice radicular independentemente das condições experimentais (Leite, 2013; Grandis, 2015). Nesta planta o estímulo ao desenvolvimento do aerênquima devido ao déficit nutricional não levou à detecção dos espaços de gás em segmentos mais apicais que S3. Ao invés disso, segmentos iguais de tratamentos distintos apresentaram valores diferentes de área do aerênquima no córtex (Fig. 10). Possivelmente em razão dos maiores trechos de raiz analisados no presente trabalho (da magnitude de centímetros, em oposição aos trabalhos citados que analisam milímetros de raiz) o desenho experimental aqui empregado não apresentou resolução suficiente para analisar sutis diferenças em relação à precisa posição em que o aerênquima se iniciou no terceiro segmento dos tratamentos aplicados.

Na literatura é sugerido que o déficit de fontes de N e P leva ao aumento da sensibilidade ao etileno e isso pode se dar por meio da regulação positiva da cascata de sinalização deste hormônio (He *et al.*, 1992; 1994). Esta hipótese foi corroborada em estudos posteriores (Ma *et al.*, 2003) e estendida para outros nutrientes, como fontes de enxofre e potássio (Bouranis *et al.*, 2003; Bouranis *et al.*, 2006; Siyiannis *et al.*, 2011). Esta é considerada uma adaptação das plantas visando reduzir os custos metabólicos para crescimento e exploração do solo (Fan *et al.*, 2003; Postma & Lynch, 2011). Provavelmente o efeito do encarceramento de etileno nos espaços de gás abertos pelo aerênquima têm um importante papel nesta resposta (Drew *et al.*, 1979), visto que a biossíntese deste hormônio não é estimulada (He *et al.*, 1992; He *et al.*, 1994; Drew *et al.*, 2000). Entretanto, a ausência de regulação positiva da biossíntese de etileno em resposta a estresse nutricional é um dado controverso e as vias específicas que ligam o monitoramento de nutrientes à sinalização deste hormônio ainda são desconhecidas (Lynch & Brown, 2000; Iqbal *et al.*, 2013).

No que diz respeito especificamente ao 1-MCP, só há um relato na literatura reportando a formação do aerênquima na presença do inibidor com padrões equiparáveis ao controle, assim como observado para cana-de-açúcar (Fig. 15). Na planta aquática *Juncus effusus* a aplicação repetitiva de 1-MCP não alterou a formação do aerênquima e resultados similares foram obtidos quando da aplicação de etileno exógeno. Diante disso, os autores concluíram que *J. effusus* não requer a sinalização por etileno para deflagrar a morte celular programada e a formação do aerênquima (Visser & Bögeman, 2006).

Conforme discutido acima, sabe-se que a sinalização por etileno é indispensável para as respostas a déficit de nutrientes específicos (Kim *et al.*, 2008). Além disso, a importância do etileno como mediador de respostas adaptativas a carência de nutrientes (sobretudo fontes de P) é sustentada pela observação que o déficit nutricional pode induzir uma variedade de mudanças na arquitetura radicular que são também provocadas por etileno. Dentre estas alterações ressaltamos o gravitropismo, formação de aerênquima e de pêlos radiculares (Lynch & Brown, 1997; Nagarajan & Smith, 2012).

A conclusão assumida para *J. effusus* no que diz respeito à formação do aerênquima independentemente do etileno não parece plausível para cana-de-açúcar visto que para esta planta o

estresse nutricional aumentou a área ocupada pelo aerênquima. Por outro lado, a questão em torno da sensibilidade ao etileno pode ser uma explicação possível para a formação do aerênquima na presença de 1-MCP. Notavelmente, foi observado que os níveis de etileno requeridos para que o aerênquima se forme em cana-de-açúcar (Fig. 16B) são cerca de 300 vezes menores que os reportados para *Z. mays* (He *et al.*, 1996a). Embora tais diferenças possam em parte refletir diferenças metodológicas ou oriundas dos distintos estágios de desenvolvimento no qual se encontravam ambas plantas, considera-se um valor bastante expressivo. Como base de comparação, em um fruto de *Cucumis melo* cultivar Charentais a concentração deste hormônio sofre acréscimos de cerca de 20 vezes durante o amadurecimento (Pech *et al.*, 2008).

Em experimentos com folhas de *Dianthus caryophyllus* em senescência foi demonstrado que, dependendo da sensibilidade do tecido a etileno, este pode requerer a aplicação simultânea de 1-MCP e AVG para bloquear o efeito deste hormônio (In *et al.*, 2013). Sendo assim é possível sugerir que as células do tecido formador de aerênquima sejam mais sensíveis ao etileno. Tal fato explicaria inclusive porque o prolongado estresse nutricional não provocou mudanças mais dramáticas na formação do aerênquima. A natureza desta suposta maior sensibilidade ao etileno pode residir em uma menor quantidade de receptores para este hormônio. Sugestões similares já haviam sido propostas explicar o aumento da sensibilidade ao etileno diante de estresse nutricional em *Z. mays* (He *et al.*, 1994).

No que concerne ainda à regulação hormonal por trás da formação do aerênquima, pode-se sugerir que não só a sensibilidade do tecido ao etileno seja relevante neste processo, como também a presença de um ou mais hormônios agindo em conjunto e não meramente a presença ou ausência de um dado hormônio único. Em plantas do tratamento controle, os mais elevados níveis de etileno sobretudo em S1 (Fig. 16B) possivelmente atuam estimulando a conjugação da auxina a aminoácidos, conforme observado em *S. lycopersicum* (Liu *et al.*, 2005) e *Capsicum chinense* L. (Böttcher *et al.*, 2010). Esta seria uma explicação plausível para os reduzidos níveis de auxina livre nestas plantas em relação às tratadas com 1-MCP neste segmento (Fig. 16A). Já no grupo ao qual foi aplicado o inibidor, reduzidas concentrações de etileno foram observadas ao longo da raiz. Dentre os efeitos já reportados do 1-MCP estão a diminuição da produção de etileno, reduzindo o acúmulo de transcritos e/ou a atividade enzimática de ACC e ACO a depender da planta empregada no tratamento e do momento em que o inibidor é aplicado (Blackenship & Dole, 2003). Dessa forma, as baixas concentrações deste hormônio em S1 favoreceram a manutenção de AIA na forma livre na região mais próxima do ápice radicular (Fig. 16A). Como consequência, ainda que o 1-MCP não tenha alterado a área do aerênquima (Fig. 15), ele afetou o balanço auxina:etileno nos dois segmentos mais apicais (Fig. 16C). Nestes, a porcentagem de auxina em relação a etileno foi de 30% em plantas controle, enquanto nas tratadas com 1-MCP foi de 60%. Isso foi resultado de altos níveis de auxina e baixos níveis de etileno em S1 e S2 em plantas tratadas com 1-MCP.

Nas plantas controle também é possível observar a queda nos níveis de ambos hormônios se comparados os segmentos 1 e 2 em relação aos demais (Fig. 16). Sabe-se que níveis reduzidos de etileno são requeridos para estimular a síntese de auxina na região onde se formam as raízes laterais (Fukaki & Tasaka, 2009; Jung & McCouch, 2013). Visto que em cana-de-açúcar algumas raízes laterais

são observadas do quinto segmento de raiz em diante (dados não mostrados), é esperado que a concentração de etileno e auxina apresente quedas de S3 a S5, voltando a acumular somente em regiões de formação destas raízes (Muday *et al.*, 2012). Interessantemente, em cana-de-açúcar o desenvolvimento do aerênquima se inicia em S3, exatamente no ponto em que a proporção entre auxina e etileno começa a se aproximar de 50%, independente do tratamento (Fig. 16C). Dessa forma, é possível que não só o etileno mas sim o balanço e entre ambos hormônios seja relevante para formação do aerênquima constitutivo em cana-de-açúcar.

A auxina é um elemento chave em processos de degradação de parede celular que são parte do desenvolvimento normal da planta, usualmente como resultado de um complexo *cross talk* com outros hormônios. A interação entre estes hormônios também é estudada no amadurecimento de frutos climatéricos como *S. lycopersicum* e *Prunus persica*. Nestes frutos, processos metabólicos dependentes de AIA se desenrolam no mesocarpo como consequência do padrão de produção de etileno. Os níveis deste hormônio são coincidentes com o amadurecimento, influenciando a expressão de PIN1 (Trainotti *et al.*, 2007). Além disso, também é sabido que auxina e etileno regulam conjuntos diferentes de genes em células em expansão e frutos em amolecimento durante o amadurecimento de frutos de *S. lycopersium* (revisado por Rose & Bennett, 1999). Processos como a formação do aerênquima, amadurecimento de frutos e outros como abscisão e mobilização de reservas seguem passos comuns no que diz respeito à sinalização e degradação de parede (Grandis *et al.*, 2014, Tavares *et al.*, 2015). Isso reforça a sugestão que, assim como já foi descrito para estes processos, os mecanismos por trás da formação constitutiva do aerênquima podem ser regulados por uma complexa interação entre auxina e etileno, ao invés de etileno unicamente.

O *cross talk* e a ação combinatória entre ambos hormônios é um tema bastante explorado no que diz respeito à morfologia radicular. Sendo o aerênquima constitutivo um processo considerado parte normal do desenvolvimento (Evans, 2003), os mecanismos que levam ao desenvolvimento do aerênquima em cana-de-açúcar podem ser indissociáveis daqueles relacionados à formação da raiz. A atuação do etileno neste órgão pode se dar como mediador do transporte acrópeto da auxina para a zona de alongamento (Ruzicka *et al.*, 2007). Por sua vez, sabe-se que a auxina tem o papel de estimular a síntese de etileno (Mittler *et al.*, 2011), corroborando a expressão aumentada de ACO que foi observada na cana-de-açúcar (Tabela 6). Estas informações, em conjunto com outras exploradas adiante no que diz respeito ao efeito de 1-MCP em raízes de cana-de-açúcar, encontram-se no modelo da Figura 41.

No caso da abscisão foliar, níveis crescentes de etileno reduzem o transporte de auxina e as células da camada de abscisão apresentam maior sensibilidade ao hormônio gasoso. Isso leva à produção de enzimas de degradação de parede celular, sobretudo aquelas destinadas à degradação da lamela média, como é o caso de endopoligalacturonases. Estas modificações consistem em etapas requeridas para formação da camada de abscisão (Beyer & Morgan, 1971; Beyer, 1973). Conforme mencionado, baixos níveis de etileno podem estimular também a síntese de auxina, promovendo a emergência da raiz lateral no ponto de máximo acúmulo de auxina (revisado por Jung & McCouch, 2013). Este estímulo se dá

por meio da regulação positiva de genes WEI 2 e 7 (WEAK ETHYLENE INSENSITIVE) (Fig. 41), que codificam para subunidades de uma enzima que catalisa a etapa limitante da biossíntese de triptofano, que leva à maior produção de auxina (Stepanova *et al.*, 2005).

Figura 41. A aplicação de 1-MCP em cana-de-açúcar resultou na queda dos níveis de etileno, conforme já observado para outras plantas (revisado por Blackenship & Dole, 2003), embora as vias envolvidas nesta regulação sejam ainda desconhecidas. Níveis mais baixos de etileno podem tanto ter estimulado a síntese de auxina via genes *WEI2/7* (Jung & McCouch, 2013) como desfavorecido a conjugação deste hormônio (Liu *et al.*, 2005; Böttcher *et al.*, 2010). Em razão disso nos segmentos de raiz S2-S5 a concentração de auxina livre é reduzida. O acúmulo de auxina em S1 pode ter sido responsável pelo aumento da expressão de genes que codificam para RBOH (Müller *et al.*, 2012), Hb (Pagnussat *et al.*, 2003), expansinas (Rose & Bennett, 1999), arabinofuranosidase (Hutchinson *et al.*, 1999) e H⁺-PPases (Fukuda & Tanaka, 2006). Essas últimas tem o papel de aumentar a expressão de genes referentes a H⁺-ATPase e PIN1, estimulando a expansão celular (Li *et al.*, 2005). Já o gene que codifica para uma hemoglobina pode regular negativamente a via de biossíntese de etileno, por meios indiretos (Igamberdiev *et al.*, 2005). Setas pretas e linhas com barras horizontais indicam respectivamente regulação positiva e negativa observada neste trabalho enquanto setas brancas e linhas pontilhadas indicam sugestões baseadas em dados da literatura. Linhas pretas indicam possíveis efeitos dos altos níveis de auxina em S1. As dimensões de formas geométricas em vermelho em azul ilustram as concentrações de auxina e etileno obtidas neste trabalho, respectivamente.

No que diz respeito à formação de aerênquima, sabe-se que este não ocorre em regiões onde há presença de raízes laterais (Gunawardena *et al.*, 2001a) nem em sistemas constitutivos (Leite, 2013). Pode-se então assumir que baixas concentrações de auxina (mais baixas do que as requeridas para

emergência da raiz lateral) são mais favoráveis à formação do aerênquima. O único trabalho reportando a influência da auxina na formação do aerênquima em *Z. mays* demonstra que a aplicação de baixos níveis de ácido 1-naftilacético (NAA, hormônio sintético da família das auxinas) levaram ao aumento da formação do aerênquima, enquanto que altas doses apresentaram efeito oposto. Segundo os autores esta resposta provavelmente se deve ao efeito indutor da auxina sob a biossíntese de etileno (Justin & Armstrong, 1991a). Todavia, o efeito deste hormônio na formação do aerênquima é pouco descrito e sua interação com o etileno é bastante negligenciada na literatura no que concerne a este processo.

2. Padrão transcricional decorrente da perturbação no âmbito hormonal

Plantas tratadas com 1-MCP demonstraram um padrão geral de aumento do perfil de expressão de diversos genes. A depender do gene em questão, atribui-se isso ao grande acúmulo de auxina em plantas tratadas com este inibidor (Fig. 16A). Sabe-se que em *H. vulgare* a expressão de genes que codificam para H⁺-PPases aumentou após tratamento com 2,4D (Fukuda & Tanaka, 2006) e que a geração transiente de ROS em resposta à auxina é bem reportada no processo de gravitropismo da raiz (Joo *et al.*, 2001). Coincidentemente, em cana-de-açúcar genes relacionados a estresse oxidativo e a bombas de prótons vacuolares foram mais expressos em plantas com maiores níveis de auxina em S1 (Tabela 6). Tais bombas têm o papel de acidificar compartimentos intracelulares usando energia da hidrólise de ATP, criando o ambiente ideal para hidrólise executada por lipases, glicosidases e proteases. A atividade desta bomba é comumente alta em tecidos jovens em intensa proliferação celular (Fukuda & Tanaka, 2006). Como possível consequência dos altos níveis de auxina observados em S1 (Fig. 16A), H⁺-PPases de cana-de-açúcar (SCEQRT1024C07.g e SCSFSD1065B09.g) foram predominantemente positivamente reguladas (Tabela 6), corroborando resultados observados em *H. vulgare* (Fukuda & Tanaka, 2006). Além disso, estas proteínas têm função relevante no transporte de auxina. Foi reportado que a superexpressão da H⁺-PPase AVP1 em *A. thaliana* resultou em divisão celular aumentada nos primeiros estágios da organogênese, hiperplasia em raízes e parte aérea, bem como acúmulo de transcritos de H⁺-ATPase e PIN1 foram observados. Nesse sentido, tais bombas são capazes de interferir indiretamente em processos regulados por auxina, como é o caso da expansão celular (Li *et al.*, 2005).

Dentre os genes envolvidos no metabolismo de ROS expressos na raiz após tratamento com 1-MCP ressaltam-se os que codificam para a hemoglobina não simbiótica (Hb; SCCCLR2001E08.g) e para a *respiratory burst oxidase homolog* (RBOH; SCAGLR2011D02.g). Esta última é uma enzima chave na via de geração de ROS e consiste na proteína homóloga à subunidade catalítica gp91phox da NADPH oxidase de mamíferos (revisado por Sagi & Fluhr, 2006), enquanto Hb é uma proteína com alta afinidade por oxigênio e que otimiza a utilização do mesmo em situações de hipóxia (revisado por Dordas *et al.*, 2003). Em plantas alagadas de *Z. mays* o acúmulo de ambos transcritos foi abolido quando as plantas foram tratadas com 1-MCP (Rajhi *et al.*, 2011), enquanto que em cana-de-açúcar não alagada (presente trabalho) o cultivo com este inibidor estimulou o acúmulo destes mRNAs em S5 (Tabela 6). Os autores propuseram que a regulação positiva de RBOH no córtex radicular de *Z. mays* levou à geração de ROS e o acúmulo

destes deflagrou o processo de estresse oxidativo que levou à formação do aerênquima (Yamauchi *et al.*, 2011). Em cana-de-açúcar a regulação positiva do transcrito referente a RBOH não correspondeu a um aumento na formação de aerênquima (Fig. 15). Logo, a hipótese elaborada para *Z. mays* não foi corroborada pelos resultados do presente trabalho. Isso pode se dever à natureza constitutiva de sua formação em cana-de-açúcar. Curiosamente, sugere-se que ROS produzido por RBOHs poderia desempenhar um papel na sinalização de auxina durante o desenvolvimento da raiz (Müller *et al.*, 2012).

No que concerne à Hb, a ligação da mesma com a formação de aerênquima pode ser devido ao seu papel no aumento do *status* antioxidante celular por meio da eliminação de óxido nítrico, produzido após tratamento com auxina (Pagnussat *et al.*, 2003). A elevação da expressão deste gene leva à queda da produção de NO e parece desligar a via de biossíntese de etileno, direta ou indiretamente. Isso pode indicar a participação de Hb em uma resposta adaptativa a baixos níveis de oxigênio, regulando os níveis de NO e etileno e retardando a formação de aerênquima (Igamberdiev *et al.*, 2005). Notavelmente, plantas de *Medicago sativa* superexpressando Hb não formaram aerênquima em hipóxia (Dordas *et al.*, 2003). É possível que a maior expressão deste gene, observada no presente trabalho, tenha contribuído junto com o 1-MCP para os menores níveis de etileno observados para plantas tratadas com o inibidor (Fig. 41), embora outros experimentos sejam necessários para suportar esta hipótese.

Ambos os genes relacionados à resposta à hipóxia em outras espécies foram mais expressos em cana-de-açúcar cultivada em normóxia. Mesmo que o suprimento de oxigênio no ambiente tenha sido normal, os níveis deste gás em tecidos internos da planta podem apresentar baixas concentrações. Em tecidos com alta taxa metabólica e consequente elevado consumo de oxigênio, o gradiente de difusão precisa ser suficientemente alto para que o oxigênio seja conduzido ao longo dos tecidos de modo a suprir seu intenso consumo (Geigenberger, 2003). Visto que o tecido radicular de plantas cultivadas em 1-MCP apresentou alta atividade metabólica possivelmente ligada ao desenvolvimento do órgão (como alta expressão de enzimas envolvidas na expansão celular, discutida posteriormente) esta pode ter sido responsável pelo intenso consumo de oxigênio, mimetizando condições de hipóxia.

Já quanto ao padrão transcricional observado entre as amostras com contrastantes níveis de formação do aerênquima oriundas do experimento de estresse nutricional, foram observados genes relacionados à morte celular programada diferencialmente expressos. Conforme descrito pela literatura, genes relacionados a cálcio são elementos chave neste processo (Fig. 2). Coincidentemente, durante a formação de aerênquima em cana-de-açúcar, o mRNA correspondente à calmodulina apresenta uma tendência de aumento em direção ao segmento 5 (Fig. 11A). Estas são proteínas ligantes de cálcio que sofrem mudanças conformacionais na presença de altos níveis deste íon de modo a expor o domínio quinase (Harper *et al.*, 2004) e que foram mais expressas na formação do aerênquima em *Z. mays* (Rajhi *et al.*, 2011).

Além da calmodulina, o gene referente à uma cisteína protease também foi mais expresso em amostras apresentando maior formação de aerênquima (Fig. 11C), com transcritos se acumulando em direção ao segmento 5. O mesmo SAS foi também identificado na proteômica com padrão coincidente

(Grandis, 2015). A aplicação de quelantes de cálcio levou à redução em 90% da atividade de uma cisteíno protease em *Z. mays* (Subbaiah *et al.*, 2000) enquanto o efeito oposto se deu mediante estresse oxidativo (Solomon *et al.*, 1999) e etileno (Tripathi *et al.*, 2009). Esta classe de proteases já foi relacionada à formação de aerênquima (Beers, 1997), sendo possivelmente responsável pelo rompimento de um vacúolo e ativação proteolítica de enzimas supostamente aprisionadas neste compartimento (revisado por Hatsugai *et al.*, 2006). Já no que diz respeito à indução hormonal por trás da atividade transcricional destes genes, já foi sugerido que duas cisteíno proteases atuem na formação do xilema de *A. thaliana* em resposta à auxina (Ismail, 2004). Coincidente com este dado, um transcrito referente a uma cisteíno protease (Tabela 6), com SAS distinto ao mencionado anteriormente, também se acumulou mais em amostras tratadas com 1-MCP, que por sua vez tem maior concentração de auxina (Fig. 16A).

No que diz respeito ao ataque à parede celular durante a formação de aerênquima após estímulo por déficit nutricional, a atividade de enzimas de modificação de parede celular foi observada em *Z. mays* em valores superiores ao constatado para plantas cultivadas com nutrientes (He *et al.*, 1996b). Similarmente, o acúmulo de transcritos relacionados à degradação de parede celular também foi observado em cana-de-açúcar (Figs. 13 e 14). Transcritos referentes a uma expansina, proteína sem atividade enzimática mas com função no afrouxamento da parede celular, foram mais acumulados em S1 que nos demais segmentos de raiz (Fig. 13A). Expansinas atuam desestabilizando interações não covalentes entre os polissacarídeos, sobretudo pontes de hidrogênio entre xiloglucano e celulose, permitindo o deslocamento destes polímeros um em relação ao outro e a expansão celular em resposta ao aumento do turgor (McQueen-Mason & Cosgrove, 1995). O perfil transcricional deste gene em amostras com maior área ocupada pelo aerênquima é um resultado esperado, dado o fato de que a expansão celular é um passo inicial na formação de aerênquima (Leite, 2013; Grandis *et al.*, 2014). Surpreendentemente, os níveis deste transcrito atingiram valores superiores para plantas cultivadas em abundância de nutrientes em S2 e S3 quando comparados ao outro tratamento, porém sem alcançar o patamar observado para plantas em déficit nutricional em S1. Os altos níveis de auxina em S1 podem estar relacionados à maior expressão de outros genes que também codificam para expansinas e arabinofuranosidases (Tabela 6), visto que o aumento na atividade transcricional destes já foi reportada em resposta a este hormônio (Hutchinson *et al.*, 1999; Rose & Bennett, 1999) (Fig. 41). Neste contexto, arabinofuranosidases seriam requeridas para aumentar a porosidade da parede celular devido à hidrólise de cadeias laterais de pectinas. Entretanto, conforme discutido a seguir, embora seja possível sugerir que o alvo preferencial da arabinofuranosidase sequenciada seja arabinoxilano, não existem evidências suficientes para afirmar categoricamente se sabe se estas enzimas estão envolvidas no ataque a este polissacarídeo, pectinas ou ambos nas raízes de cana-de-açúcar.

Também relacionada à atividade de expansão celular, a enzima xiloglucano transglicosilase/hidrolase (XTH) já foi apontada como sendo elemento chave na formação de aerênquima em *Z. mays*, juntamente com a endoglucanase (He *et al.*, 1996a, Saab & Sachs, 1999; Rajhi *et al.*, 2011), embora não existam relatos no que concerne o ataque aos polissacarídeos alvo. No presente trabalho, o

perfil transcricional de XTH foi constante ao longo da raiz de plantas submetidas ou não ao estresse nutricional, sugerindo que esta enzima pode não estar relacionada à formação de aerênquima (Fig. 13B). Já os níveis de transcritos referentes à endoglucanase decresceram a partir de S1 até S3 em plantas controle. Este acumulou-se mais em plantas cultivadas na presença de nutrientes em S2 (Fig. 14D). Em cana-de-açúcar evidências sugerem manutenção da celulose, xiloglucano e arabinoxilano nos restos de parede celular que permanecem colapsados no cortex de S5 (Leite, 2013). Supõe-se que nas etapas finais da degradação do aerênquima se dê um ataque específico à celulose em regiões amorfas e que o remanescente deste polímero interaja com o arabinoxilano após este sofrer desramificação, formando um composto recalcitrante à hidrólise (Leite *et al.*, em preparação). Na ausência de degradação pronunciada da celulose, a atividade intensa de β -glucosidase não é requerida, conforme o perfil de expressão do gene correspondente sugere (Fig. 14C).

No que diz respeito ao arabinoxilano, arabinofuranosidases podem atuar na remoção dos resíduos de arabinose presentes neste polissacarídeo. Grandes quantidades de AX estão presentes em gramíneas, sobretudo as com possível emprego para produção de etanol de segunda geração (Pauly & Keegstra, 2008), contabilizando 40% das paredes celulares de cana-de-açúcar (estimado de acordo com De Souza *et al.*, 2013). Entretanto, os exemplos de degradação endógena de arabinoxilano são escassos (Tavares *et al.*, 2015).

A arabinofuranosidase sequenciada no presente trabalho apresenta um módulo predito de ligação a carboidrato com atividade descrita de ligação a xilanos (Lombard *et al.*, 2013), e seu padrão de mapeamento pelo transcriptoma ao longo da formação do aerênquima (Fig. 25A) sugere maior expressão nos segmentos 3 e 4. Nesse sentido, o papel de remoção das decorações de arabinose do arabinoxilano poderia ser desempenhado por esta proteína, corroborando a hipótese de formação do composto de xilano e celulose ao final da formação do aerênquima. Entretanto, a caracterização funcional desta enzima se faz necessária para confirmar a atividade sobre arabinoxilano ou pectinas.

Por outro lado, transcritos referentes a enzimas seguramente envolvidas no ataque a pectinas foram diferencialmente expressos entre as amostras oriundas de plantas cultivadas com distintas disponibilidade de nutrientes. O acúmulo de transcritos relacionados à degradação de pectina apresentaram um platô a partir de S2 (Fig. 13C-D), sendo mais expressos em plantas em déficit nutricional. Pectina acetil esterase e endopoligalacturonase podem agir sinergisticamente sobre o homogalacturonano, primeiro catalisando a desacetilação e, subsequentemente, quebrando as ligações glicosídicas entre os resíduos de ácido galacturônico (Micheli, 2001; Mohnen *et al.*, 2008). De fato, durante a formação de aerênquima em cana-de-açúcar a reação cruzada com o anticorpo anti-homogalacturonano diminuiu à medida que o córtex perdeu a sua integridade (Leite, 2013). Modificações neste polissacarídeo também foram reportadas nas etapas iniciais da formação do aerênquima em *Z. mays* assim como foi observado que o transcrito da endopoligalacturonase acumula-se em plantas com aerênquima em formação (Rajhi *et al.*, 2011). Ainda nesta planta, experimentos de estresse nutricional mostram que indivíduos do grupo controle apresentam coloração de Ca^{2+} no córtex da raiz visivelmente concentrada na lamela média

(Siyiannis *et al.*, 2011). Isso indica a integridade do homogalacturonano, uma vez que íons cálcio são aprisionados na estrutura deste polissacarídeo (McCann & Carpita, 2000). Em contraste, em plantas cultivadas na ausência de fontes de N, P ou S a marcação de cálcio é fraca e não mostra pontos óbvios de acúmulo (Siyiannis *et al.*, 2011).

Conforme sugerido por nosso grupo (Grandis *et al.*, 2014), a degradação da parede celular que ocorre na formação do aerênquima parece compartilhar muitas características com outros eventos em que a parede é atacada, como resposta a patógenos (Ohme-Takagi *et al.*, 2000), abscisão (Agustí *et al.*, 2008; Li & Yuan, 2008), senescência (Brummel *et al.*, 2004), germinação (Bewley, 1997) e amadurecimento de frutos. No caso deste último, a remoção de pectina abre acesso aos polissacarídeos restantes, permitindo o afrouxamento da parede celular e amolecimento do fruto, assim como os dados do presente trabalho em conjunto com relatos da literatura apontam para a formação do aerênquima. Em frutos em amadurecimento enzimas de degradação da pectina e fatores de resposta ao etileno são elementos chave bem descritos (Fischer, 1991; Karakurt & Huber, 2003; Giovannoni, 2004; Nishiyama *et al.*, 2007; Wang *et al.*, 2007).

No que diz respeito aos fatores de resposta a etileno, sabe-se que membros desta família de transcrição são responsáveis por promover a adaptação a várias condições de estresse como hipóxia em *O. sativa* e *A. thaliana* (Hinz *et al.*, 2010; Yang *et al.*, 2011; Xu *et al.*, 2006), seca, injúria e ataque de patógenos em *A. thaliana* e *G. hirsutum* (Fujimoto *et al.*, 2000; Jin & Liu, 2008), muitos dos quais também possivelmente levam à formação de aerênquima. De acordo com a sua sequência, o SAS de cana-de-açúcar SCAGLR2011B02.g é referente a uma proteína pertencente à família de fatores de transcrição RAV, a menor das sub-famílias da superfamília ERF, com apenas 6 membros em *O. sativa* e *A. thaliana* (Nakano *et al.*, 2006). Em raízes de *Z. mays*, transcritos correspondentes a RAV, 95% idêntico ao supracitado SAS, se acumularam no córtex radicular durante a formação do aerênquima (Rajhi *et al.*, 2011). Similarmente, o transcrito referente a RAV em cana-de-açúcar apresentou níveis mais elevados em plantas com formação de aerênquima favorecida (tratamento NPK), especificamente no primeiro segmento de raiz (Fig. 12D), sugerindo que possa estar relacionado com a regulação dos passos iniciais da formação de aerênquima. Tendo em vista a atuação de ERFs sobre genes que codificam para endopoligalacturonases em outros sistemas, este trabalho se voltou para obtenção da sequência *full length* destes genes, visando avaliar se em sistema heterólogo, ScRAV1 e/ou ScERF1 regula a degradação de parede por meio da expressão de enzimas como a endopoligalacturonase *ScEPG1* e/ou arabinofuranosidase, elucidando um possível elo entre a sinalização por etileno e a degradação de parede celular durante a formação do aerênquima.

3. Caracterização dos *loci* ocupados pelos genes alvo sequenciados

Com o objetivo de analisar o processo de regulação transcricional da degradação de parede celular em raiz de cana-de-açúcar, os genes referentes a dois fatores de transcrição e duas glicosil hidrolases expressos neste órgão foram sequenciados. Para tanto, a biblioteca de BACs da variedade R570 foi triada e um total de 38 BACs foram sequenciados. Em razão da alta similaridade entre as sequências codantes obtidas (Anexos 6-9), bem como a presença de versões de genes aparentemente não funcionais,

aspectos desde a conservação da estrutura gênica em relação à *S. bicolor* até a identificação de aminoácidos possivelmente críticos nas proteínas preditas foram explorados. Esta análise foi utilizada como critério para a escolha da sequência a ser empregada em experimentos posteriores. Como consequência foi possível acoplar os padrões estruturais com a atividade transcricional e o evento fisiológico.

Os BACs selecionados carregando as cópias hom(e)ólogas dos genes alvo foram analisadas quanto à sintonia e colinearidade em relação à *S. bicolor*. Tais análises revelaram maiores extensões atingidas pelos BACs de cana-de-açúcar em relação aos *loci* dos genes ortólogos em *S. bicolor* (Fig. 18-21). Embora em alguns casos a extensão dos genes fosse distinta entre cana-de-açúcar e *S. bicolor*, regiões intergênicas foram as grandes responsáveis por este fato. Dados similares obtidos na caracterização de 317 BACs de cana-de-açúcar apontaram a expansão do genoma desta planta (De Setta *et al.*, 2014), corroborada pelo presente trabalho.

De Setta e colaboradores (2014) observaram um bloco sintênico contendo o gene *rpa1* em 12 BACs e em *S. bicolor*. Já no presente trabalho, as sequências genômicas selecionadas se apresentaram estruturalmente mais variáveis em relação a *S. bicolor*, já que não foram observadas regiões em que a sequência de um dado BAC contemplasse em sua totalidade a região selecionada do genoma de *S. bicolor*. Todavia, blocos sintênicos correspondentes a parte dos genes de *S. bicolor* foram também identificados (Fig. 18-21). No caso dos BACs carregando o gene *ScEPG1*, o BAC SHCRBa_224_P16 é o que cobre uma maior extensão da região selecionada do cromossomo 1 de *S. bicolor* (Fig. 19). Neste BAC foram identificados os prováveis ortólogos de 12 dos 16 genes representados de *S. bicolor*. A sequência compreendida entre os genes 2 e 16 têm extensão de pouco mais que 114 Mb. Já em *S. bicolor* os respectivos ortólogos encontram-se a pouco mais de 95 Mb um do outro. Para os BACs carregando as versões dos genes *ScRAV1*, *ScERF1* e *ScARA1*, aqueles com maior cobertura da região ortóloga de *S. bicolor* são SHCRBa_023_O12 (Fig. 18), 247_B02 (Fig. 20) e 124_A06 (Fig. 21), respectivamente. A sequência compreendida entre os genes 2 a 9 (em SHCRBa_023_O12), 1 a 9 (em SHCRBa_247B02) e 1 a 4 (em SHCRBa_124A06) são de 41 a 82 Mb superiores às sequências entre os respectivos ortólogos em *S. bicolor*.

A presença de extensos TEs nas regiões intergênicas em cana-de-açúcar parece ser a maior responsável pelos maiores comprimentos de blocos sintênicos. Embora o conteúdo de TEs em BACs individuais seja consideravelmente heterogêneo, o conteúdo médio de 30% deste tipo de sequência (Tabela 7) é menor do que o estimado para o genoma, em torno de 50% (De Setta *et al.*, 2014). Isso seguramente se deve à busca voltada para sequências codantes realizada no presente trabalho. Em contrapartida, a prevalência de TEs contendo LTRs (Fig.18-21) é um dado consonante entre o presente trabalho e o De Setta e colaboradores (2014).

A presença de TEs em regiões eucromáticas de um dado genoma são associados a eventos como rearranjos entre cromossomos (Zhang *et al.*, 2011), perda e ganho de sequências gênicas (Cormack & Peterson, 1994) e outros eventos de natureza epigenética (Stotkin & Martienssen, 2007). A influência de tais elementos para a evolução do genoma é tema frequente em genomas de plantas, com destaque para

as gramíneas (Bennetzen, 2007). A pesquisa em Epigenética revelou que genomas menores apresentam uma espécie de "monitoramento epigenético" mais agressivo, de modo a conter a atividade de TEs. Em contrapartida, em genomas maiores este mecanismo é menos ativo (Michael, 2014). Sendo assim, a proliferação de TEs é um elemento chave para a expansão dos genomas complexos como o da cana-de-açúcar.

No que diz respeito a sequências codantes, as análises estruturais realizadas com os *loci* de *ScRAV1*, *ScEPG1*, *ScARA1* e *ScERF1* indicam que os quatro conjuntos de BACs representam regiões hom(e)ólogas (Fig. 18-21). Os genes hom(e)ólogos e o provável ortólogo em *S. bicolor* foram empregados na análise filogenética. Em muitos casos a topologia da árvore obtida refletiu diferenças no conteúdo de TEs na região examinada. Ambos os BACs do agrupamento I na análise de sintenia de *ScRAV1*, SHCRBa_199_K01 e 170_P07, apresentam extensos retrotransposons contendo LTR entre os genes 9 e 10 (Fig. 18). O mesmo não é observado em *S. bicolor* e nem nos demais BACs. Os dois BACs do agrupamento I na análise de sintenia do gene *ScEPG1* (SHCRBa_023_N06 e 191_M21) também apresentam longos retrotransposons com LTR entre os genes 6 e o gene alvo (Fig. 19). A presença do gene B é uma outra característica única a este agrupamento. Notavelmente, em SHCRBa_191_M21 a região compreendida entre os genes 12 e 15 está invertida em relação a *S. bicolor* e aos demais hom(e)ólogos (sombreado em laranja; Fig 18), representando mais uma característica particular a este agrupamento. Já nos BACs contendo o gene *ScERF1* e *ScARA1* as características únicas de vários BACs de um mesmo agrupamento não são tão evidentes (Figs. 20 e 21). Sobreretudo no caso de *ScARA1* (Fig. 21) as populações de genes em cada BAC são muito heterogêneas, impossibilitando a identificação de padrões.

Em razão desta elevada heterogeneidade e pela inexistência de blocos sintênicos que cubram todas as quatro regiões selecionadas do genoma de *S. bicolor*, somente as sequências dos genes alvo (para *ScRAV1* e *ScERF1*) ou do alvo concatenado ao gene vizinho (para *ScEPG1* e *ScARA1*) puderam ser empregadas na filogenia. O padrão repetitivo de dois ou três BACs formando um clado basal entre as demais sequências de cana-de-açúcar nas filogenias é notável.

Resultado similar foi observado nas análises filogenéticas empregando sequências promotoras de genes que codificam para arabinofuranosidase e endopoligalacturonase. Os padrões de distribuição dos sítios para fatores de transcrição das famílias RAV e ERF são distintos nos promotores acompanhando de modo geral a topologia da árvore. No caso da arabinofuranosidase, a presença de nove destes sítios de modo geral regularmente espaçados nas sequências promotoras foi identificada nos BACs SHCRBa_109_O23 e 133_H06, alocados no agrupamento I (Fig. 23A). Este mesmo padrão no qual ambos BACs divergem dos demais é observado nas filogenias empregando sequência promotora e sequência do bloco sintênico (concatenado *ScARA1* e gene 4; Fig. 21), conforme discutido acima. Já no agrupamento 2 os padrões de ocorrência deste sítio são menos claros. Entretanto, os promotores de *ScARA1* nos BACs SHCRBa_048_E22, 111_H15, 143_N21 e 191_E06 podem ser considerados um sub-grupo mais conservado, visto que a distribuição de sítios GCCCORE segue um padrão bem similar. Embora tais

distintos padrões existam, a expressão dos diferentes genes hom(e)ólogos não é diferencial (Fig. 27A), conforme será melhor explorado adiante.

No caso de endoglicanases, levando em conta o número e localização sítios de ERF e RAV, é possível observar que a topologia da árvore dos hom(e)ólogos *ScEPG1* dos BACs SHCRBa_023N06 e 191_M21 divergem dos demais, consolidando o agrupamento I. Por sua vez os BACs do agrupamento II só apresentam cinco sítios para o domínio AP2 de RAV, enquanto no agrupamento III, mais heterogêneo, tais sítios são menos numerosos ou intercalados por motivos GCCCORE (Fig. 24B)

A manutenção do padrão de dois ou três BACs divergirem precocemente na filogenia é observado tanto nas análises de sintonia (Figs.18-21) como nas sequências promotoras de *ScARA1* e *ScEPG1* (Figs. 23 e 24). É possível sugerir que os BACs ocupando os agrupamentos I para os quatro conjuntos de sequências consistam na contribuição de *S. spontaneum* no genoma de cana-de-açúcar. Por outro lado, os agrupamentos II (e III, quando houver) representariam as sequências oriundas do outro parental, *S. officinarum*. Embora os altos valores de *bootstrap* sustentem esta hipótese, o sequenciamento da região ortóloga nas duas espécies parentais pode efetivamente confirmá-la. Essa é a primeira vez que uma análise de várias regiões genômicas ortólogas indicam a contribuição das espécies parentais para o genoma híbrido de cana-de-açúcar por métodos filogenéticos.

4. Análise comparativa das proteínas codificadas por *ScARA1a* e *ScEPG1a* e o possível emprego biotecnológico

Arabinofuranosidases e endopoligalacturonases apresentam ampla caracterização funcional no que diz respeito à degradação de parede celular e regulação de várias etapas do desenvolvimento vegetal. Além disso, a estrutura gênica das GH sequenciadas de cana-de-açúcar é coincidente com o observado para outras plantas. Além disso, o padrão transcricional de ambas é bem descrito devido a ação destas enzimas durante o amadurecimento de frutos. No que diz respeito a arabinofuranosidases, a expressão de *FaARA1* de *Fragaria ananassa* foi detectada em frutos com padrão coincidente com o aumento da atividade da enzima correspondente durante o amadurecimento (Rosli *et al.*, 2009). O mesmo incremento no acúmulo de transcritos de *PcARF1* e *LeARAF1* foi observado para frutos de *Pyrus communis* (Sekine *et al.*, 2006) e *Lycopersicum esculentum*, sendo este último negativamente regulado por etileno (Itai *et al.*, 2003). Por fim, em *A. thaliana* os transcritos de *AtADS1* e 2, referentes a arabinofuranosidases, são ubíquos, sendo o primeiro sobretudo identificado em regiões de divisão e expansão celular próximas ao ápice radicular (Fulton & Cobbett, 2003). No que diz respeito à regulação hormonal e ao padrão transcricional, *ScARA1* apresenta características similares às de *LeARAF1* de *L. esculentum*, visto que sua expressão foi mais pronunciada em amostras com menor concentração média de etileno (Fig. 16B). Por outro lado, assim como os genes *AtADS1* e 2 de *A. thaliana*, *ScARA1* foi mais expresso próximo ao ápice radicular (Tabela 6), em regiões com marcante atividade de expansão celular (Leite, 2013).

Já quanto às endopoligalacturonases, acredita-se que só o genoma de *A. thaliana* codifique 69 possíveis enzimas com esta atividade. Diante disso, essas plantas dispõem de um padrão espaço-

temporal de expressão gênica de endopoligalacturonases bastante complexo (González-Carranza *et al.*, 2007). A literatura mostra que endopoligalacturonases têm amplos papéis, pois estas enzimas atuam na maturação do grão de pólen e formação do tubo polínico em *Brassica napus* (Roberts *et al.*, 1993), deiscência da antera e abscisão floral em *A. thaliana* (Jenkins *et al.*, 1999) e amadurecimento de frutos em diversas espécies (Roberts *et al.*, 2002; Kim *et al.*, 2006). Além disso, possuem altos níveis de expressão ou atividade enzimática em folhas senescentes e região da qual emergem raízes laterais (González-Carranza *et al.*, 2007). Em vários destes eventos já é consensual que estas enzimas exerçam sua função promovendo expansão e/ou separação celular por meio da degradação da lamela média (Jenkins *et al.*, 1999; Rose *et al.*, 2003).

Sabe-se que o a expressão de *CkPGA* e B de *Actinidia chinensis* ocorre somente diante do aumento dos níveis endógenos de etileno (Wang *et al.*, 2000), fato coincidente com a presença de elementos de resposta no promotor de *AtEPG* em *A. thaliana* (Jenkins *et al.*, 1999). Por outro lado, a auxina também tem um papel indutor na expressão de genes membros desta família, como observado no caso do gene *ADPG2* de *A. thaliana* (González-Carranza *et al.*, 2007). O presente trabalho não gerou dados a respeito da concentração de etileno existente nas amostras em que *ScEPG1* foi diferencialmente expressa. Entretanto, visto que a escassez nutricional provavelmente leva a uma maior sensibilidade a etileno (He *et al.*, 1992; He *et al.*, 1994) e que o presente trabalho sugere que um balanço hormonal entre auxina e etileno seja responsável por deflagrar a formação do aerênquima, é plausível sugerir que a expressão de *ScEPG1* possa ser regulada por ambos hormônios. Um outro fato que sustenta esta sugestão é a presença de elementos em *cis* para ligação de fatores de resposta à etileno e auxina (Fig. 38) na sequência promotora do hom(e)ólogo mais expresso deste gene de acordo com o mapeamento de *reads* do RNAseq sem normalização (Fig. 25B).

Assim como endopoligalacturonases, arabinofuranosidases também são requeridas em outros processos que não o amadurecimento. Enquanto alguns genes que codificam para arabinofuranosidases em *Hordeum vulgare* têm papel reportado na depolimerização de arabinoxilanos do tecido materno durante o desenvolvimento da semente (Laidlaw *et al.*, 2012), outras demonstraram capacidade de hidrolisar tanto arabinoxilano como arabinano e arabinogalactano (Lee *et al.*, 2001; Chaves Montes *et al.*, 2008). A proteína codificada pelo gene *AXAH2* de *H. vulgare* (GenBank JQ303076) expressa em sistema heterólogo, com sequência codificante 85-86% idêntica às de cana-de-açúcar, apresenta valores relevantes de atividade enzimática sobre arabinoxilano, 1,5- α -L-arabinopentaose, arabinano não ramificado e arabinogalactano, o que sugere atividade sobre ramnogalacturonano I e hemiceluloses (Laidlaw *et al.*, 2012). Sugere-se que funcionalmente esta enzima possa estar inserida no contexto de deposição, expansão ou *turnover* de parede celular, sendo um consenso de que a mesma possa apresentar um papel bifuncional na degradação de hemiceluloses e pectina (Lee *et al.*, 2001).

Visando o levantamento das demais análises comparativas que sugerissem a funcionalidade tanto de *ScARA1* quanto das sequências obtidas, uma sequência dentre cada um dos genes hom(e)ólogos foi selecionada. Inicialmente o critério estabelecido consistiu em selecionar a sequência mais

expressa, razão pela qual os BACs SHCRBa_021_A10, 111_H15, 188_L14 e 224_P14 carregando um hom(e)ólogo de *ScRAV1*, *ScARA1*, *ScERF1* e *ScEPG1*, respectivamente foram escolhidos. As primeiras análises disponibilizadas consistiram no mapeamento de *reads* do transcriptoma de quatro dos cinco segmentos de raiz sobre todas as sequências codantes dos hom(e)ólogos dos genes alvo. Tal mapeamento se deu manualmente (para *ScARA1* e *ScEPG1*) e por *Bowtie* (para *ScERF1* e *ScARA1*).

O mapeamento empregando as duas primeiras técnicas mencionadas levou à seleção da sequência de *ScARA1* presente no BAC SHCRBa_111_H15 (*ScARA1a*) (Fig. 27B). O alto conteúdo de arabinoxilano na parede celular de cana-de-açúcar (De Souza *et al.*, 2013) pode ser um dos fatores responsáveis pelos altos níveis de transcritos de *ScARA1* mapeados pelo transcriptoma, sendo este polissacarídeo um dos possíveis substratos desta enzima, conforme discutido acima. Por sua vez, o perfil de expressão do hom(e)ólogo de *ScEPG1* no BAC SHCRBa_224_P16 (*ScEPG1a*) (Fig. 25B) coincide com hipóteses correntes para formação do aerênquima, que se inicia na transição S2-S3 com a degradação da pectina presente na lamela média (Gunawardena *et al.*, 2001b; Leite, 2013) e a proteína correspondente é forte candidata a desempenhar esta função. Além disso, processos que dependem de perda de adesão celular a participação de ambos hormônios é requerida (Agustí *et al.*, 2008).

Na sequência protéica codificada pelo gene *ScARA1a* foi predito o domínio α -L-AF_C (Tabela 8; Figs. 23C e 35). Dentre as enzimas com o mesmo domínio e com atividade já descrita na família GH51 encontra-se a AF de *Bacillus stearothermophilus* T-6 (Gilead & Shoram, 1995) e AF de *Clostridium acetobutylicum* ATCC 824 (Lee & Forsberg, 1987). A primeira é produzida em grandes quantidades quando a bactéria é cultivada em L-arabinose, arabinano de beterraba ou farelo de aveia. Já a AF de *C. acetobutylicum* apresentou valores expressivos de atividade sobre arabinano de beterraba quando comparado à enzima de *B. stearothermophilus*. No que diz respeito à ação sinérgica, a incubação de xilano de aveia com arabinofuranosidase e xilanase de ambas as bactérias liberaram como produto da reação L-arabinose, xilose e xilobiase, sugerindo ação das AF sobre arabinoxilano (Lee & Forsberg, 1987; Gilead & Shoram, 1995).

No que diz respeito a aspectos estruturais, arabinofuranosidases de microrganismos foram empregadas como base de comparação, visto que a estrutura tridimensional desta enzima em plantas ainda não foi descrita (Lombard *et al.*, 2013). Embora haja regiões de alta conservação entre a enzima bacteriana e a de plantas, grandes *gaps* foram também notados no alinhamento (Fig. 36). Possivelmente o dobramento desta enzima em plantas é similar ao de microrganismos, embora seja possível que os 15 kDa adicionais da proteína predita a partir do gene *ScARA1a* (Tabela 8) resultem em uma estrutura mais complexa em relação à AbfA de *G. stearothermophilus* (Hövel *et al.*, 2003).

O alinhamento múltiplo das sequências de proteínas oriundas dos trabalhos supracitados demonstrou que a sequência predita de *ScARA1* tem um alto grau de conservação em relação às demais, sobretudo nos resíduos que cercam o sítio catalítico, de acordo com o observado para AbfA de *G. stearothermophilus* (Fig. 36). Dos dez aminoácidos considerados como cruciais para arabinofuranosidases do clã GH-A, presentes em AbfA de *G. stearothermophilus* (Hövel *et al.*, 2003) (Glu29, Arg69, Asn74,

Asn174, Glu175, His244, Tyr246, Glu294, Gln351 e Trp344), sete foram identificados em cana-de-açúcar (Glu61, Arg256, Asn378, Glu379, His421, Glu457 e Trp461; retângulos amarelos e verdes na Figura 36).

Na arabinofuranosidase bacteriana o resíduo Glu175 faz parte da díade catalítica e estabiliza uma ponte de hidrogênio com His224. Em ScARA1a o resíduo de Glu está conservado na posição 379 (retângulo em verde; Fig.36) e a histidina em 421 (retângulo em amarelo; Fig. 36). Na enzima de *G. stearothermophilus* o resíduo Asn174 também estabiliza pontes de hidrogênio com a díade catalítica (Hövel *et al.*, 2003) e a mutação Asn→Asp resulta em quase total perda da atividade enzimática (Navas & Béguin, 1992). Aparentemente este resíduo atua na formação do intermediário enzima-glicosil (Ryttersgaard *et al.*, 2002). Já em cana-de-açúcar, este resíduo encontra-se na posição 378. Já o resíduo Arg69 de AbfA encontra-se no fundo do bolso formado pelo sítio ativo, tendo a função de manter deprotonado o resíduo Glu294. Este último por sua vez forma pontes de hidrogênio com Tyr246 para estabilizar o subsítio -1 do arabinofuranosil. Na região alinhada da sequência de cana-de-açúcar em relação à enzima bacteriana encontra-se Arg256 e Tyr430 (retângulos amarelos; Fig. 36).

Embora localizado no extremo N-terminal da proteína, a mutação do resíduo Glu28 apresentou efeitos severos à atividade catalítica da α -L-arabinofuranosidase de *Thermobacillus xylanilyticus*, possivelmente representando um terceiro resíduo catalítico (Debeche *et al.*, 2002). Este resíduo é relevante para estabilização de pontes de hidrogênio com outros dois aminoácidos, Asn74 e Gln351, auxiliando na ligação ao substrato (Hövel *et al.*, 2003). Na sequência predita a partir de ScARA1a, na posição correspondente a Asn74 observa-se Cys261 e um *gap* na região que alinha com o aminoácido Gln351. O mesmo foi observado para a sequência de *S. bicolor* também empregada no alinhamento (retângulos azuis; Fig. 36). É possível que o trecho de posições 350 a 356 (AQLVNVI) da proteína de *G. stearothermophilus* seja exclusiva de microrganismos e sua ausência em plantas não afete a funcionalidade das arabinofuranosidases de eucariotos superiores.

Por fim, a manutenção do resíduo Trp461 em ScARA1a em relação a Trp344 de AbfA é relevante visto que este aminoácido parece ser responsável pela discriminação de substratos furanosídicos reconhecidos pela família GH51 em relação a substratos piranosídicos (retângulos amarelos; Fig. 36; Hövel *et al.*, 2003). Entretanto, em teoria AbfA é capaz de acomodar também um substrato xilopiranosídico, o que é corroborado pela existência de enzimas bi-funcionais α -L-arabinofuranosidases/ β -D-xilosidases, embora a eficiência catalítica seja bastante distinta entre as duas atividades (Hövel *et al.*, 2003).

A partir das análises comparativas executadas neste trabalho sugere-se que ScARA1a codifique uma arabinofuranosidase funcional, permitindo a execução de etapas posteriores de expressão heteróloga e caracterização funcional da enzima em interação com seu substrato, experimentos estes que são almejados pelo nosso grupo de pesquisa.

No que concerne a proteína predita a partir do hom(e)ólogo ScEPG1a, esta também apresenta conservação da maioria dos sítios apontados como críticos para a atividade da endopoligalacturonase II de *A. niger*. Experimentos de mutação sítio dirigida empregando esta enzima revelaram que alterações na identidade de Asp180, Asp201 e Asp202 resultaram em queda dramática nos

valores de atividade específica. Já os resíduos Arg256 e Lys258 aparentemente estão envolvidos na ligação ao substrato. Por fim, His223, também relevante para a catálise, provavelmente compartilhando prótons com Asp201. Como consequência, este último permanece em seu estado de ionização adequado para protonar o produto da reação (Armand *et al.*, 2000). Destes seis resíduos, quatro estão presentes na proteína predita a partir de *ScEPG1a* (Asp229, Asp261, Asp262 e Arg316, em retângulos amarelos e verdes na Figura 30). Além destes, a presença e um resíduo de triptofano vizinho ao aminoácido correspondente à Lys258 em plantas é apontado como responsável pela insensibilidade de endopoligalacturonases de planta em relação à ação de proteínas inibidoras de poligalacturonases (*polygalacturonase inhibiting protein*, PGIP; retângulo azul na Figura 30). Este inibidor atua de forma competitiva, impedindo a interação necessária para ligação ao substrato e, concomitantemente, cobre a fenda onde se situa o sítio ativo (Federici *et al.*, 2001). Dois resíduos de triptofano foram identificados na endopoligalacturonase predita de cana-de-açúcar nas proximidades da posição esperada, Trp317 e Trp320 (retângulos azuis; Fig. 30) sugerindo que *ScEPG1a* seja insensível ao inibidor, tal qual enzimas de plantas.

Os resíduos His223 e Lys258 da enzima de *A. niger* não são conservados na sequência predita da endopoligalacturonase de cana-de-açúcar (retângulos azuis; Fig. 30). Estudos de mutação sítio dirigida demonstraram a relevância destes resíduos. Endopoligalacturonases carregando a mutação H223Q resultaram em valores de atividade específica de até 10^4 vezes inferior à enzima selvagem. Entretanto, mutações menos severas (H223A/S) levaram a alterações mínimas na atividade enzimática. As enzimas preditas a partir de *ScEPG1a* e do provável ortólogo de *S. bicolor* apresentam um resíduo de tirosina alinhado à His223 de *A. niger*. Visto que a alteração His→Ser pouco alterou as características funcionais da enzima do fungo, espera-se que a existência de um outro aminoácido polar não carregado nesta posição (como ocorre na sequência da endopoligalacturonase das duas gramíneas) não afete a funcionalidade da proteína.

Entretanto, no que diz respeito ao resíduo Lys258 de *A. niger* a situação parece mais complexa, visto que na proteína de cana-de-açúcar este resíduo não é conservado, embora o seja na enzima de *S. bicolor* (retângulo azul; Fig. 30). Embora dentre as enzimas mutantes de *A. niger* aqueles carregando a mutação no resíduo Lys258 tenham apresentado os maiores valores de Km, os autores especulam que a falta de interação do substrato com um dos resíduos Arg ou Lys pode alterar a geometria de distorção ideal do mesmo (Armand *et al.*, 2000), afetando consideravelmente a funcionalidade da enzima.

A ausência do resíduo de lisina correspondente a Lys258 de *A. niger* em *ScEPG1a* pode levar à perda de funcionalidade na mencionada proteína. A observação do alinhamento da sequência codificante de *ScEPG1a* com a sequência genômica correspondente mostra que o códon correspondente ao 318º resíduo de aminoácido (Fig. 29). Este localiza-se na borda do sétimo éxon da sequência codificante e 63 nucleotídeos após o fim do elemento de transposição predito na sequência genômica (Fig. 24C). Esta distinção entre a sequência de *ScEPG1* presente no BAC SHCRBa_224_P16 em relação aos demais pode ser o fator responsável pela proteína predita a partir do mencionado BAC não apresentar o resíduo crítico

de lisina. A ausência de tal resíduo pode tanto se dever a uma variante protéica não funcional ou de propriedades distintas. Ainda é possível que tenham ocorrido erros na predição de genes e, diante desta possibilidade é necessário que a predição seja validada por RT-PCR e sequenciamento.

5. Aspectos estruturais e funcionais preditos dos fatores de transcrição sequenciados

O domínio de ligação ao DNA AP2 ocorre em ambos fatores de transcrição sequenciados. Entretanto, as proteínas codificadas pelos genes *ScERF1a* e *ScRAV1a* são alocadas em subfamílias distintas. Elementos da família ERF apresentam somente um domínio AP2, em oposição a proteínas da família AP2, com dois destes domínios (Nakano *et al.*, 2006).

Primeiramente descrito no contexto de seu papel exercido no desenvolvimento floral em *A. thaliana* (Yanofsky 1995; Weigel, 1995), o domínio AP2 se estende por cerca de 60 aminoácidos, sendo essencial para a atividade sequência-específica de ligação ao DNA (Ohme-Takagi & Shinshi, 1995). A partir da análise do alinhamento na região do domínio AP2 predito nas proteínas RAV e ERF de cana-de-açúcar observamos relativa conservação em relação à mesma região de EREBPs de *A. thaliana* e *N. benthamiana*, dentre estes a proteína empregada no trabalho de Allen e co-autores (1998). Dentro deste domínio se situam dois elementos, nomeados RAYD e YRG, ambos altamente conservados nas proteínas preditas a partir de *ScERF1a* e *ScRAV1a* (Fig. 33). A presença da região *core* do elemento RAYD é crítica para a funcionalidade de fatores de transcrição com domínio AP2, visto que cerca de 18 resíduos dentro deste elemento formam uma α -hélice anfipática (Okamuro *et al.*, 1997; Allen *et al.*, 1998), possivelmente participando de interações proteína-proteína (Jofuku *et al.*, 1994).

A evolução do domínio AP2 e características dos genes que os apresentam também são fatos já descritos na literatura. Assim como a maioria dos membros da família RAV e ERF, os hom(e)ólogos de cana que codificam para estas proteínas apresentaram apenas um (para *ScERF1*; Fig. 22A) ou nenhum (para *ScRAV1*; Fig.22B) íntron (Sharoni *et al.*, 2012). Acredita-se que a transferência lateral de vírus ou bactérias para plantas tenha dado origem a esta superfamília de fatores de transcrição. Endonucleases de bactérias chamadas *homing endonucleases* são capazes de clivar fitas duplas de DNA de um dado hospedeiro, inserindo uma sequência molde em um alelo alvo. O hospedeiro repara o DNA clivado usando a sequência inserida como molde, consequentemente duplicando-a. Interessantemente, tais endonucleases apresentam o domínio AP2 capaz de ligar-se a sequências poli(dG)/poli(dC) GCCGCC (Magnani *et al.*, 2004), estando presentes também no genoma e cloroplastos e mitocôndrias (Gimble *et al.*, 2000).

A atividade de ligação ao DNA do domínio AP2 também foi estruturalmente caracterizada. A elucidação da estrutura tridimensional AtERF1 de *A. thaliana* demonstrou que resíduos de arginina e triptofano são os principais responsáveis durante o contato com o DNA (Allen *et al.*, 1998). Esses dois resíduos ocorrem nas proteínas preditas de cana-de-açúcar (retângulos em azul nas Figuras 31 e 32), estando contidos em uma região altamente conservada e que assume a conformação em folha β na

proteína de *A. thaliana* (Fig. 33). A atuação do fator de transcrição se dá por meio de uma região em folha β tripla antiparalela que contacta o ácido nucléico (Allen *et al.*, 1998).

Por fim, a ocorrência de um domínio AP2 e um domínio B3 na mesma proteína são típicos de membros da família RAV (do inglês *related to ABI3/VP1*) (Giraudat *et al.*, 1992; Kagaya *et al.*, 1999). Proteínas ABI3/VP1 (*Abscisic acid-insensitive 3/Viviparous 1*) são fatores de transcrição que regulam genes responsivos a ABA durante o desenvolvimento da semente (Lazarova *et al.*, 2002; Suzuki *et al.*, 2003). Estas proteínas apresentam três domínios de ligação ao DNA formados majoritariamente por aminoácidos de caráter básico (B1, B2 e B3) e compartilham com RAV o terceiro destes domínios, B3. Embora originalmente tenha sido reportado que a ligação da proteína foi observada quando ambos sítios estavam separados entre si por no máximo 8 pb e posicionados um em relação ao outro em duas orientações possíveis, trabalhos posteriores demonstraram que a presença de só um dos sítios é suficiente para a regulação transcricional ocorra (Li *et al.*, 2011), ainda que a constante de dissociação entre DNA e proteína seja relativamente baixa (Yamasaki *et al.*, 2004).

Embora a caracterização de aspectos evolutivos, estruturais e funcionais de fatores de transcrição com domínio AP2 seja extensa, alvos específicos destes fatores de transcrição no contexto da degradação da parede celular ainda são pouco explorados.

6. Regulação transcricional de *ScEPG1a* por *ScRAV1a*

A identificação de possíveis elementos de ligação a fatores de transcrição nos promotores de genes que codificam para ambas glicosil hidrolases sugeriu que os fatores de transcrição sequenciados podem ter um papel na regulação transcricional destas enzimas (Figs. 37 e 38). A presença de sítios RAV1AAT é uma característica comum dos promotores dos hom(e)ólogos *ScARA1a* e *ScEPG1a*. Por sua vez, os padrões de mapeamento no transcriptoma e também o obtido por qPCR indica que genes de glicosil hidrolases (Figs. 25A, 27A e 13C) são mais expressos quando *ScRAV1* é menos expresso (Figs 26A e 12D). Tais perfis opostos de *ScRAV1* e genes de glicosil hidrolases sugerem um papel repressor do primeiro sobre os segundos, o que foi corroborado pelos ensaios de transativação para *ScEPG1a* (Fig. 40D). Estes demonstraram que em folhas de *N. benthamiana* o fator de transcrição *ScRAV1* apresentou efeito repressor sobre o gene repórter clonado em fase com o promotor de *ScEPG1a*, reduzindo cerca de 50% da atividade da enzima codificada pelo mesmo (Fig. 40D). Em conjunto ambos resultados apontam fortemente para o papel repressor de *ScRAV1* sobre a expressão do gene referente a endopoligalacturonase.

O papel de RAV como repressor transcricional já foi previamente reportado. Plantas de *N. benthamiana* superexpressando o fator RAV de *Glycine max* (GmRAV) apresentaram alterações morfológicas e fisiológicas que incluíam reduzidas taxa de crescimento da parte aérea e de alongamento da raiz, desenvolvimento floral retardado, queda na atividade fotossintética e no conteúdo de clorofila (Sohn *et al.*, 2006). Ainda que RAV1 tenha demonstrado ser um indutor da senescência foliar em *A. thaliana* (Zhao *et al.*, 2008; Woo *et al.*, 2010), o mesmo gene foi apontado como um regulador negativo geral do desenvolvimento (Hu *et al.*, 2004). De fato, a proteína predita codificada pelo gene *ScRAV1a* carrega um

motivo RLFGV na região C-terminal (sublinhado em verde na Figura 32), uma assinatura de fatores de transcrição com atividade repressora (Ikeda & Ohme-Takagi, 2009).

Um papel mais específico para RAV foi proposto por Taken e co-autores (2010) para *Malus domestica*. Visando à identificação de reguladores potenciais de uma endopoligalacturonase chave no processo de maturação, foram analisados vários membros da superfamília AP2/ERF no que diz respeito aos perfis transcricionais e ensaios de transativação em *N. benthamiana*. Entre eles, um fator de transcrição RAV foi responsável por um dos valores mais baixos de ativação do repórter após a co-infiltração com o promotor da endopoligalacturonase em questão. Curiosamente, após tratamento com etileno os níveis de transcrição do mencionado RAV diminuíram, enquanto que aqueles para endopoligalacturonase apresentaram-se elevados.

Devido ao fato que fatores de transcrição da sub-família RAV possuem dois domínios de ligação ao DNA, estes podem se ligar especificamente a um motivo paritário (*bipartite recognition sequence*), composto por um par de elementos em *cis*. O domínio AP2 se liga ao elemento CACAA (RAV1AAT), enquanto B3 reconhece CACCTG (RAV1BAT). A caracterização do comportamento de reconhecimento de um motivo paritário de fatores de transcrição da família RAV demonstrou restrições para o espaçamento e orientação dos mesmos (Kagaya *et al.*, 1999). Mais tarde, outro estudo mostrou que a ligação a somente um destes elementos pode levar à regulação transcricional de um dado gene alvo (Li *et al.*, 2011), ainda que com uma constante de dissociação relativamente baixa (Yamasaki *et al.*, 2004). De acordo com este dado, ainda que somente elementos RAV1AAT tenham sido identificados na sequência de EPGpm (Fig. 38) isso parece ter sido suficiente para interação de ScRAV1 com o promotor de ScEPG1a (Fig. 40D).

Considerando a interação entre os dois, é possível sugerir que o fator de transcrição RAV atue modulando a expressão do gene da endopoligalacturonase, uma vez que ambos parecem ter um papel durante a formação do aerênquima em raízes de cana-de-açúcar. Visto que ScRAV1a não aparenta ser um repressor forte (Jacqueline Grimma-Pettenati, comunicação pessoal), parece mais provável que esse repressor atue modulando ou evitando níveis demasiadamente elevados do transcrito de ScEPG1a ao invés de desligá-lo completamente. Quanto ao fato dos níveis do transcrito de ScEPG1a atingirem o pico quando as células estão iniciando a formação do aerênquima, parece razoável supor que RAV atue no sentido de reduzir ligeiramente ou impedir que os níveis dos transcritos de ScEPG1a permaneçam ascendentes em um ponto em que a atividade de endopoligalacturonase parece não ser requerida. Deve-se lembrar que a região onde se forma o aerênquima é precisamente localizada, sendo portanto necessário manter a integridade de parte do tecido do córtex durante o desenvolvimento. Conforme os níveis de transcritos de ScRAV1 decaem, o efeito repressor sobre ScEPG1 deve tornar-se mais brando e, consequentemente, a digestão da pectina presente na lamela média pode ocorrer. Além disso, o papel repressor operado por RAV é provavelmente parte de uma intrincada rede de ativadores e outros repressores atuando de modo sinérgico. De fato, o motivo RLFGV presente em proteínas desta família são requeridos para interação destes com TPL/TRP (TOPLESS e TPL-related; Causier *et al.*, 2012), que juntos

provavelmente atuam como um complexo, reprimindo a transcrição de reguladores florais. Assim, o equilíbrio entre todos os membros deste complexo de regulação seria responsável pelo o destino das células no contexto do aerênquima.

7. A estrutura e funcionalidade do genoma de cana-de-açúcar para os quatro genes alvo

Após finalizadas as análises *in silico* das quatro proteínas preditas, bem como os ensaios de transativação, o procedimento de mapeamento foi revisto. Neste terceiro procedimento o mapeamento por *Bowtie* foi normalizado de acordo com o tamanho das sequências codantes e das bibliotecas de *reads*. Como resultado, observou-se que a maioria das sequências dos hom(e)ólogos dos genes alvo é expressa em níveis similares, dentre elas os BACs selecionados supracitados.

Como as análises normalizadas não apontam uma sequência mais expressa (Fig. 25A-28A), é possível sugerir que (i) de fato a cana-de-açúcar expresse igualmente os genes hom(e)ólogos destes quatro genes ou (ii) o peso dado pelo procedimento computacional às diferenças entre os comprimentos das sequências codantes (Anexo 10A) e/ou do número de reads em cada biblioteca do transcriptoma (Anexo 10B) mascarou sutis diferenças em número de *hits* do alinhamento. Para os quatro genes o padrão de expressão ao longo dos segmentos é muito similar se comparado com os respectivos padrões de expressão gênica obtidos por qPCR empregando oligonucleotídeos desenhados com base nos respectivos SAS. Tal comparação é mais evidente no caso de *ScEPG1* (comparar Figuras 13C e 25A) e *ScRAV1* (comparar Figuras 12D e 26A), pois para estes existem dados quanto à expressão gênica ao longo de todos segmentos de raiz. Sendo assim, a divisão estrutural dada pela filogenia parece não refletir um padrão de expressão diferencial entre distintos grupos de BACs.

Os eventos de cruzamento entre *S. spontaneum* e *S. officinarum* dando origem à cana-de-açúcar comercializada para produção de etanol são recentes. Os dados deste trabalho sugerem que, ao menos para os hom(e)ólogos dos genes-alvo, o período decorrido desde o surgimento da cana-de-açúcar moderna foi curto para que esta planta desenvolvesse estratégias de regulação precisas o suficiente para coordenar a expressão de um hom(e)ólogo em detrimento dos demais. Embora eventos de recombinação entre os cromossomos das espécies parentais tenham sido descritos (D'Hont, 2005), a regulação fina da atividade transcricional diferencial entre genes hom(e)ólogos altamente similares parece não existir para os genes-alvo nestre trabalho. Dentre os hom(e)ólogos mapeados, somente o gene *ScEPG1* do BAC SHCRBa_237_A10 não aparenta ter sido expresso (Fig. 25A). Tal fato possivelmente se deve a presença de um TE dentro da sequência codificante (Fig. 24C).

O presente trabalho teve por objetivo a identificação e sequenciamento de genes chave na degradação de parede celular durante a formação do aerênquima em raízes de cana-de-açúcar. A seleção destes genes se deu por meio de experimentos que levaram à obtenção de resultados referentes à influência da disponibilidade de nutrientes e do balanço hormonal na formação do aerênquima em raízes de cana-de-açúcar, caracterizando mais amplamente a formação constitutiva. O sequenciamento de BACs hom(e)ólogos carregando os genes de interesse forneceu informações a respeito da estruturação destes *loci* em relação a *S. bicolor*. A reconstrução filogenética empregando blocos sintênicos e regiões promotoras apontou sequências possivelmente oriundas da contribuição diferencial dos genomas parentais. A confirmação da contribuição de cada parental sugerida pela filogenia será validada posteriormente. No entanto, mesmo diante das diferenças estruturais observadas nos *loci* de interesse, parece não haver um padrão de expressão diferencial entre os hom(e)ólogos.

Já do ponto de vista fisiológico, dados oriundos deste trabalho sugerem que os genes relacionados à morte celular programada e à degradação de parede celular são expressos em resposta a um balanço hormonal mais complexo do que anteriormente previsto. Dentre estes genes destacam-se aqueles que codificam para a endopoligalacturonase e α -arabinofuranosidase que, de acordo com a análise comparativa realizada neste trabalho, são referentes a enzimas que atuam no ataque a homogalacturonano no caso da primeira e possivelmente arabinoxilano no caso da segunda. Além disso, foi observado que o gene referente à endopoligalacturonase é negativamente regulado pelo fator de transcrição da sub-família RAV, que apresenta um padrão de expressão oposto à do gene regulado.

Em conjunto, os dados obtidos sugerem que nos segmentos de raiz mais apicais, RAV é mais expresso, o que parece refletir nos menores níveis de transcritos da endopoligalacturonase. Nesta região as concentrações de etileno parecem favorecer a expressão deste gene. Conforme os níveis de etileno decaem, o mesmo ocorre aos transcritos de RAV. Por conseguinte, aqueles referentes à endopoligalacturonase se acumulam quando a proporção entre auxina e etileno tende a se estabilizar em 50%, concomitantemente à elevação dos transcritos da arabinofuranosidase. Uma vez produzidas as enzimas correspondentes, acredita-se que ambas atuem de forma sinérgica. Sugere-se que a hidrólise da pectina presente na lamela média no terceiro segmento leve à diminuição da adesão celular, com aumento da porosidade do tecido e expansão celular. Este último evento foi possivelmente já iniciado em S1, visto que a expressão de expansinas é mais marcante no primeiro segmento de raiz.

A perturbação hormonal decorrente da aplicação de 1-MCP que levou ao aumento da expressão de expansinas e arabinofuranosidase sugere que tais genes respondam à auxina. Na ausência deste inibidor, os níveis normais de auxina (inferiores em relação aos observados quando da aplicação de 1-MCP) em ação combinatória a concentrações mais elevadas de etileno possivelmente leva à menor atividade transcricional destes genes. A coordenação da expressão destes genes executada pela auxina pode ser parte componente da formação do aerênquima, porém, diante do desbalanço hormonal provocado pelo 1-MCP, os níveis de genes regulados por ambos os hormônios foram desacoplados, desorganizando as etapas iniciais de separação e expansão celular.

A manutenção da proporção auxina:etileno em torno de 40-50% após S3 parece desfavorecer a expressão do gene da endopoligalacturonase, enquanto a expressão daquele referente arabinofuranosidase aumenta, sofrendo regulação positiva pela auxina. Assim, em um tecido cortical previamente submetido à perda de adesão celular e afrouxamento de parede celular, a ação de arabinofuranosidases é facilitada. Estas parece ter importante papel na desramificação do arabinosilano, uma das etapas iniciais da formação do compósito provavelmente formado por xilano e celulose cristalina, constituindo o agregado de polissacarídeos altamente recalcitrantes que compõem os remanescentes de parede celular após a formação do aerênquima.

Figura 42. Modelo de formação do aerênquima com foco na regulação hormonal por trás das etapas de separação e expansão celular e morte celular programada. Setas pretas e linhas com barras horizontais indicam respectivamente regulação positiva e negativa observada neste trabalho. Numerais indicam a porcentagem de auxina em relação à etileno nos segmentos de raiz. As dimensões das formas geométricas cinza, azul e vermelha ilustram a área ocupada pelo aerênquima no córtex da raiz e a abundância de etileno e auxina, respectivamente. O tamanho da fonte referente à RAV e EPG reflete a abundância relativa dos mesmos ao longo da formação do aerênquima e a espessura do *loop* repressor indica a intensidade com a qual RAV reprime EPG.

Por fim, visto que a caracterização enzimática de glicosil hidrolases de plantas ainda está aquém do seu potencial biotecnológico, análise comparativa das sequências preditas destas enzimas forneceu dados que sugerem a viabilidade de seu uso no pré-tratamento biológico. Experimentos que visam validar as predições dos genes correspondentes ainda serão realizados. Mediante expressão heteróloga, visa-se a ampla caracterização funcional destas hidrolases, bem como a investigação de aspectos relevantes da interação enzima-substrato.

REFERÊNCIAS BIBLIOGRÁFICAS

- Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2003. *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. **Plant Cell** 15:63-78
- Abeles FB, Morgan PW, Saltveit ME. 1992. **Ethylene in Plant Biology**. 2 ed. San Diego Academic. 414 pp.
- Abiko T, Obara M. 2014. Enhancement of porosity and aerenchyma formation in nitrogen-deficient rice roots. **Plant Science** 215-216,76-83.
- Aguilar EA, Turner DW, Sivasithamparam K. 1999. Aerenchyma formation in roots of four banana (*Musa* spp.) cultivars. **Scientia Horticulturae** 80, 57-72.
- Agustí J, Merelo P, Cercós M, Tadeo FR & Talón M. 2008. Ethylene-induced differential gene expression during abscission of citrus leaves. **Journal of Experimental Botany** 59 (10), 2717–2733.
- Aharoni A, Keizer LCP, Van Den Broeck HC, Blanco-Portales R, Muñoz-Blanco J, Bois G, Smit P, De Vos RCH, O'Connell AP. 2002. Novel insight into vascular, stress, and auxin-dependent and -independent gene expression programs in strawberry, a non-climacteric fruit. **Plant Physiology** 129 (3), 1019-1031.
- Allen MD, Yamasaki K, Ohme-Takagi M, Tateno M, Suzuki M. 1998. A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. **Journal of the European Molecular Biology Organization** 17 (18), 5484–5496.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. **Journal of Molecular Biology** 215 (3), 403–410.
- Arikado H, Adachi Y. 1955. Anatomical and ecological responses of barley and some forage crops to the flooding treatment. **Bulletin Faculty Agriculture** 11, 1–29.
- Armand S, Wagemaker MJ, Sánchez-Torres P, Kester HC, van Santen Y, Dijkstra BW, Visser J, Benen JA. 2000. The active site topology of *Aspergillus niger* endopolygalacturonase II as studied by site-directed mutagenesis. **Journal of Biology and Chemistry** 275 (1), 691-696.
- Atwell BJ, Drew MC, Jackson MB. 1988. The influence of oxygen deficiency on ethylene synthesis, l-aminocyclopropane-1-carboxylic acid levels and aerenchyma formation in roots of *Zea mays* L. **Physiologia Plantarum** 72 (1), 15-22.
- Bailey-Serres J, Chang R. 2005. Sensing and signalling in response to oxygen deprivation in plants and other organisms. **Annals of Botany** 96 (4), 507-518.
- Balbi V, Lomax TL. 2003. Regulation of early tomato fruit development by the diageotropica gene. **Plant Physiology** 131(1), 186–197.
- Ballas N, Wong LM, Theologis A. 1993. Identification of the auxin-responsive element, AuxRE, in the primary indoleacetic acid-inducible gene, PS-IAA4/5, of pea (*Pisum sativum*). **Journal of Molecular Biology** 233, 580-596.
- Beers EP. 1997. Programmed cell death during plant growth and development. **Cell Death & Differentiation** 4, 649-661.
- Begum MK, Alam MR, Islam MS. 2013. Adaptive mechanisms of sugarcane genotypes under flood stress condition. **World Journal of Agricultural Sciences** 1 (2), 56-64.

- Bennetzen JL. 2007. Patterns in grass genome evolution. **Current Opinion in Plant Biology** 10 (2), 176-181.
- Bewley JD. 1997. Seed Germination and Dormancy. **Plant Cell** 9 (7), 1055–1066.
- Beyer EM, Morgan PW. 1971. The role of ethylene modification of auxin transport1. **Plant Physiology** 48, 208-212.
- Beyer EM. 1976. A potent inhibitor of ethylene action in plants. **Plant Physiology** 58, 268-271.
- Beyer EM. 1976. Silver ion: a potent anti-ethylene agent in cucumber and tomato. **HortScience** 11 (3), 175-196.
- Bissing DR. 1974. Haupt's gelatin adhesive mixed with formal in for affixing paraffin sections to slides. **Stain Technology** 49 (2), 116-117.
- Blakeship SM, Dole JM. 2003. 1-Methylcyclopropene: a review. **Postharvest Biology and Technology** 28 (1), 1-25.
- Bleecker AB, Jeffrey JE, Hall AE, Rodriguez F, Binder BM. 1998. The ethylene-receptor family from *Arabidopsis*: structure and function. **Philosophical Transactions of the Royal Society** 353, 1405-1412
- Böttcher C, Keyzers RA, Boss PK, Davies C. 2010. Sequestration of auxin by the indole-3-acetic acid-amido synthetase GH3-1 in grape berry (*Vitis vinifera* L.) and the proposed role of auxin conjugation during ripening. **Journal of Experimental Botany** 61 (13), 3615-3625.
- Bouranis DL, Chorianopoulou SN, Kollias C, Maniou P, Protonotarios VE, Siyiannis V, Hawkesford MJ. 2006. Dynamics of aerenchyma distribution in the cortex of sulfate-deprived adventitious roots of maize. **Annals of Botany** 97:695-704.
- Bouranis DL, Chorianopoulou SN, Siyiannis VF, Protonotarios VE, Hawkesford MJ. 2003. Aerenchyma formation in roots of maize during sulphate starvation. **Planta** 217, 382-391.
- Bragina TV, Martinovich LI, Rodionova NA, Bezborodov AM, Grineva GM. 2001. Ethylene-induced activation of xylanase in adventitious roots of maize as a response to the stress effect of root submersion. **Applied Biochemistry and Microbiology** 37(6):722-725
- Bragina TV, Rodionova NV, Grineva GM. 2003. Ethylene production and activation of hydrolytic enzymes during acclimation of maize seedlings to partial flooding. **Russian Journal of Plant Physiology** 50 (6): 794-798.
- Brummel DA, Cin VD, Crisosto CH, Labavitch JM. 2004. Cell wall metabolism during maturation, ripening and senescence of peach fruit. **Journal of Experimental Botany** 55 (405), 2029-2039.
- Brummell DA. 2006. Cell wall disassembly in ripening fruit. **Functional Plant Biology** 33, 103–119.
- Buckeridge MS, dos Santos WD, Tiné MAS, De Souza AP. 2015. The cell wall architecture of sugarcane and its implications to cell wall recalcitrance. In Press. In: Lam, E. **Compendium of Bioenergy Crops for Science Publishers**, CRC Press of Taylor and Francis Group.
- Buckeridge MS, Rayon C, Urbanowicz B, Tiné MAS, Carpita NC. 2004. Mixed linkage (1→3),(1→4)-β-d-glucans of grasses. **Cereal Chemistry** 81 (1), 115–127.

- Caffall KH, Mohnen D. 2009. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. **Carbohydrate Research** 344 (14), 1879-1900.
- Campbell R and Drew MC. 1983. Election microscopy of gas space (aerenchyma) formation in adventitious roots of *Zea mays* L. subjected to oxygen shortage. **Planta** 157 (4), 350-357.
- Cao Y, Song F, Goodman RM, Zheng Z. 2006. Molecular characterization of four rice genes encoding ethylene-responsive transcriptional factors and their expressions in response to biotic and abiotic stress. **Journal of Plant Physiology** 163 (11), 1167-1178.
- Cao Y, Song F, Goodman RM, Zheng Z. 2006. Molecular characterization of four rice genes encoding ethylene-responsive transcriptional factors and their expressions in response to biotic and abiotic stress. **Journal of Plant Physiology** 163(11), 1167-1178.
- Carpita N, McCann M. 2000. The Cell Wall. In: **Biochemistry & Molecular Biology of Plants**, Buchanan B, Gruissem W, Jones R, Eds. Poole, UK: Orca Book Services and American Society of Plant Physiologists. 1408pp.
- Carpita N, Tierney M, Campbell M. 2001. Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics. **Plant Molecular Biology** 47(1-2):1-5.
- Carpita NC, Gibeaut DM. 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. **The Plant Journal** 3 (1), 1-30.
- Causier B, Ashworth M, Guo W, Davies B. 2012. The TOPLESS interactome: a framework for gene repression in *Arabidopsis*. **Plant Physiology** 158, 423-438.
- Chakravarthy S, Tuori RP, D'Ascenzo MD, Fobert PR, Despres C, Martin GB. 2003. The tomato transcription factor *pti4* regulates defense-related gene expression via GCC box and non-GCC box cis elements. **Plant Cell** 15 (12): 3033-3050.
- Chávez-Montes RA, Ranocha P, Martinez Y, Minic Z, Jouanin L, Marquis M, Saulnier L, Fulton LM, Cobbett CS, Bitton F, Renou JP, Jauneau A, Goffner D. 2008. Cell wall modifications in *Arabidopsis* plants with altered α-L-arabinofuranosidase activity. **Plant Physiology** 147: 63-77.
- Christianson JA, Llewellyn DJ, Dennis ES, Wilson IW. 2010. Global gene expression responses to waterlogging in roots and leaves of cotton (*Gossypium hirsutum* L.). **Plant Cell Physiology** 51(1): 21-37
- Chung MY, Vrebalov J, Alba R, Lee JM, McQuinn R, Chung JD. Klein P, Giovannoni J. 2010. A tomato (*Solanum lycopersicum*) APETALA2/ERF gene, SIAP2a, is a negative regulator of fruit ripening. **Plant Journal** 64 (6), 936-947.
- Colmer TD, Cox MCH, Voisenek LACJ. 2006. Root aeration in rice (*Oryza sativa*): evaluation of oxygen, carbon dioxide, and ethylene as possible regulators of root acclimatizations. **New Phytologist** 170 (4), 767-778.
- Colmer TD. 2002. Aerenchyma and an inducible barrier to radial oxygen loss facilitate root aeration in upland, paddy and deep-water rice (*Oryza sativa* L.). **Annals of Botany** 91 (2), 301-309.

- Cormack J, Peterson PA. 1994. Gene marker loss induced by the transposable element, En, in maize. **Genetics** 136 (3), 1151-1156.
- D'Hont A. 2005. Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. **Cytogenetic and Genome Research** 109 (1-3), 27-33.
- Dal-Bianco M, Carneiro MS, Hotta CT, Chapola RG, Hoffmann HP, Garcia AAF, Souza GM. 2012. Sugarcane improvement: how far can we go? **Current Opinion in Biotechnology** 23 (2), 265-270.
- Dantas BF, Aragão CA & Alves JD. 2001. Cálcio e o desenvolvimento de aerênquimas e atividade de celulase em plântulas de milho submetidas à hipóxia. **Scientia Agricola** 58 (2), 251-257.
- Davies C, Boss PK, Robinson SB. 1997. Treatment of crape berries, a nonclimacteric fruit with a synthetic auxin, retards ripening and alters the expression of developmentally regulated genes. **Plant Physiology** 115: 1155-1161.
- De Setta N, Monteiro-Vitorello CB, Metcalfe CJ, Cruz GM, Del Bem LE, Vicentini R, Nogueira FT, Campos RA, Nunes SL, Turrini PC, Vieira AP, Ochoa Cruz EA, Corrêa TC, Hotta CT, de Mello Varani A, Vautrin S, da Trindade AS, de Mendonça Vilela M, Lembke CG, Sato PM, de Andrade RF, Nishiyama MY Jr, Cardoso-Silva CB, Scortecci KC, Garcia AA, Carneiro MS, Kim C, Paterson AH, Bergès H, D'Hont A, de Souza AP, Souza GM, Vincentz M, Kitajima JP, Van Sluys MA. 2014. Building the sugarcane genome for biotechnology and identifying evolutionary trends. **BMC Genomics** 30;15:540.
- De Souza AP, Grandis A, Leite DCC, Buckeridge MS. 2014. Sugarcane as a bioenergy source: history, performance, and perspectives for second-generation bioethanol. **BioEnergy Research** 7 (1), 24-35.
- De Souza AP, Leite DCC, Pattathil S, Hahn MG, Buckeridge MS. 2013. Composition and structure of sugarcane cell wall polysaccharides: implications for second-generation bioethanol production. **BioEnergy Research** 6 (2), 564-579.
- Debeche T, Bliard C, Debeire P, O'Donohue MJ. 2002. Probing the catalytically essential residues of the α -L-arabinofuranosidase from *Thermobacillus xylanilyticus*. **Protein Engineering** 15 (1), 21-28.
- Dennis ES, Dolferus R, Ellis M, Rahman M, Wu Y, Hoeren FU, Grover A, Ismond KP, Good AG, Peacock WJ. 2000. Molecular strategies for improving waterlogging tolerance in plants. **Journal of Experimental Botany** 51(342), 89-97.
- Dordas C1, Hasinoff BB, Igamberdiev AU, Manac'h N, Rivoal J, Hill RD. 2003. Expression of a stress-induced hemoglobin affects NO levels produced by alfalfa root cultures under hypoxic stress. **Plant Journal**. 35 (6), 763-770.
- Drew MC, He CJ, Morgan PW. 1989. Decreased ethylene biosynthesis, and induction of aerenchyma, by nitrogen- or phosphate-starvation in adventitious roots of *Zea mays* L. **Plant Physiology** 91 (1), 266-271.
- Drew MC, He CJ, Morgan PW. 2000. Programmed cell death and aerenchyma formation in roots. **Trends in Plant Science** 5 (3), 123-127.

REFERÊNCIAS BIBLIOGRÁFICAS

- Drew MC, Jackson MB, Giffard SC, Campbell R. 1981. Inhibition by silver ions of gas space (aerenchyma) formation in adventitious roots of *Zea mays* L. subjected to exogenous ethylene onto oxygen deficiency. **Planta** 153 (3), 217-224.
- Drew MC. 1997. Oxygen deficiency and root metabolism: injury and acclimation under hypoxia and anoxia. **Annual Review in Plant Physiology and Plant Molecular Biology** 48, 223-250.
- Drew MC¹, Jackson MB, Giffard S. 1979. Ethylene-promoted adventitious rooting and development of cortical air spaces (aerenchyma) in roots may be adaptive responses to flooding in *Zea mays* L. **Planta**. 147 (1), 83-88.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research** 32 (5), 1792-1797.
- El-Sharkawy I, Sherif S, Mila I, Bouzayen M, Jarasankar S. 2009. Molecular characterization of seven genes encoding ethylene-responsive transcriptional factors during plum fruit development and ripening. **Journal of Experimental Botany** 60(3), 907-922.
- Evans, DE. 2003. Aerenchyma formation. **New Phytologist** 161 (1), 35–49.
- Fagerstedt KV. Programmed cell death and aerenchyma formation under hypoxia. 2010. In: Mancuso S, Shabala S (eds), **Waterlogging signalling and tolerance in plants**, Heidelberg: Springer Science, 99-118.
- Fan M, Zhi J, Richards C, Brown K & Lynch JP. 2003. Physiological roles for aerenchyma in phosphorus-stressed roots. **Functional Plant Biology** 30 (5), 493-506.
- Federici L, Caprari C, Mattei B, Savino C, Di Matteo A, De Lorenzo G, Cervone F, Tsernoglou D. 2001 Structural requirements of endopolygalacturonase for the interaction with PGIP (polygalacturonase-inhibiting protein). **Proceedings of the National Academy of Sciences USA** 98 (23), 13425-13430.
- Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M. 2014. The Pfam protein families database. **Nucleic Acids Research** 42, D222-D230.
- Fischer RL. 1991. Role of cell wall hydrolases in fruit ripening. **Annual Reviews in Plant Molecular Biology** 42, 675-703.
- Fujimoto SY, Ohta M, Usui A, Shinshi H, Ohme-Takagi M. 2000. *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. **Plant Cell** 12(3), 393-404.
- Fujimoto SY, Ohta M, Usui A, Shinshi H, Ohme-Takagi M. 2000. *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. **Plant Cell** 12 (3), 393-404.
- Fukaki H, Tasaka M. 2009. Hormone interactions during lateral root formation. **Plant Molecular Biology** 69, 437–449.

- Fukuda A, Tanaka Y. 2006. Effects of ABA, auxin, and gibberellin on the expression of genes for vacuolar H⁺-inorganic pyrophosphatase, H⁺-ATPase subunit A, and Na⁺/H⁺ antiporter in barley. **Plant Physiology and Biochemistry** 44 (5-6), 351-358.
- Fulton LM, Cobbett CS. 2003. Two α -L-arabinofuranosidase genes in *Arabidopsis thaliana* are differentially expressed during vegetative growth and flower development. **Journal of Experimental Botany** 54, 2467–2477.
- Garthwaite AJ, von Bothmer R, Colmer TD. Diversity in root aeration traits associated with waterlogging tolerance in the genus *Hordeum*. **Functional Plant Biology** 30 (8), 875-889.
- Geigenberger P. 2003 Response of plant metabolism to too little oxygen. **Current Opinion in Plant Biology** 6 (3), 247-256
- Gilbert RA, Rainbolt CR, Morris DR, Bennett AC. 2007. Morphological responses of sugarcane to long-term flooding. **Agronomy Journal** 99 (6), 1622–1628.
- Gilead S, Shoham Y. 1995. Purification and characterization of α -L-arabinofuranosidase from *Bacillus stearothermophilus* T-6. **Applied Environmental Microbiology** 61(1), 170-174.
- Gimble FS. 2000. Invasion of a multitude of genetic niches by mobile endonuclease genes. **Federation of European Microbiological Societies Letters** 185, 99–107.
- Giovannoni JJ. 2004. Genetic regulation of fruit development and ripening. **Plant Cell** 16: 170-180.
- Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM. 1992. Isolation of the *Arabidopsis* ABI3 gene by positional cloning. **Plant Cell** 4 (10), 1251-1261.
- Given NK, Venis MA, Gierson D. 1988. Hormonal regulation of ripening in the strawberry, a non-climacteric fruit. **Planta** 174(3), 402-406.
- Glaz B, Morrisa DR, Daroubb SH. Sugarcane photosynthesis, transpiration, and stomatal conductance due to flooding and water table. **Crop Science** 44 (5), 1633-1641.
- Gomez LD, Steele-King CG, McQueen-Mason SJ. 2008. Sustainable liquid biofuels from biomass: the writing's on the walls. **New Phytologist** 178 (3), 473–485.
- González-Carranza ZH, Elliott KA, Roberts JA. 2007. Expression of polygalacturonases and evidence to support their role during cell separation processes in *Arabidopsis thaliana*. **Journal of Experimental Botany** 58 (13), 3719-3730.
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS. 2012. Phytozome: a comparative platform for green plant genomics. **Nucleic Acids Research** 40 (D1), D1178-D1186.
- Grandis A, De Souza AP, Tavares EQP, Buckeridge MS. 2014. Using natural plant cell wall degradation mechanisms to improve second generation bioethanol. In: McCann MC, Buckeridge MS, Carpita NC, eds. **Plants and bioenergy**. New York: Springer, 211–230.
- Grandis, A. 2015. Papel das enzimas de degradação da parede celular na formação do aerênquima em raízes de cana-de-açúcar. **Tese (Doutorado)**. Instituto de Biociências da Universidade de São Paulo. Departamento de Botânica. 122p.

REFERÊNCIAS BIBLIOGRÁFICAS

- Gunawardena A, Lan H, Pearce DM, Jackson MB, Hawes CR, Evans DE. 2001a. Characterization of programmed cell death during aerenchyma formation induced by ethylene or hypoxia in roots of maize (*Zea mays* L.). **Planta** 212 (2), 205-214.
- Gunawardena AHLAN, Pearce DME, Jackson MB, Hawes CR, Evans DE. 2001b. Rapid changes in cell wall pectic polysaccharides are closely associated with early stages of aerenchyma formation, a spatially localized form of programmed cell death in roots of maize (*Zea mays* L.) promoted by ethylene. **Plant, Cell & Environment** 24 (12), 1369–1375.
- Hall B, Shakeel S, Schaller G. 2007. Ethylene receptors: ethylene perception and signaling transduction. **Journal of Plant Growth Regulation** 26 (2):118-130.
- Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. **Nucleic Acids Symposium** 41, 95-98.
- Harper JF, Breton G & Harmon A. 2004. Decoding Ca²⁺ Signals through plant protein kinases. **Annual Reviews in Plant Biology** 55, 263–288.
- Hatsugai N, Kuroyanagi M, Nishimura M & Hara-Nishimura I. 2006. A cellular suicide strategy of plants: vacuole-mediated cell death. **Apoptosis** 11, 905-911.
- He CJ, Drew M, Morgan PW. 1994. Induction of enzymes associated with lysigenous aerenchyma formation in roots of *Zea mays* during hypoxia or nitrogen starvation. **Plant Physiology** 105:861-865
- He CJ, Morgan P, Drew MC. 1992. Enhanced sensitivity to ethylene in nitrogen- of phosphate-starved roots of *Zea mays* L. during aerenchyma formation. **Journal of Plant Physiology**. 98:137-142.
- He CJ, Morgan PW, Drew MC. 1996a. Transduction of an ethylene signal is required for cell death and lysis in the root cortex of maize during aerenchyma formation induced by hypoxia. **Journal of Plant Physiology** 112: 463-472.
- He CJ, Pinlayson SA, Drew MC, Jordan WR, Morgan PW. 1996b. Ethylene biosynthesis during aerenchyma formation in roots of *Zea mays* subjected to mechanical impedance and hypoxia. **Journal of Plant Physiology** 112: 1679-1685.
- Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. **Genome Biology** 8(2), R19.
- Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant cis-acting regulatory DNA elements (PLACE) database. **Nucleic Acids Research** 27 (1), 297-300.
- Himmel, M. E. 2009. **Biomass recalcitrance: deconstructing the plant cell wall for bioenergy**. Oxford, United Kingdom: Wiley-Blackwell. 505 pp.
- Hinz M, Wilson IW, Yang J, Buerstenbinder K, Llewellyn D, Dennis ES, Sauter M, Dolferus R. 2010. *Arabidopsis* RAP2.2: an ethylene response transcription factor that is important for hypoxia survival. **Plant Physiology** 153 (2), 757-772.

REFERÊNCIAS BIBLIOGRÁFICAS

- Hiraiwa N, Nishimura M, Hara-Nishimura I. 1999. Vacuolar processing enzyme is self-catalytically activated by sequential removal of the C-terminal and N-terminal propeptides. **Federation of European Biochemical Societies Letters** 447 (2-3), 213–216.
- Ho MS, Tsai PI, Chien CT. 2006. F-box proteins: the key to protein degradation. **Journal of Biomedical Sciences** 13, 181-191.
- Hoagland DR, Arnon DI. 1950. **The water-culture method for growing plants without soil**. Berkeley, University of California, College of Agriculture, Agricultural Experiment Station. 31p.
- Hoeren FU, Dolferus R, Wu Y, Peacock WJ & Dennis ES. 1998. Evidence for a role for AtMYB2 in the induction of the *Arabidopsis* alcohol dehydrogenase gene (ADH1) by low oxygen. **Genetics** 149 (2), 479-490.
- Holm S. 1979. A simple sequentially rejective multiple test procedure. **Scandinavian Journal of Statistics** 6 (2), 65–70.
- Hövel K, Shallom D, Niefind K, Belakhov V, Shoham G, Baasov T, Shoham Y, Schomburg D. 2003. Crystal structure and snapshots along the reaction pathway of a family 51 α -L-arabinofuranosidase. **Journal of the European Molecular Biology Organization** 22 (19), 4922-4932.
- Hu YX, Wang YX, Liu XF, Li JY. 2004. *Arabidopsis* RAV1 is down-regulated by brassinosteroid and may act as a negative regulator during plant development. **Cell Research** 14 (1), 8-15.
- Hutchison KW, Singer PB, McInnis S, Diaz-Sala C, Greenwood MS. expansins are conserved in conifers and expressed in hypocotyls in response to exogenous auxin1. **Plant Physiology** 120 (3), 827–832.
- Igamberdiev AU, Baron K, Manac'h-Little N, Stoimenova M, Hill RD. 2005. The haemoglobin/nitric oxide cycle: involvement in flooding stress and effects on hormone signalling. **Annals of Botany** 96 (4), 557-564.
- Ikeda M, Ohme-Takagi M. 2009. A novel group of transcriptional repressors in *Arabidopsis*. **Plant Cell Physiology** 50 (5), 970-975.
- Ikeda M, Ohme-Takagi M. 2009. A novel group of transcriptional repressors in *Arabidopsis*. **Plant Cell Physiology** 50 (5), 970-975.
- Im DH, Kimura KI, Hayasaka F, Tanaka T, Noguchi M, Kobayashi A, Shoda S, Miyazaki K, Wakagi T, Fushinobu S. 2012. Crystal structures of glycoside hydrolase family 51 α -L-arabinofuranosidase from *Thermotoga maritima*. **Bioscience, Biotechnology and Biochemistry** 76, 423-428
- In BC, Strable J, Binder BM, Falbel TG, Patterson SE. 2013. Morphological and molecular characterization of ethylene binding inhibition in carnations. **Postharvest Biology and Technology** 86, 272–279.
- International Energy Agency. Disponível em < <http://www.iea.org/> >. Acessado em 10 jan. 2015.
- International Rice Genome Sequencing Project. 2005. The Map-Based Sequence of the Rice Genome. **Nature** 436, 793–780.
- IPCC. 2014. Climate change 2014: Impacts, adaptation, and vulnerability. Working Group II Contribution to the IPCC 5th Assessment Report. Disponível em < <http://www.ipcc.ch/> >. Acessado em 05 jan. 2015.

- Iqbal N, Trivellini A, Masood A, Ferrante A, Khan NA. 2013. Current understanding on ethylene signaling in plants: the influence of nutrient availability. **Plant Physiology and Biochemistry** 73, 128-138.
- Iskandar HM, Simpson RS, Casu RE, Bonnett GD, MacLean DJ, Manners JM. 2004. Comparison of reference genes for quantitative real-time polymerase chain reaction analysis of gene expression in sugarcane. **Plant Molecular Biology Reporter** 22: 325-337.
- Ismail IO. 2004. Function and regulation of xylem cysteine protease 1 and xylem cysteine protease 2 in *Arabidopsis*. **Tese (Doutorado)**. Virginia Polytechnic Institute and State University. 100 pp.
- Itai A, Ishihara K, Bewley JD. 2003. Characterization of expression, and cloning, of beta-D-xylosidase and alpha-L-arabinofuranosidase in developing and ripening tomato (*Lycopersicon esculentum* Mill.) fruit. **Journal of Experimental Botany** 54 (393), 2615-2622.
- Jackson MB, Armstrong W. 1999. Formation of aerenchyma and the processes of plant ventilation in relation to soil flooding and submergence. **Plant Biology** 1 (3), 274-287.
- Jackson MB, Fenning TM, Drew MC, Saker LR. 1985. Stimulation of ethylene production and gas-space (aerenchyma) formation in adventitious-roots of *Zea mays* L. by small partial pressures of oxygen. **Planta** 165 (4), 486-492.
- Jenkins ES, Paul W, Craze M, Whitelaw CA, Weigand A, Roberts JA. 1999. Dehiscence-related expression of an *Arabidopsis thaliana* gene encoding a polygalacturonase in transgenic plants of *Brassica napus*. **Plant Cell & Environment** 22 (2), 159-167.
- Jiang YM, Joyce DC, Macnish AJ. 1999. Responses of banana fruit to treatment with 1-methylcyclopropene. **Plant Growth Regulation** 28 (2), 77-82.
- Jin LG, Liu JY. 2008. Molecular cloning, expression profile and promoter analysis of a novel ethylene responsive transcription factor gene GhERF4 from cotton (*Gossypium hirsutum*). **Plant Physiology and Biochemistry** 46 (1), 46-53.
- Jofuku KD, den Boer BGW, Van Montagu M, Okamuro JK. 1994. Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. **Plant Cell** 6 (9), 1211-1225.
- Jones AM. 2001. Programmed cell death in development and defense. **Plant Physiology** 125 (1), 94-97.
- Jones B, Frasse P, Olmos E, Zegzouti H, Li ZG, Latché A, Pech JC, Bouzayen M. 2002. Down-regulation of DR12, an auxin-response-factor homolog, in the tomato results in a pleiotropic phenotype including dark green and blotchy ripening fruit. **The Plant Journal** 32 (4), 603-613.
- Joo JH, Bae YS, Lee JS. 2001. Role of auxin-induced reactive oxygen species in root gravitropism. **Plant Physiology** 126 (3), 1055-1060.
- Jung JKH, McCouch S. 2013. Getting to the roots of it: Genetic and hormonal control of root architecture **Frontiers Plant Science** 4, 186.
- Justin SHF, Armstrong W. 1991a. A reassessment of the influence of NAA on aerenchyma formation in maize roots. **New Phytologist** 117 (4), 607-618.
- Justin SHF, Armstrong W. 1991b. Evidence for the involvement of ethene in aerenchyma formation in adventitious roots of rice (*Oryza sativa* L.). **New Phytologist** 118 (1), 49-62.

- Kagaya Y, Ohmiya K, Hattori T. 1999. RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. **Nucleic Acids Research** 27 (2), 470–478.
- Karakurt Y, Huber DJ. 2003. Activities of several membrane and cell-wall hydrolases, ethylene biosynthetic enzymes, and cell wall polyuronide degradation during low-temperature storage of intact and fresh-cut papaya (*Carica papaya*) fruit. **Postharvest Biology and Technology** 28 (2), 219-229.
- Karimi M, Inzé D, Depicker A. 2002. GATEWAY vectors for Agrobacterium-mediated plant transformation. **Trends Plant Science** 7 (5), 193-195.
- Kawai M, Samarajeewa PK, Barrero RA, Nishiguchi M, Uchimiya H. 1998. Cellular dissection of the degradation pattern of cortical cell death during aerenchyma formation of rice roots. **Planta** 204 (3), 277-287.
- Kieber JJ. 1997. The Ethylene Response Pathway In *Arabidopsis*. **Annual Review of Plant Physiology and Plant Molecular Biology** 48, 277-296.
- Kim HJ, Lynch JP, Brown KM. 2008. Ethylene insensitivity impedes a subset of responses to phosphorus deficiency in tomato and petunia. **Plant Cell & Environment** 31 (12), 1744-1755.
- Kim J, Shiu S-H, Thoma S, Li W-H, Patterson SE. 2006. Patterns of expansion and expression divergence in the plant polygalacturonase gene family. **Genome Biology** 7, R87.
- Klok EJ, Wilson IW, Wilson D, Chapman SC, Ewing RM, Somerville SC, Peacock WJ, Dolferus R, Dennis ES. 2002. Expression profile analysis of the low-oxygen response in *Arabidopsis* root cultures. **The Plant Cell** 14 (10), 2481-2494.
- Konings H, Verschuren G. 1980. Formation of aerenchyma in roots of *Zea mays* in aerated solutions, and its relation to nutrient supply. **Physiologia Plantarum** 49 (3), 265-270.
- Konings H. 1982. Ethylene-promoted formation of aerenchyma in seedling roots of *Zea mays* L. under aerated and nonaerated conditions. **Physiologia Plantarum** 54 (2), 119-124.
- Kosugi S., Hasebe M., Tomita M., and Yanagawa H. 2009. Systematic identification of yeast cell cycle-dependent nucleocytoplasmic shuttling proteins by prediction of composite motifs. **Proceedings of the National Academy of Sciences USA** 106 (25), 10171-10176.
- Kozlova LV, Ageeva MV, Ibragimova NN, Gorshkova TA. 2014. Arrangement of mixed-linkage glucan and glucuronoarabinoxylan in the cell walls of growing maize roots. **Annals of Botany** 114 (6), 1135–1145.
- Kraus JE, Arduin M. 1997. **Manual básico de métodos em morfologia vegetal**. Seropédica: Edur. 198 p.
- Laidlaw HK1, Lahnstein J, Burton RA, Fincher GB, Jobling SA. 2012. Analysis of the arabinoxylan arabinofuranohydrolase gene family in barley does not support their involvement in the remodelling of endosperm cell walls during development. **Journal of Experimental Botany** 63 (8), 3031-3045.
- Lam E. 2005. Vacuolar proteases livening up programmed cell death. **Trends in Cell Biology** 15 (3), 124-127.
- Lazarova G, Zeng Y, Kermode AR. 2002. Cloning and expression of an ABSCISIC ACID-INSENSITIVE 3 (ABI3) gene homologue of yellow-cedar. **Journal of Experimental Botany** 53 (371), 1219-1221.

- Lee RC, Burton RA, Hrmova M, Fincher GB. 2001. Barley arabinoxylan arabinofuranohydrolases: purification, characterization and determination of primary structures from cDNA clones. **Biochemistry Journal** 356(Pt 1):181-189.
- Lee SF, Forsberg CW. 1987. Purification and characterization of an α -L-arabinofuranosidase from *Clostridium acetobutylicum* ATCC 824. **Canadian Journal of Microbiology** 33 (11), 1011-1016.
- Lee TG, Jang CS, Kim JY, Kim DS, Park JH, Kim DY & Seo YW. 2007. A Myb transcription factor (TaMyb1) from wheat root is expressed during hypoxia: roles in response to the oxygen concentration in root environment and abiotic stresses. **Physiologia Plantarum** 129 (2), 375-385.
- Leite, DCC. 2013. Modificações da parede celular durante a formação de aerênquima em raízes de cana-de-açúcar. **Dissertação (Mestrado)**. Instituto de Biociências da Universidade de São Paulo. Departamento de Botânica. 106 pp.
- Lembke CG, Nishiyama Jr MY, Sato PM, de Andrade RF, Souza GM. 2012. Identification of sense and antisense transcripts regulated by drought in sugarcane. **Plant Molecular Biology** 79(4-5), 461-477.
- Lenochová Z, Soukup A, Votrubová O. 2009. Aerenchyma formation in maize roots. **Biologia Plantarum** 53 (2), 263–270.
- Li CW, Su RC, Cheng CP, Sanjaya, You SJ, Hsieh TH, Chao TC, Chan MT. 2011. Tomato RAV transcription factor is a pivotal modulator involved in the AP2/EREBP-mediated defense pathway. **Plant Physiology** 156 (1), 213–227.
- Li J, Yang H, Peer WA, Richter G, Blakeslee J, Bandyopadhyay A, Titapiwantakun B, Undurraga S, Khodakovskaya M, Richards EL, Krizek B, Murphy AS, Gilroy S, Gaxiola R. 2005. *Arabidopsis* H⁺-PPase AVP1 regulates auxin-mediated organ development. **Science**. 310 (5745), 121-125.
- Li J, Yuan R. 2008. NAA and ethylene regulate expression of genes related to ethylene biosynthesis, perception, and cell wall degradation during fruit abscission and ripening in 'delicious' apples. **Journal of Plant Growth Regulation** 27, 283–295.
- Lieth H. 1975. Productivity of the Biosphere Ecological Studies. In: **Ecological Studies and Synthesis**. Eds. Lieth H, Whittaker RH. New York: Springer, vol 14. 339 pp.
- Lima DU, Santos HP, Tiné MA, Molle FRD, Buckeridge MS. 2001. Patterns of expression of cell wall related genes in sugarcane. **Genetics and Molecular Biology** 24(1-4), 191-198.
- Liu F, van Toai T, Moy LP, Bock G, Linford LD, Quackenbush J. 2005. Global transcription profiling reveals comprehensive insights into hypoxic response in *Arabidopsis*. **Plant Physiology** 137 (3), 1115-1129.
- Liu K, Kang BC, Jiang H, Moore SL, Li H, Watkins CB, Setter TL, Jahn MM. 2005. A GH3-like gene, CcGH3, isolated from *Capsicum chinense* L. fruit is regulated by auxin and ethylene. **Plant Molecular Biology** 58 (4), 447-464.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The Carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research** 42, D490–D495.
- Lu Y, Wassmann R, Neue HU, Huang C. 1999. Impact of phosphorus supply on root exudation, aerenchyma formation and methane emission of rice plants. **Biogeochemistry** 47 (2), 203–218.

- Ludwig-Müller J, Georgiev M, Bley T. 2008. Metabolite and hormonal status of hairy root cultures of Devil's claw (*Harpagophytum procumbens*) in flasks and in a bubble column bioreactor. **Process Biochemistry** 43 (1), 15-23.
- Lynch JP, Brown KM. 1997. Ethylene and plant responses to nutritional stress. **Physiologia Plantarum** 100, 613-619.
- Lynch JP, Brown KM. 2001. Topsoil foraging – an architectural adaptation of plants to low phosphorus availability. **Plant and Soil** 237 (2), 225-237.
- Ma Z, Baskin TI, Brown KM, Lynch JP. 2003. Regulation of root elongation under phosphorus stress involves changes in ethylene responsiveness. **Plant Physiology** 131 (3), 1381–1390.
- Macrelli S, Galbe M, Wallberg O. 2014. Effects of production and market factors on ethanol profitability for an integrated first and second generation ethanol plant using the whole sugarcane as feedstock. **Biotechnology for Biofuels** 7, 26.
- Magnani E, Sjölander K, Hake S. 2004 From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. **Plant Cell** 16 (9), 2265-22677.
- Malik AI, Colmer TD, Lambers H, Schortemeyer M. 2003. Aerenchyma formation and radial O₂ loss along adventitious roots of wheat with only the apical root portion exposed to O₂ deficiency. **Plant Cell & Environment** 26 (10), 1713-1722.
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. **Nucleic Acids Research** 43 (D), 222-226.
- Margeot A, Hahn-Hagerdal B, Edlund M, Slade R, Monot F. 2009. New improvements for lignocellulosic ethanol. **Current Opinion in Biotechnology** 20 (3), 372-380.
- McCann MC, Carpita NC. 2008. Designing the deconstruction of plant cell walls. **Current Opinion in Plant Biology** 11 (3), 314–320.
- McQueen-Mason SJ & Cosgrove DJ. 1995. Expansin mode of action on cell walls: analysis of wall hydrolysis, stress relaxation, and binding. **Plant Physiology** 107, 87-100.
- Michael TP. 2014. Plant genome size variation: bloating and purging DNA. **Briefings in Functional Genomics** 13, 308-317.
- Micheli F. 2001. Pectin methylesterases: cell wall enzymes with important roles in plant physiology. **Trends in Plant Science** 6 (9), 414–419.
- Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, Van Breusegem F. 2011. ROS signaling: the new wave? **Trends in Plant Science** 16 (6), 300-309.
- Mohnen D. 2008. Pectin structure and biosynthesis. **Current Opinion in Plant Biology** 11 (3), 266-277.
- Moshe Sagi, Robert Fluhr. 2006. Production of reactive oxygen species by plant NADPH oxidases. **Plant Physiology** 141 (2), 336–340.

- Muday GK, Rahman A, Binder BM. 2012. Auxin and ethylene: collaborators or competitors? **Trends in Plant Science** 17 (4), 181-195.
- Müller K, Linkies A, Leubner-Metzger G, Kermode AR. 2012. Role of a respiratory burst oxidase of *Lepidium sativum* (cress) seedlings in root development and auxin signalling. **Journal of Experimental Botany** 63 (18), 6325-6334.
- Nagarajan VK, Smith AP. 2012. Ethylene's role in phosphate starvation signaling: more than just a root growth regulator. **Plant Cell Physiology** 53 (2), 277-286.
- Nakano T, Suzuki K, Fujimura T & Shinshi H. 2006. Genome-Wide Analysis of the ERF Gene Family in *Arabidopsis* and Rice. **Plant Physiology** 140 (2), 411-432.
- Nanjo Y, Maruyama K, Yasue H, Yamaguchi-Shinozaki K, Shinozaki K & Komatsu S. 2011. Transcriptional responses to flooding stress in roots including hypocotyl of soybean seedlings. **Plant Molecular Biology** 77 (1-2), 129-144.
- Navas J, Béguin P. 1992. Site-directed mutagenesis of conserved residues of *Clostridium thermocellum* endoglucanase CelC. **Biochemical and Biophysical Research Communications** 189 (2), 807-812.
- Nishiyama K, Guis M, Rose JK, Kubo Y, Bennett KA, Wangjin L, Kato K, Ushijima K, Nakano R, Inaba A, Bouzayen M, Latche A, Pech JC, Bennett AB. 2007. Ethylene regulation of fruit softening and cell wall disassembly in *Charentais melon*. **Journal of Experimental Botany** 58 (6), 1281-1290.
- O'Donoghue EM, Huber DJ, Timpa JD, Erdos GW, Brecht JK. 1994. Influence of avocado (*Persea americana*) Cx cellulase on the structural features of avocado cellulose. **Planta** 194 (4), 573-584.
- Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S. 2003. Gibberellin biosynthesis and response during *Arabidopsis* seed germination. **Plant Cell** 15, 1591-1604.
- Ohme-Takagi M, Shinshi H. 1995. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. **Plant Cell** 7(2), 173-182.
- Ohme-Takagi M, Suzuki K, Shinshi H. 2000. Regulation of ethylene-induced transcription of defense genes. **Plant Cell and Physiology** 41 (11), 1187-1192
- Okamuro JK, Caster B, Villarroel R, van Montagu M, Jofuku KD. 1997. The AP2 domain of APETALA2 defines a large new family of DNA binding proteins in *Arabidopsis*. **Proceedings of the National Academy of Sciences USA** 94, 7076-7081.
- Oñate-Sánchez L, Anderson JP, Young J & Singh K. 2007. AtERF14, a member of the ERF family of transcription factors, plays a nonredundant role in plant defense. **Plant Physiology** 143 (1), 400-409.
- Paes G., Skov LK, O'Donoghue MJ, Remond C, Kastrup JS, Gajhede M, Mirza O. 2008. The structure of the complex between a branched pentasaccharide and *Thermobacillus xylanilyticus* GH-51 arabinofuranosidase reveals xylan-binding determinants and induced fit. **Biochemistry** 47, 7441-7451
- Pagnussat GC, Alandete-Saez M, Bowman JL, Sundaresan V. 2009. Auxin-dependent patterning and gamete specification in the *Arabidopsis* female gametophyte. **Science** 324 (5935), 1684-1689.
- Paterson AH1, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J,

REFERÊNCIAS BIBLIOGRÁFICAS

- Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otilar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob-ur-Rahman, Ware D, Westhoff P, Mayer KF, Messing J, Rokhsar DS. 2009. The Sorghum bicolor genome and the diversification of grasses. **Nature** 29; 457 (7229), 551-556.
- Pauly M, Keegstra K. 2008. Cell-wall carbohydrates and their modification as a resource for biofuels. **Plant Journal** 54 (4), 559–568.
- Pech JC, Bouzayen M, Latché A. 2008. Climacteric fruit ripening: Ethylene-dependent and independent regulation of ripening pathways in melon fruit. **Plant Science** 175 (1-2), 114-120.
- Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods** 8 (10), 785-786.
- Petersen TN, Kauppinen S, Larsen S. 1997. The crystal structure of rhamnogalacturonase A from *Aspergillus aculeatus*: a right-handed parallel beta helix. **Structure** 5(4), 533-544.
- Pickersgill R, Smith D, Worboys K, Jenkins J. 1998. Crystal structure of polygalacturonase from *Erwinia carotovora* ssp. *carotovora*. **Journal of Biological Chemistry** 273(38), 24660-24664.
- Postma JA, Lynch JP. 2011a. Theoretical evidence for the functional benefit of root cortical aerenchyma in soils with low phosphorus availability. **Annals of Botany** 107: 829–841.
- Postma JA, Lynch JP. 2011b Root cortical aerenchyma enhances the growth of maize on soils with suboptimal availability of nitrogen, phosphorus and potassium. **Journal of Plant Physiology** 156 (3), 1190-1201.
- Prestridge DS. 1991. SIGNAL SCAN: A computer program that scans DNA sequences for eukaryotic transcriptional elements. **Computer Applications in the Biosciences** 7 (2), 203-206.
- Promkhambut A, Polthanee A, Akkasaeng C, Younger A. 2011. Growth, yield and aerenchyma formation of sweet and multipurpose sorghum (*Sorghum bicolor* L. Moench) as affected by flooding at different growth stages. **Australian Journal of Crop Science** 5(8), 954-965.
- Rajhi I, Yamauchi T, Takahashi H, Nishiuchi S, Shiono K, Watanabe R, Mliki A, Nagamura Y, Tsutsumi N, Nishizawa NK, Nakazono M. 2011. Identification of genes expressed in maize root cortical cells during lysigenous aerenchyma formation using laser microdissection and microarray analyses. **New Phytologist** 190 (2), 351-368.
- Raven PH, Evert RF, Eichhorn SE. 2007. **Biologia Vegetal**, 7a. ed. Rio de Janeiro: Guanabara Koogan. 850pp.
- Renewable Fuels Association. Disponível em < <http://www.ethanolrfa.org/> >. Acessado em 05 jan. 2015.
- Renewable Fuels Standard. Disponível em < <http://www.epa.gov/oms/fuels/renewablefuels/> >. Acessado em 05 jan. 2015.
- Riechmann JL, Meyerowitz EM. 1998. The AP2/EREBP family of plant transcription factors. **Journal of Biological Chemistry** 379 (6), 633-646.

- Robert LS, Allard S, Gerster JL, Cass L, Simmonds J. 1993. Isolation and characterization of a polygalacturonase gene highly expressed in *Brassica napus* pollen. **Plant Molecular Biology** 23, 1273–1278.
- Roberts JA, Elliot KA, Gonzalez-Carranza Z. 2002. Abscission, dehiscence and other cell separation processes. **Annual Reviews in Plant Biology** 53, 131–158.
- Rose JKC, Bennett AB. 1999. Cooperative disassembly of the cellulose-xyloglucan network of plant cell walls: parallels between cell expansion and fruit ripening. **Trends in Plant Science** 4 (5).
- Rose JKC, Catala C, Gonzalez-Carranza ZH, Roberts JA. 2003. Cell wall disassembly. **Annual Plant Reviews** 8, 264–324.
- Rosli HG, Civello PM, Martínez GA. 2009. α -L-Arabinofuranosidase from strawberry fruit: cloning of three cDNAs, characterization of their expression and analysis of enzymatic activity in cultivars with contrasting firmness. **Plant Physiology and Biochemistry** 47(4), 272-281.
- Růžicka K, Ljung K, Vanneste S, Podhorská R, Beeckman T, Friml J, Benková E. 2007. Ethylene regulates root growth through effects on auxin biosynthesis and transport-dependent auxin distribution. **Plant Cell** 19 (7), 2197-2212.
- Ryttersgaard C, Leggio LL, Coutinho PM, Henrissat B, Larsen S. 2002. *Aspergillus aculeatus* β -1,4-galactanase: substrate recognition and relations to other glycoside hydrolases in clan GH-A. **Biochemistry** 41 (51), 15135–15143.
- Rzewuski G & Sauter M. 2008. Ethylene biosynthesis and signaling in rice. **Plant Science** 175 (1-2), 32-42.
- Saab IN & Sachs MM. 1996. A flooding-induced xyloglucan endo-transglycosylase homolog in maize is responsive to ethylene and associated with aerenchyma. **Plant Physiology** 112 (1), 385-391.
- Sachs JA. 1882. **A text book of botany**. Oxford, UK: Oxford University Press.
- Sakon J, Adney WS, Himmel ME, Thomas SR, Karplus PA. 1996. Crystal structure of thermostable family 5 endocellulase E1 from *Acidothermus cellulolyticus* in complex with cellotetraose. **Biochemistry** 35 (33), 10648–10660.
- Scavetta RD, Herron SR, Hotchkiss AT, Kita N, Keen NT, Benen JA, Kester HC, Visser J, Jurnak F. 1999. Structure of a plant cell wall fragment complexed to pectate lyase C. **Plant Cell** 11(6),1081-1092.
- Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S,

- Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddeloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK. 2009. The B73 maize genome: complexity, diversity, and dynamics. **Science** 20; 326 (5956),1112-1125.
- Sekine D, Munemura I, Gao M, Mitsuhashi W, Toyomasu T, Murayama H. 2006. Cloning of cDNAs encoding cell-wall hydrolases from pear (*Pyrus communis*) fruit and their involvement in fruit softening and development of melting texture. **Physiologia Plantarum** 126 (2), 163–174.
- Sharoni AM, Nuruzzaman M, Satoh K, Shimizu T, Kondoh H, Sasaya T, Choi IR, Omura T, Kikuchi S. 2011. Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. **Plant Cell and Physiology** 52 (2), 344-360.
- Shimada TL, Shimada T, Hara-Nishimura I. 2010. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. **Plant Journal** 61 (3), 519-528.
- Shimizu T, Nakatsu T, Miyairi K, Okuno T, Kato H. 2002. Active-site architecture of endopolygalacturonase I from *Stereum purpureum* revealed by crystal structures in native and ligand-bound forms at atomic resolution. **Biochemistry** 41(21), 6651-6659.
- Shiono K, Takahashi H, Colmer TD, Nakazono M. 2008. Role of ethylene in acclimations to promote oxygen transport in roots of plants in waterlogged soils. **Plant Science** 175: 52–58
- Shussler E, Longstreth DJ.1996. Aerenchyma develops by cell lysis in roots and cell separation in leaf petioles in *Sagittaria lancifolia* (Alismataceae). **American Journal of Botany** 83 (1-2), 1266–1273.
- Silva-Cardenas RI, Ricard B, Saglio P, Hill RD. 2003. Hemoglobin and hypoxic acclimation in maize root tips. **Russian Journal of Plant Physiology** 50 (6), 821–826.
- Siyiannis VF, Protonotarios VE, Zechmann B, Chorianopoulou SN, Müller M, Hawkesford MJ, Bouranis DL. 2011. Comparative spatiotemporal analysis of root aerenchyma formation processes in maize due to sulphate, nitrate or phosphate deprivation. **Protoplasma** 249 (3), 671-686.
- Skowrya D, Craig KL, Tyers M, Elledge SJ, Harper JW. 1997. F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. **Cell** 91, 209-219.
- Smith DM, Inman-Bamber NG, Thorburn PJ. 2005. Growth and function of the sugarcane root system. **Field Crops Research** 92 (2-3), 169-183.
- Sohn KH1, Lee SC, Jung HW, Hong JK, Hwang BK. 2006. Expression and functional roles of the pepper pathogen-induced transcription factor RAV1 in bacterial disease resistance, and drought and salt stress tolerance. **Plant Molecular Biology** 61 (6), 897-915.

REFERÊNCIAS BIBLIOGRÁFICAS

- Solano R, Stepanova A, Chao Q, Ecker J. 1998. Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. **Genes & Development** 12 (23), 3703-3714.
- Solomon M, Belenghi B, Delledonne M, Menachem E, Levine A. 1999. The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. **Plant Cell** 11 (3), 431-444.
- Souza TA, Santos CR., Souza AR., Oldiges DP, Ruller R., Prade RA, Squina FM, Murakami MT. 2011. Structure of a novel thermostable GH51 α -L-arabinofuranosidase from *Thermotoga petrophila* RKU-1. **Protein Science** 20, 1632-1637.
- Sozzi GO, Fraschina AA, Navarro AA, Cascone O, Greve LC, Labatovich JM. 2002. α -L-Arabinofuranosidase activity during development and ripening of normal and ACC synthase antisense tomato fruit. **Hortscience** 37 (3), 564–566.
- Stepanova AN, Alonso JM. 2009. Ethylene signaling and response: where different regulatory modules meet. **Current Opinion in Plant Biology** 12 (5), 548-555.
- Stepanova AN, Hoyt JM, Hamilton AA, Alonso JM. 2005. A link between ethylene and auxin uncovered by the characterization of two root-specific ethylene-insensitive mutants in *Arabidopsis*. **Plant Cell** 17, 2230 – 2242.
- Stotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. **Nature Reviews Genetics** 8, 272–285.
- Subbaiah CC, Kollipara KP, Sachs MM. 2000. A Ca^{2+} -dependent cysteine protease is associated with anoxia-induced root tip death in maize. **Journal of Experimental Botany** 51 (345), 721-730.
- Subbaiah CC, Sachs MM. 2003. Molecular and cellular adaptations of maize to flooding stress. **Annals of Botany** 91 (2), 119-127.
- Suzuki M, Ketterling MG, Li QB, McCarty DR. 2003. Viviparous1 alters global gene expression patterns through regulation of abscisic acid signaling. **Plant Physiology** 132 (3), 1664-1677.
- Tacken E, Ireland H, Gunaseelan K, Karunairetnam S, Wang D, Schultz K, Bowen J, Atkinson RG, Johnston JW, Putterill J, Hellens RP, Schaffer RJ. 2010. The role of ethylene and cold temperature in the regulation of the apple POLYGALACTURONASE1 gene and fruit softening. **Plant Physiology**. 153 (1), 294–305.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Molecular Biology and Evolution** 28 (10), 2731–2739.
- Tavares EQP, De Souza AP, Buckeridge MS. 2015. How endogenous plant cell-wall degradation mechanisms can help achieve higher efficiency in saccharification of biomass. **Journal of Experimental Botany** 66. doi:10.1093/jxb/erv171.
- Tenenbaum DJ. 2008. Food vs. Fuel: Diversion of Crops Could Cause More Hunger. **Environmental Health Perspective** 116(6), 254-257.

- Tetsushi H, Karim MA. 2007. Flooding tolerance of sugarcane in relation to growth, physiology and root structure. **South Pacific Studies** 28 (1), 9-22.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research** 22(22), 4673–4680.
- Tiné MA, Cortelazzo AL, Buckeridge MS. 2000. Xyloglucan mobilisation in cotyledons of developing plantlets of *Hymenaea courbaril* L. (Leguminosae-Caesalpinoideae). **Plant Science** 154 (2), 117–126.
- Tomkins JP, Yu Y, Miller-Smith H, Frisch DA, Woo SS, Wing RA. 1999. A bacterial artificial chromosome library for sugarcane. **Theoretical and Applied Genetics** 99 (3-4), 419-424.
- Trainotti L, Tadiello A, Casadoro G. 2007. The involvement of auxin in the ripening of climacteric fruits comes of age: the hormone plays a role of its own and has an intense interplay with ethylene in ripening peaches. **Journal of Experimental Botany** 58 (12), 3299-3308.
- Trépant H, Kombargi R, Hoteit L, Nordahl H. 2009. the upstream survivors of 2009: how to live to tell the tale. Booz&Co. Disponível em < http://www.booz.com/media/uploads/Upstream_Survivors_2009.pdf >. Acessado em 20 jan. 2015.
- Tripathi SK, Singh AP, Sane AP & Nath P. 2009. Transcriptional activation of a 37 kDa ethylene responsive cysteine protease gene, RbCP1, is associated with protein degradation during petal abscission in rose. **Journal of Experimental Botany** 60 (7), 2035-2044.
- Trought MCT, Drew MC. 1980. The development of waterlogging damage in young wheat plants in anaerobic solution cultures. **Journal of Experimental Botany** 31 (6), 1573–1580.
- União da Indústria da Cana-de-Açúcar. Disponível em < <http://www.unicadata.com.br/> >. Acessado em 07 mar. 2015.
- van Pouderooyen 1, Snijder HJ, Benen JA, Dijkstra BW. 2003. Structural insights into the processivity of endopolygalacturonase I from *Aspergillus niger*. **Federation of the Societies of Biochemistry and Molecular Biology Letters** 554 (3), 462-466.
- van Santen Y, Benen JA, Schröter KH, Kalk KH, Armand S, Visser J, Dijkstra BW. 1999. 1.68-Å crystal structure of endopolygalacturonase II from *Aspergillus niger* and identification of active site residues by site-directed mutagenesis. **Journal of Biological Chemistry** 274(43), 30474-30480.
- Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W. 2010. Lignin Biosynthesis and Structure. **Plant Physiology** 153 (3), 895-905.
- Virolainen E, Blokhina O, Fagerstedt KV. 2002. Ca²⁺-induced high amplitude swelling and cytochrome c release from wheat (*Triticum aestivum* L.) mitochondria under anoxic stress. **Annals of Botany** 90 (4), 509-516.
- Visser EJ, Bögemann GM. 2006. Aerenchyma formation in the wetland plant *Juncus effusus* is independent of ethylene. **New Phytologist** 171 (2), 305-314.
- Voesenek LACJ, Colmer TD, Pierik R, Millenaar FF, Peeters AJM. 2006. How plants cope with complete submergence. **New Phytologist** 170 (2), 213–226.

- Wang A, Tan D, Takahashi A, Li TZ, Harada T. 2007. MdERFs, two ethylene-response factors involved in apple fruit ripening. **Journal of Experimental Botany** 58 (13), 3743-3748.
- Wang ZY1, MacRae EA, Wright MA, Bolitho KM, Ross GS, Atkinson RG. 2000. Polygalacturonase gene expression in kiwifruit: relationship to fruit softening and ethylene production. **Plant Molecular Biology** 42 (2) 317-328.
- Wang, Tan D, Takahashi A, Li TZ, Harada T. 2007. MdERFs, two ethylene-response factors involved in apple fruit ripening. **Journal of Experimental Botany** 58 (13), 3743-3748.
- Weigel D. 1995. The genetics of flower development: from floral induction to ovule morphogenesis. **Annual Reviews in Genetics** 29, 19-39.
- Woo HR, Kim JH, Kim J, Kim J, Lee U, Song IJ, Kim JH, Lee HY, Nam HG, Lim PK. 2010. The RAV1 transcription factor positively regulates leaf senescence in *Arabidopsis*. **Journal of Experimental Botany** 61 (14), 3947-3957.
- Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ. 2006. Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. **Nature** 442, 705-708.
- Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, Seki E, Matsuda T, Tomo Y, Hayami N, Terada T, Shirouzu M, Osanai T, Tanaka A, Seki M, Shinozaki K, Yokoyama S. 2004. Solution structure of the B3 DNA binding domain of the *Arabidopsis* cold-responsive transcription factor RAV1. **Plant Cell** 16 (12), 3448-3459.
- Yamauchi T, Rajhi I, Nakazono M. 2011. Lysigenous aerenchyma formation in maize root is confined to cortical cells by regulation of genes related to generation and scavenging of reactive oxygen species. **Plant Signaling & Behavior** 6 (5), 759-761.
- Yamauchi T, Shimamura S, Nakazono M, Mochizuki T. 2013. Aerenchyma formation in crop species: a review. **Field Crops Research** 152, 8–16.
- Yang CY, Hsu FC, Li JP, Wang NN, Shih MC. 2011 The AP2/ERF transcription factor AtERF73/HRE1 modulates ethylene responses during hypoxia in *Arabidopsis*. **Plant Physiology** 156 (1), 202-212.
- Yang X, Li Y, Ren B, Ding L, Gao C, Shen Q, Guo S. 2012. Drought-induced root aerenchyma formation restrains water uptake in nitrate-supplied rice seedlings. **Plant Cell Physiology** doi: 10.1093/pcp/pcs003.
- Yanofsky MF. 1995. Floral meristems to floral organs: Genes controlling early events in *Arabidopsis* flower development. **Annual Reviews on Plant Physiol and Plant Molecular Biology** 46, 167–188.
- Yin X, Allan AC, Chen K, Ferguson IB. 2010. Kiwifruit EIL and ERF genes involved in regulating fruit ripening. **Plant Physiology** 153 (3), 1280–1292.
- Yu YB, Yang SF. 1979. Auxin-induced ethylene production and its inhibition by aminoethoxyvinylglycine and cobalt ion. **Plant Physiology** 64 (6), 1074–1077.
- Yukiyoshi K, Karahara I. 2014. Role of ethylene signalling in the formation of constitutive aerenchyma in primary roots of rice. **AoB PLANTS** 6, plu043

REFERÊNCIAS BIBLIOGRÁFICAS

- Zhang J, Yu C, Krishnaswamy L, Peterson T. 2011. Transposable elements as catalysts for chromosome rearrangements. **Methods in Molecular Biology** 701, 315-326.
- Zhang Z, Zou X, Tang W, Zheng Y. 2006. Revelation on early response and molecular mechanism of submergence tolerance in maize roots by microarray and suppression subtractive hybridization. **Environmental and Experimental Botany** 58 (1-3), 53-63.
- Zhao L, Luo Q, Yang C, Han Y, Li W 2008 A RAV-like transcription factor controls photosynthesis and senescence in soybean. **Planta** 227 (6), 1389-1399.

Anexo 1. Genes diferencialmente expressos no segmento 1 em resposta a 1-MCP.

Transcript	Function category	Structural category	Category	Sub category 1	Expression*
SCBGLR1002B09.g	Unknown	no matches	.	.	1.2609
SCEQRT1024C07.g	Transport	kinase/phosphatase	.	Pyrophosphate-energized vacuolar membrane proton pump; Pyrophosphate-energized inorganic pyrophosphatase; H(+)-PPase; proton translocating pyrophosphatase [<i>O. sativa</i>]	2.2483
SCEQHR1082B01.g	Carbohydrate metabolism	several pathways (several putative substrates)	Cell wall	Glycosyl hydrolase family 1; hypothetical protein OsI_030804 [<i>O. sativa</i> (indica cultivar-group)]	1.3234
SCJFRZ2012A12.g	Development	uncharacterized	.	Cortical cell-delineating protein precursor (Root-specific protein ZRP3); unknown [<i>Z. mays</i>]	1.5283
SCEQRT3018G04.g	Cell wall metabolism	Lignin metabolism	Cell wall	dirigent-like protein; Os10g0335000 [<i>O. sativa</i> (japonica cultivar-group)] Os10g0335000 [<i>O. sativa</i> (japonica cultivar-group)]	-1.2535
SCACRZ3036E03.g	Unknown	no matches	No matches	.	2.8444
SCBFRT3093E07.g	.	.	Secondary metabolism	anthocyaninless2 [<i>Z. mays</i>] Leucoanthocyanidin dioxygenase; LDOX; Leucocyanidin oxygenase; Leucoanthocyanidin hydroxylase A2 [<i>Z. mays</i>]	-2.5337
SCEQLB1063D10.g	Unknown	uncharacterized	.	Protein RUPTURED POLLEN GRAIN 1; unknown [<i>Z. mays</i>]	3.3119
SCJFRT1007E01.g	.	.	Stress	dioxygenase; IDS2	3.1613
SCVPRT2084A03.g	.	.	.	unknown [<i>Z. mays</i>]	2.2613
SCCART1C04G12.g	.	.	.	hypothetical protein OsJ_000657 [<i>O. sativa</i> (japonica cultivar-group)]	2.4181
SCEQRT2091H09.g	Unknown	.	.	unknown [<i>Z. mays</i>]	1.3938
SCJFRZ2034H02.g	Protein metabolism	zinc-binding protein	.	"RING-H2 finger protein ATL5B;"; Os02g0106600 [<i>O. sativa</i> (japonica cultivar-group)] zinc finger-like [<i>O. sativa</i> Japonica Group] Os02g0106600 [<i>O. sativa</i> (japonica cultivar-group)]	1.4942
SCEQRT1029F09.g	AminoAcid and nitrogen metabolism	decarboxylase	Amino acid metabolism	Glutamate decarboxylase (EC 4.1.1.15); unknown [<i>Z. mays</i>]	1.9153
SCCSD1003E06.g	Others	inhibitor	Pathogenicity	Protease inhibitor; thaumatin	-2.2697
SCBFRZ2048F01.g	Unknown	uncharacterized	.	Ser/Thr-rich protein T10 in DGCR region; unknown [<i>Z. mays</i>]	1.6617
SCQSHR1021B05.g	Unknown	Uncharacterized	.	Putative protein TPRXL; unknown [<i>Z. mays</i>]	1.4809

Anexo 1. Continuação

SCEQRT1033G03.g	.	.	.	Chemocyanin precursor (Basic blue protein) (Plantacyanin); hypothetical protein Osl_022834 [<i>O. sativa</i> (indica cultivar-group)]	2.7350
SCRLFL4025G12.g	Unknown	uncharacterized	.	unknown [<i>Z. mays</i>]	1.2649
SCJLRT1018C09.g	.	.	.	hypothetical protein LOC100192619 [<i>Z. mays</i>] unknown [<i>Z. mays</i>]	1.4718
SCCART2002G08.g	Transport	transporter	.	Protein RUPTURED POLLEN GRAIN 1; hypothetical protein LOC100192602 [<i>Z. mays</i>] unknown [<i>Z. mays</i>]	1.2450
SCEPRT2043E12.g	Transport	transporter	Transporters	Sulfate transporters (Sulfate_transp); Os03g0161200 [<i>O. sativa</i> (japonica cultivar-group)] Sulfate transporter 3.1, putative, expressed [<i>O. sativa</i> (japonica cultivar-group)] Os03g0161200 [<i>O. sativa</i> (japonica cultivar-group)]	1.2543
SCQSLR1061C11.g	DNA metabolism	uncharacterized	.	DNA-damage-inducible protein F; unknown [<i>Z. mays</i>]	2.3321
SCCART2C08E12.g	Others	others	.	UPF0497 membrane protein At4g03540; unknown [<i>Z. mays</i>]	1.4303
SCSBFL1038E08.g	.	.	.	hypothetical protein [<i>S. bicolor</i>]	2.2134
SCCCRZ1001A02.g	Unknown	uncharacterized	.	unknown [<i>Z. mays</i>]	1.7048
SCQSHR1024F04.g	Transport	Membrane transporters ABC transporter	.	Probable white-brown complex homolog protein 16; unknown [<i>Z. mays</i>]	1.5009
SCCCCL3003E07.b	Pathogen Resistance	induced by biotic and abiotic stress and hormone	Cell wall	chitinase; chitinase B [<i>S. officinarum</i>]	-3.0691
SCJFRT2060A10.g	Unknown	Uncharacterized	No matches	unknown [<i>Z. mays</i>]	1.2612
SCCART1002B11.g	Unknown	others	Carbohydrate metabolism	(Q40007) Putative 32.7 kDa jasmonate-induced protein; hypothetical protein Osl_035081 [<i>O. sativa</i> (indica cultivar-group)]	2.2446
SCQSRT2033A02.g	.	.	.	unknown [<i>Z. mays</i>]	1.4822
SCCART1002A07.g	.	.	.	Lysine histidine transporter 2 OS=A. <i>thaliana</i> ; unknown [<i>Z. mays</i>]	1.5632
SCCCCL2001H12.b	Unknown	others	.	hypothetical protein OsJ_010070 [<i>O. sativa</i> (japonica cultivar-group)]	-1.3359
SCJLST1019A04.g	Unknown	uncharacterized	.	.	3.2667

* Log₂ dos valores de expressão 1-MCP/controle.

Anexo 2. Genes diferencialmente expressos no segmento 2 em resposta a 1-MCP

Transcript	Function category	Structural category	Category	Sub category 1	Expression*
SCBFLR1083G04.g	Transposable element-related	.	.	Retrotransposable element Tf2 155 kDa protein type 3; polyprotein [<i>S. bicolor</i>]	-1.9150
SCMCRZ3068E07.g	.	.	.	unknown [<i>Z. mays</i>]	1.7650
SCEZRZ3098F04.g	Redox metabolism	Transferase family	.	Probable glutathione S-transferase BZ2; Protein bronze-2; bronze-2 protein [<i>Z. mays</i>]	-1.4550
SCBGLR1119H08.g	Protein metabolism	peptidase	.	Serine carboxypeptidase-like 35; hypothetical protein Osl_016140 [<i>O. sativa</i> (indica cultivar-group)]	1.2252
SCJLRT1022H01.g	.	.	.	Cycloartenol synthase; 2,3-epoxysqualene--cycloartenol cyclase; unknown [<i>Z. mays</i>]	1.1524
SCMCST1049H08.g	AminoAcid and nitrogen metabolism	synthase	.	Nicotianamine synthase 3 (S-adenosyl-L-methionine:S-adenosyl-L-methionine:S- adenosyl-methionine 3-amino-3-carboxypropyltransferase 3) (OsNAS3) Nicotianamine synthase 3 (S-adenosyl-L-methionine:S-adenosyl-L-methionine:S- adenosyl-methionine 3-amino-3; nicotianamine synthase 3 [<i>Z. mays</i>]	1.1351
SCEQRT1024C07.g	Transport	kinase/ phosphatase	.	Pyrophosphate-energized vacuolar membrane proton pump; Pyrophosphate-energized inorganic pyrophosphatase; H(+)-PPase; proton translocating pyrophosphatase [<i>O. sativa</i>]	3.2419
SCEPRT2043B01.g	Cell wall metabolism	Cell wall growth/extension	Cell wall	Expansin; unknown [<i>Z. mays</i>]	2.3962
SCRLB2032D09.g	Others	uncharacterized	.	Gamma-thionin homolog PPT precursor; unknown [<i>Z. mays</i>]	-1.0406
SCACRZ3036E03.g	Unknown	no matches	No matches	.	1.0551
SCBFRT3093E07.g	.	.	Secondary metabolism	Leucoanthocyanidin dioxygenase (LDOX); anthocyaninless2 [<i>Z. mays</i>] Leucoanthocyanidin dioxygenase; LDOX; Leucocyanidin oxygenase; Leucoanthocyanidin hydroxylase A2 [<i>Z. mays</i>]	-2.0683
SCCCLB1C04B12.g	Others	.	.	RING finger and CHY zinc finger domain-containing protein 1; Zinc finger protein 363; CH-rich-interacting match with PLAG1; Androgen receptor N-terminal-interacting protein; unknown [<i>Z. mays</i>]	1.3723

Anexo 2. Continuação

SCCCLB1C03G03.g	Unknown	uncharacterized	.	unknown [<i>Z. mays</i>]	1.3416
SCCCRT2004E10.g	.	.	.	Anther-specific proline-rich protein APG; Protein CEX; unknown [<i>Z. mays</i>]	1.5361
SCJFRT1007E01.g	.	.	Stress	dioxygenase; IDS2	1.2550
SCVPRT2084A03.g	.	.	.	unknown [<i>Z. mays</i>]	1.7558
SCEZHR1048C09.g	Others	uncharacterized	.	Vignain; Bean endopeptidase; Cysteine proteinase EP-C1; unknown [<i>Z. mays</i>]	1.6057
SCEQRT1025C04.g	Others	cytochrome P450 family protein	Stress	Cytochrome P450; Cyp only	1.3396
SCQGLR2032E05.g	Others	.	Development	development; ZmGR1a [<i>Z. mays</i>]	1.0489
SCCCCL7002A04.g	.	.	Transcription Factor	hormone-related/ethylene/AP2/EREBP; EREB-like protein	1.0489
SCCCRT1C04G12.g	.	.	.	hypothetical protein OsJ_000657 [<i>O. sativa</i> (japonica cultivar-group)]	3.1392
SCJFRT2060A05.g	.	.	.	O-methyltransferase ZRP4 (OMT); hypothetical protein LOC100193561 [<i>Z. mays</i>] unknown [<i>Z. mays</i>]	1.9452
SCEQRT2091C06.g	.	.	.	hypothetical protein OsJ_015414 [<i>O. sativa</i> (japonica cultivar-group)]	1.4825
SCVPRZ2042G03.g	Carbohydrate metabolism	several pathways (several putative substrates)	Cell wall	Glycosyl hydrolase family 1; unknown [<i>Z. mays</i>]	1.6650
SCVPRT2076E10.g	Signal Transduction	transcription factor	Transcription Factor	AP2/EREBP; hypothetical protein Osl_007303 [<i>O. sativa</i> (indica cultivar-group)]	1.5626
SCCCRZ3004F06.g	Unknown	uncharacterized	.	.	1.4858
SCRURT2006B11.g	Unknown	uncharacterized	.	.	2.4432
SCVPLB1020G05.g	Carbohydrate metabolism	Sacarose/trealose /sorbitol and other mono-oligosaccharides	Cell wall	α -L-arabinofuranosidase; unknown [<i>Z. mays</i>]	1.4916
SCMCRT2086H06.g	Unknown	uncharacterized	.	hypothetical protein LOC100192824 [<i>Z. mays</i>]; unknown [<i>Z. mays</i>]	1.2967
SCJLST1019A04.g	Unknown	uncharacterized	.	.	1.8897

Anexo 2. Continuação

SCVPLB1018H01.g	Others	uncharacterized	.	Protein regulator of cytokinesis 1; unknown [<i>Z. mays</i>]	-1.9669
SCVPLR2005A02.g	Unknown	.	.	.	-0.9790
SCEZLB1011H05.g	Unknown	uncharacterized	.	unknown [<i>Z. mays</i>]	1.3907
SCEZLB1006G08.g	Carbohydrate metabolism	uncharacterized	Carbohydrate metabolism	Putative glucan 1,3-beta-glucosidase; unknown [<i>Z. mays</i>]	1.7018
SCEPLR1030E06.g	Signal Transduction	kinase/ phosphatase	Receptors	receptor Ser/Thr kinase; caneRLK-D14 (LRR)	1.8632
SCVPRT2073A12.g	Unknown	uncharacterized	.	unknown [<i>Z. mays</i>]	1.3570
SCRFLR2038G09.g	Others	transferase	.	Thiosulfate sulfurtransferase; Rhodanese; Senescence-associated protein; Sulfurtransferase protein 16; AtStr16; unknown [<i>Z. mays</i>]	2.6769
SCCART2001D04.g	.	.	Development	Root-specific; RCc3	-2.0532

* Log₂ dos valores de expressão 1-MCP/controle.

Tabela 3. Genes diferencialmente expressos no segmento 3 em resposta a 1-MCP

Transcript	Function category	Structural category	Category	Sub category 1	Expression*
SCCCRT3001A02.g	Pathogen Resistance	induzida por stress biotico e abiotico e por hormônios	Cell wall	chitinase; Endochitinase A precursor (Seed chitinase A)	2.1000
SCEQRT2026A06.g	Cell wall metabolism	Cell wall growth/extension	Cell wall	Expansin; Putative selT-like protein F28H7.4 precursor	2.8501
SCQSRT2031H06.g	Signal Transduction	protein-protein interaction	Stress	Water stress; induced	-1.5884
SCEZRZ3098F04.g	Redox metabolism	Transferase family	.	Probable glutathione S-transferase BZ2; Protein bronze-2; bronze-2 protein [<i>Z. mays</i>]	-1.5089
SCJLRT3076A06.g	.	.	Unknown protein	H0410G08.14 [<i>O. sativa</i> (indica cultivar-group)] hypothetical protein OsI_016948 [<i>O. sativa</i> (indica cultivar-group)]	2.1912
SCEQRT1024C07.g	Transport	kinase/phosphatase	.	Pyrophosphate-energized vacuolar membrane proton pump; Pyrophosphate-energized inorganic pyrophosphatase; H(+)-PPase; proton translocating pyrophosphatase [<i>O. sativa</i>]	3.3795
SCQGLR1085F11.g	Signal Transduction	dehydrin	Stress	drought-induced; Dehydrin	-1.6809
SCJLRT1023E06.g	Cell wall metabolism	Other glycan degradation	Cell wall	beta-1,3-glucanase Glycosyl hydrolases family 17; pathogenesis-related protein 6 [<i>Z. mays</i> subsp. <i>parviglumis</i>]	1.5048
SCEQRT1024F02.g	Redox metabolism	peroxidase	Stress	response to oxidative stress; Peroxidase 1 precursor	-1.6100
SCJLFL4183G10.g	Unknown	uncharacterized	.	hypothetical protein OsI_014192 [<i>O. sativa</i> (indica cultivar-group)]	1.5436
SCCCRT1C04G12.g	.	.	.	hypothetical protein OsJ_000657 [<i>O. sativa</i> (japonica cultivar-group)]	3.6190
SCACAD1035G11.g	.	.	.	LOB domain-containing protein 29 (ASYMMETRIC LEAVES 2-like protein 16) (AS2-like protein 16); Os08g0563400 [<i>O. sativa</i> (japonica cultivar-group)] putative LOB domain protein 17 [<i>O. sativa</i> Japonica Group]	-1.0209
SCCCSD1003E06.g	Others	inhibitor	Pathogenicity	Protease inhibitor; thaumatin	2.3418

Anexo 3. Continuação

SCEQRT2101A06.g	Pathogen Resistance	.	.	Nectarin-1; Superoxide dismutase [Mn]; Os01g0952000 [<i>O. sativa</i> (japonica cultivar-group)] putative Rhicadhesin receptor precursor [<i>O. sativa</i> Japonica Group] Os01g0952000 [<i>O. sativa</i> (japonica cultivar-group)] hypothetical protein OsI_005095 [<i>O. sativa</i>]	1.1896
SCVPLB1017D05.g	Signal Transduction	Uncharacterized	Stress	induced	1.0642
SCVPRT2080B02.g	Cell wall metabolism	Structural proteins of the cell wall		hypothetical protein OsI_032717 [<i>O. sativa</i> (indica cultivar-group)]	2.2216
SCEZLB1012C07.g	Signal Transduction	kinase/phosphatase	Protein kinase	SNF-like kinase; caneCIPK-17	-1.3075
SCSGAM2103C12.g	.	.	.	HVA22-like protein f; AtHVA22f; hypothetical protein OsJ_003236 [<i>O. sativa</i> (japonica cultivar-group)]	1.2923
SCRFSD2023E07.g	Pathogen Resistance	peptidase	Pathogenicity	R-genes transduction; PR	1.4797
SCEPLR1030E06.g	Signal Transduction	kinase/phosphatase	Receptors	receptor Ser/Thr kinase; caneRLK-D14 (LRR)	1.0359
SCVPLB1018H01.g	Others	uncharacterized	.	Protein regulator of cytokinesis 1; unknown [<i>Z. mays</i>]	-1.6725
SCSBFL1038E08.g	.	.	.	hypothetical protein [<i>S. bicolor</i>]	5.3723
SCSGHR1069F04.b	Others	cytochrome P450	Stress	cytochrome P450; CYP71C4	2.3618
SCRURT2009D02.g	.	.	.	OSJNBa0091D06.16 [<i>O. sativa</i> (japonica cultivar-group)] hypothetical protein OsI_016227 [<i>O. sativa</i> (indica cultivar-group)] hypothetical protein OsJ_014957 [<i>O. sativa</i>]	1.1952
SCACHR1035G08.g	Others	uncharacterized	.	14 kDa proline-rich protein DC2.15; unknown [<i>Z. mays</i>]	2.2071
SCSGLR1045B04.g	Unknown	Uncharacterized	.	unknown [<i>Z. mays</i>]	-1.0467
SCSBSD1058E10.g	Unknown	uncharacterized	Receptors	Receptor Ser/Thr kinase-unclassified; with LRR	-1.0503
SCSFSD1065B09.g	.	.	.	Pyrophosphate-energized vacuolar membrane proton pump; Pyrophosphate-energized inorganic pyrophosphatase; H(+)-PPase; unknown [<i>Z. mays</i>]	-1.6482
SCVPRT2073C10.g	Others	uncharacterized	.	Wheatwin-2; Pathogenesis-related protein 4b; unknown [<i>Z. mays</i>]	1.9193

* Log₂ dos valores de expressão 1-MCP/controle.

Anexo 4. Genes diferencialmente expressos no segmento 4 em resposta a 1-MCP

Transcript	Function category	Structural category	Category	Sub category 1	Expression*
SCEQRT2026A06.g	Cell wall metabolism	Cell wall growth/extension	Cell wall	Expansin; Putative selT-like protein F28H7.4 precursor	1.1878
SCSGHR1072F11.g	Signal Transduction	transcription factor	Transcription Factor	Zinc finger proteins; GATA	1.5703
SCEQRT1024C07.g	Transport	kinase/ phosphatase	.	Pyrophosphate-energized vacuolar membrane proton pump; H(+)-PPase; [<i>O. sativa</i>]	3.5658
SCJFRT2057H11.g	.	.	.	unknown [<i>Z. mays</i>]	2.1804
SCVPRT2073F02.g	.	.	.	unknown [<i>Z. mays</i>]	1.8975
SCCART3004A10.g	.	.	No matches	.	1.9384
SCQGLB1029D12.g	Unknown	uncharacterized	.	unknown [<i>Z. mays</i>]	1.4169
SCJLFL4183G10.g	Unknown	uncharacterized	.	hypothetical protein Osl_014192 [<i>O. sativa</i>]	1.2228
SCVPRT2084A03.g	.	.	.	unknown [<i>Z. mays</i>]	0.8050
SCCART1C04G12.g	.	.	.	hypothetical protein OsJ_000657 [<i>O. sativa</i>]	4.3313
SCCCL3001F06.g	.	.	.	Sulfite oxidase; Moco-containing protein AtMCP; At-SO; AtSOX	1.2823
SCMCRT2086A02.g	Nucleotide metabolism	synthase	.	Phosphoribosylamine-glycine ligase, chloroplastic; Glycinamide ribonucleotide synthetase; GARS; Phosphoribosylglycinamide synthetase; hypothetical protein Osl_036507 [<i>O. sativa</i>]	2.2205
SCVPRT2080B02.g	Cell wall metabolism	Structural proteins of the cell wall	.	Os12g0442800 [<i>O. sativa</i> (japonica cultivar-group)] Sulfite oxidase, putative, expressed [<i>O. sativa</i> (japonica cultivar-group)] Os12g0442800 [<i>O. sativa</i> (japonica cultivar-group)]; hypothetical protein Osl_032717 [<i>O. sativa</i>]	2.3652
SCJLRT1015G08.g	.	.	No matches	.	1.0603
SCBGLR1023D10.g	Signal Transduction	transcription factor	Transcription Factor	Zinc finger proteins; GATA	1.4422
SCSGAM2103C12.g	.	.	.	HVA22-like protein f; AtHVA22f; hypothetical protein OsJ_003236 [<i>O. sativa</i> (japonica cultivar-group)]	1.1154
SCEZRT2019F10.g	Redox metabolism	peroxidase	.	Peroxidase 1; unknown [<i>Z. mays</i>]	1.2323
SCCCLR1072D05.g	Carbohydrate metabolism	translocator	.	Protein RUPTURED POLLEN GRAIN 1; unknown [<i>Z. mays</i>]	0.9374

Anexo 4. Continuação

SCBGFL5077B07.g	Unknown	.	.	.	1.1759
SCJLRZ1020B01.g	.	.	Nucleic acid metabolism	Putative prpol	1.6465
SCJFRT1059G06.g	Unknown	no matches	.	.	1.2712
SCEZLB1005E07.g	.	.	Lipid, fatty-acid and isoprenoid metabolism	Phospholipase D p1 (EC 3.1.4.4) (AtPLDp1) (Phospholipase D1 PHOX and PX containing domain); putative nitrate transporter [<i>O. sativa</i> Japonica Group]	1.4970
SCRLRT3032A09.g	Pathogen Resistance	induced by biotic and abiotic stress and hormone	Cell wall	chitinase; hypothetical protein Osl_001428 [<i>O. sativa</i> (indica cultivar-group)]	2.1352
SCEPAM2056G07.g	Redox metabolism	peroxidase	Stress	response to oxidative stress; Peroxidase 64 precursor (Atperox P64) (PRXR4) (ATP17a)	1.9397
SCEQRT2100C04.g	Cell wall metabolism	cell wall differentiation	Cell wall	Fucosyltransferase; hypothetical protein Osl_021353 [<i>O. sativa</i> (indica cultivar-group)]	1.8011
SCQGLR2032G02.g	Redox metabolism	peroxidaseCell wall	.	Peroxidase 15; Prx15; Anionic peroxidase; unknown [<i>Z. mays</i>] unknown [<i>Z. mays</i>]	2.4812
SCSGHR1069F04.b	Others	cytochrome P450 family proteinCell wall	Stress	cytochrome P450; CYP71C4	3.0488
SCACHR1035G08.g	Others	uncharacterized	.	14 kDa proline-rich protein DC2.15; unknown [<i>Z. mays</i>]	2.1422
SCRFLR2038G09.g	Others	transferase	.	Thiosulfate sulfurtransferase; Rhodanese; Senescence-associated protein; Sulfurtransferase protein 16 [<i>Z. mays</i>]	2.8056
SCRFSD1019E08.g	.	.	.	Putative transporter arsB OS=Dictyostelium discoideum hypothetical protein OsJ_031088 [<i>O. sativa</i>]	-1.1549
SCCART2001D04.g	.	.	Development	Root-specific; RCc3	1.0124
SCJLLR1054A06.g	Hormone biosynthesis	oxidase	Hormone biosynthesis	ethylene; ACC oxidase	1.1619
SCSFSD1065B09.g	.	.	.	Pyrophosphate-energized vacuolar membrane proton pump; Pyrophosphate-energized inorganic pyrophosphatase; H(+)-PPase; unknown [<i>Z. mays</i>]	-1.7465
SCCCCL3004A08.b	Others	others	.	Early nodulin 93; N-93; unknown [<i>Z. mays</i>]	2.7904
SCEQRT1029H04.g	Unknown	Uncharacterized	Library-specific	By domain; Tryp_alpha_amyl	1.8805

* Log₂ dos valores de expressão 1-MCP/controle.

Anexo 5. Genes diferencialmente expressos no segmento 5 em resposta a 1-MCP.

Transcript	Function category	Structural category	Category	Sub category 1	Expression*
SCBFAD1048A12.g	.	.	.	unknown [<i>Z. mays</i>]	-0.8932
SCCCLR1024E01.g	Pathogen Resistance	others	.	unknown [<i>Z. mays</i>]	-0.9530
SCCCHR1001C06.g	Transport	Membrane transporters ABC transporter	Transporters	ABC transporter (ABC_tran); unknown [<i>Z. mays</i>]	-0.7641
SCSGHR1072E12.g	Lipid metabolism	hydrolase	.	Abhydrolase domain-containing protein 6-A; unknown [<i>Z. mays</i>]	-0.9493
SCCCT3002G10.g	Cell wall metabolism	Lignin metabolism	Cell wall	dirigent-like protein; hessian fly response gene 1 protein [<i>T. aestivum</i>]	-1.9128
SCSGAD1006B03.g	Others	Glicosil-transferase	Cell wall	glucosyltransferase; unknown [<i>Z. mays</i>]	-0.8822
SCEQRT1024C07.g	Transport	kinase/phosphatase	.	Pyrophosphate-energized vacuolar membrane proton pump; Pyrophosphate-energized inorganic pyrophosphatase; H(+)-PPase; proton translocating pyrophosphatase [<i>O. sativa</i>]	3.7218
SCCST3C04A05.g	Cell wall metabolism	redistribution of phospholipids after cell activation or injury.	.	Uncharacterized protein At5g01750; unknown [<i>Z. mays</i>]	-0.8512
SCCCLR2001E08.g	.	.	.	Non-symbiotic hemoglobin; Hbt; ZEAm GLB1; hemoglobin [<i>Z. mays</i>] Non-symbiotic hemoglobin; Hbt; ZEAm GLB1 Chain A, B and C	1.1590
SCVPRT2073F02.g	.	.	.	unknown [<i>Z. mays</i>]	1.6208
SCCCT3004A10.g	.	.	No matches	.	-3.8448
SCSBLB1035F03.g	Others	Leucine Rich Repeat	Receptors	receptor Ser/Thr kinase-unclassified; caneURLK-119 (LRR)	-0.8086
SCEZLB1013C10.g	AminoAcid and nitrogen metabolism	synthase	.	Non-specific lipid-transfer protein-like protein At5g64080; unknown [<i>Z. mays</i>]	-1.0312
SCCCL3001F06.g	.	.	.	Sulfite oxidase; Moco-containing protein AtMCP; At-SO; AtSOX; Os12g0442800 [<i>O. sativa</i> (japonica cultivar-group)] Sulfite oxidase, putative, expressed [<i>O. sativa</i>]	1.2940

Anexo 5. Continuação

SCMCRT2086A02.g	Nucleotide metabolism	synthase	.	Phosphoribosylamine-glycine ligase, chloroplastic; Glycinamide ribonucleotide synthetase; GARS; Phosphoribosylglycinamide synthetase; hypothetical protein OsI_036507 [<i>O. sativa</i> (indica cultivar-group)]	2.4069
SCCART1C04G12.g	.	.	.	hypothetical protein OsJ_000657 [<i>O. sativa</i>]	4.5988
SCAGLR2011D02.g	.	.	Bioenergetics	Respiratory burst oxidase homolog protein B; NADPH oxidase RBOHB; StRBOHB; gb AAT35117.1 respiratory burst oxidase protein B [<i>O. sativa</i> (indica cultivar-group)]	2.2170
SCJLRT1015G08.g	.	.	No matches	.	-1.3950
SCBFRZ2048F01.g	Unknown	uncharacterized	.	Ser/Thr-rich protein; unknown [<i>Z. mays</i>]	1.1987
SCVPLR1028C05.g	.	.	.	unknown [<i>Z. mays</i>]	-0.9573
SCEZLB1012E05.g	Redox metabolism	peroxidase	.	Peroxidase 19; Atperox P19; ATP51; hypothetical protein OsJ_002847 [<i>O. sativa</i> (japonica cultivar-group)]	0.9200
SCCCLR2004G12.g	Unknown	others	.	unknown [<i>Z. mays</i>]	0.9732
SCEQRT3017F02.g	.	.	.	putative 32.7 kDa jasmonate-induced protein [<i>Hordeum vulgare</i>] 32 kDa protein [<i>Hordeum vulgare</i>]	-2.9281
SCSGHR1069F04.b	Others	cytochrome P450 family proteinCell wall	Stress	cytochrome P450; CYP71C4	2.4217
SCJFRT1059B01.g	Transport	Transporters	Transporters	(xan_ur_permease); permease 1 [<i>Z. mays</i>] Nucleobase-ascorbate transporter LPE1; Leaf permease protein 1 permease 1 [<i>Z. mays</i>]	-1.0196
SCSGRT2061H07.g	.	.	.	unknown [<i>Z. mays</i>]	-0.9165
SCACLR1129A10.g	Unknown	others	.	unknown [<i>Z. mays</i>]	1.4081
SCRFSD1019E08.g	.	.	.	Putative transporter arsB OS=Dictyostelium discoideum; hypothetical protein OsJ_031088 [<i>O. sativa</i> (japonica cultivar-group)]	-1.2126
SCQSRT2033A02.g	.	.	.	unknown [<i>Z. mays</i>]	-1.0233
SCVPRT2081A02.g	Unknown	uncharacterized	No matches	.	-1.2800

* Log₂ dos valores de expressão 1-MCP/controle.

Anexo 6. Distâncias par a par entre as sequências codantes de ScARA1

SHCRBa_048_E22												
SHCRBa_054_G11	0,022											
SHCRBa_109_O23	0,015	0,028										
SHCRBa_111_H15	0,001	0,023	0,016									
SHCRBa_124_A06	0,019	0,029	0,026	0,020								
SHCRBa_125_A09	0,013	0,016	0,020	0,013	0,020							
SHCRBa_133_H06	0,016	0,029	0,002	0,018	0,028	0,022						
SHCRBa_143_N21	0,001	0,023	0,016	0,000	0,020	0,013	0,018					
SHCRBa_191_E06	0,001	0,023	0,016	0,000	0,020	0,013	0,018	0,000				
SHCRBa_220_C01	0,014	0,012	0,021	0,015	0,020	0,006	0,023	0,015	0,015			
Sobic.005G022500	0,030	0,040	0,028	0,031	0,032	0,032	0,029	0,031	0,031	0,033		

Anexo 7. Distâncias par a par entre as sequências codantes de *ScEPG1*

SHCRBa_023_N06									
SHCRBa_028_C15	0,011								
SHCRBa_036_E16	0,011	0,005							
SHCRBa_095_L16	0,012	0,001	0,006						
SHCRBa_069_B14	0,012	0,001	0,006	0,000					
SHCRBa_191_M21	0,000	0,011	0,011	0,012	0,012				
SHCRBa_224_P16	0,014	0,002	0,007	0,004	0,004	0,014			
SHCRBa_237_A10	0,011	0,000	0,005	0,001	0,001	0,011	0,002		
Sobic.001G525000	0,051	0,047	0,050	0,049	0,049	0,051	0,050	0,047	

Anexo 8A. Distâncias par a par entre as sequências codantes de *ScERF1*

SHCRBa_056_P16												
SHCRBa_062_K05	0,015											
SHCRBa_084_H10	0,016	0,005										
SHCRBa_095_D17	0,015	0,003	0,002									
SHCRBa_168_I11	0,015	0,003	0,002	0,000								
SHCRBa_180_G11	0,015	0,007	0,008	0,007	0,007							
SHCRBa_188_L14	0,015	0,000	0,005	0,003	0,003	0,007						
SHCRBa_191_C20	0,011	0,016	0,015	0,013	0,013	0,016	0,016					
SHCRBa_193_M18	0,015	0,003	0,002	0,000	0,000	0,007	0,003	0,013				
SHCRBa_216_I12	0,011	0,003	0,005	0,003	0,003	0,007	0,003	0,013	0,003			
SHCRBa_247_B02	0,000	0,015	0,016	0,015	0,015	0,015	0,015	0,011	0,015	0,011		
Sobic.002G260800	0,067	0,072	0,070	0,068	0,068	0,068	0,072	0,063	0,068	0,068	0,067	

Anexo 9A. Distâncias par a par entre as sequências codantes de ScRAV1

Anexo III - Estatísticas para o teste de significância estatística de GOSTAR										
SHCRBa_021_A10										
SHCRBa_023_O12	0,003									
SHCRBa_042_C10	0,000	0,003								
SHCRBa_052_D15	0,004	0,003	0,004							
SHCRBa_056_A24	0,003	0,000	0,003	0,003						
SHCRBa_170_P07	0,016	0,015	0,016	0,016	0,015					
SHCRBa_174_O12	0,002	0,001	0,002	0,002	0,001	0,014				
SHCRBa_198_M03	0,003	0,000	0,003	0,003	0,000	0,015	0,001			
SHCRBa_199_K01	0,017	0,016	0,017	0,017	0,016	0,003	0,015	0,016		
Sobic.003G265200	0,037	0,036	0,037	0,037	0,036	0,033	0,035	0,036	0,034	

Anexo 10A. Número de *reads* sequenciados nas bibliotecas do transcriptoma de raiz

	Segmento 1	Segmento 2	Segmento 3	Segmento 4
Réplica A	63.500.049	40.787.158	57.029.342	61.017.224
Réplica B	56.630.107	54.447.857	56.062.773	34.702.063
Réplica C	70.669.046	51.684.071	51.968.410	55.060.071
Média	63.599.734	48.973.029	55.020.175	50.259.786
Desvio padrão	7.020.000	7.222.603	2.686.730	13.798.695

Anexo 10B. Comprimento das sequências codantes dos genes alvo usadas como referência para mapeamento de *reads* do transcriptoma

	056_P16	062_K05	084_H10	095_D17	168_I11	180_G11	188_L14	191_C20	193_M18	216_I12	247_B02
BACs ScERF1 CDS (pb)	607	601	601	598	631	568	569	569	598	568	607
BACs ScARA1 CDS (pb)	048_E22 1903	054_G11 2128	109_O23 1903	111_H15 1903	124_A06 1870	125_A09 1903	133_H06 1822	143_N21 1903	191_E06 1903	220_C01 2272	
BACs ScRAV1 CDS (pb)	021_A10 1141	023_O12 1141	042_C10 1141	052_D15 973	056_A24 1141	170_P07 1123	174_O12 1142	198_M03 1141	199_K01 1132		
BACs ScEPG1 CDS (pb)	023_N06 1315	028_C15 1309	036_E16 1309	069_B14 1309	095_L16 1504	191_M21 1315	224_P16 1273	237_A10 1003			

Anexo 11. Mapa do vetor RAVpFAST-G02

Anexo 12. Mapa do vetor ERFpFAST-G02

Anexo 13. Mapa do vetor EPGpm-pKGWFST7

Anexo 14. Mapa do vetor ARApM-pKGWFST7

REVIEW PAPER

How endogenous plant cell-wall degradation mechanisms can help achieve higher efficiency in saccharification of biomass

Eveline Q. P. Tavares, Amanda P. De Souza and Marcos S. Buckeridge*

Laboratory of Plant Physiological Ecology (LAFIECO), Department of Botany, Institute of Biosciences, University of São Paulo, Rua do Matão 277, São Paulo, SP, Brazil

* To whom correspondence should be addressed. E-mail: msbuck@usp.br

Received 1 December 2014; Revised 9 March 2015; Accepted 16 March 2015

Abstract

Cell-wall recalcitrance to hydrolysis still represents one of the major bottlenecks for second-generation bioethanol production. This occurs despite the development of pre-treatments, the prospect of new enzymes, and the production of transgenic plants with less-recalcitrant cell walls. Recalcitrance, which is the intrinsic resistance to breakdown imposed by polymer assembly, is the result of inherent limitations in its three domains. These consist of: (i) porosity, associated with a pectin matrix impairing trafficking through the wall; (ii) the glycomic code, which refers to the fine-structural emergent complexity of cell-wall polymers that are unique to cells, tissues, and species; and (iii) cellulose crystallinity, which refers to the organization in micro- and/or microfibrils. One way to circumvent recalcitrance could be by following cell-wall hydrolysis strategies underlying plant endogenous mechanisms that are optimized to precisely modify cell walls *in planta*. Thus, the cell-wall degradation that occurs during fruit ripening, abscission, storage cell-wall mobilization, and aerenchyma formation are reviewed in order to highlight how plants deal with recalcitrance and which are the routes to couple prospective enzymes and cocktail designs with cell-wall features. The manipulation of key enzyme levels *in planta* can help achieving biologically pre-treated walls (i.e. less recalcitrant) before plants are harvested for bioethanol production. This may be helpful in decreasing the costs associated with producing bioethanol from biomass.

Keywords: Abscission, aerenchyma, bioenergy, cell wall, cell-wall polysaccharide, fruit ripening, pre-treatment, recalcitrance storage mobilization.

Introduction

The current increase in the world's demand for energy and the need to reduce the dependence on fossil fuel, as well as the establishment of alternative fuels as a cost-effective production, has induced an increase in scientific efforts in plant biology. Although 23.5 million of gallons of first-generation (1G) bioethanol produced from sucrose and starch were generated in 2013 (Renewable Fuels Association, 2014, <http://www.ethanolrfa.org/>), the fuel industry can no longer rely solely on 1G technologies to supply the future demands. As a result, second-generation

(2G) bioethanol is an alternative to overcome this challenge. Production of 2G bioethanol is based on plant biomass, which is considered the most abundant and highly wasted organic raw material available (Pauly and Keegstra, 2008). Grasses such as sugarcane, sorghum, switchgrass, maize, and *Miscanthus* have an appealing biotechnological potential to be used as feedstocks for 2G bioethanol production (Kim and Dale, 2004; Carroll and Somerville, 2009). However, none of them fulfils all the needs for an ideal biofuel crop (De Souza *et al.*, 2014).

Abbreviations: 1G, first generation; 2G, second generation; AX, arabinoxylan; GH, glycosyl hydrolase; HG, homogalacturonan; MLG, mixed-linkage (1→3),(1→4)-β-D-glucan; XG, xyloglucan; XTH, xyloglucan transglycosylase/hydrolase.

© The Author 2015. Published by Oxford University Press on behalf of the Society for Experimental Biology. All rights reserved.
For permissions, please email: journals.permissions@oup.com

In order to convert plant biomass into fermentable sugars for bioethanol production, it is necessary to access the polysaccharides (e.g. cellulose, hemicelluloses, and pectins) that form plant cell walls, which means circumventing cell-wall recalcitrance. In industry, the breakdown of these polymers is usually accomplished by chemical biomass pre-treatment followed by hydrolysis with glycosyl hydrolase (GH) cocktails. Although considerable improvement in microbial fermentation and GH efficiencies have been made (Gilbert, 2014) and a few companies around the world are already producing 2G bioethanol, cell-wall recalcitrance to hydrolysis remains a major barrier. This is so because 2G bioethanol has not yet achieved full economic sustainability to complement the 1G technology and compete with other fuels (Lynd *et al.*, 2008). Furthermore, the large amount of enzyme required for polysaccharide hydrolysis corresponds to 9% of the total cost for biomass conversion (Aden *et al.*, 2002), which can render 2G-bioethanol unfeasible in a larger scale (Banerjee *et al.*, 2010). In this sense, any scientific information that can help decrease cell-wall recalcitrance and, consequently, the amount of enzyme used by the industry will probably have an important impact on bioethanol production as a renewable and economically sustainable biofuel.

In this article, we discuss details of how the composition and assembly of wall domains can affect recalcitrance and saccharification, highlighting the key cell-wall aspects contributing to its inaccessibility to GHs. We report the main drawbacks of pre-treatment and molecular biology approaches as a result of lack of knowledge on how the selective and progressive removal of polysaccharides targeted by one reagent or enzyme activity can influence the overall cell-wall structure. Next, we discuss possible key features of endogenous plant mechanisms of cell-wall degradation that could be used as a guide to better complement and/or help direct the ongoing strategies to improve saccharification of biomass. Thus, the cell-wall hydrolysis underlying mobilization of cell-wall storage polysaccharides, abscission, fruit ripening, and aerenchyma formation are highlighted as possible ways to optimize the design of enzyme hydrolysis processing of plant biomass in the industry. In addition, the focus on how plant cells degrade their own walls can better direct the selection of candidate genes to be employed in the production of transgenic plants. This could provide a less-recalcitrant biomass input to the industry in order to contribute to reducing the costs of bioenergy production.

Recalcitrance: a whole-cell-wall problem

One of the key steps in comprehending cell-wall recalcitrance is the understanding of cell-wall composition and architecture. Cell walls are composed of cellulose, hemicelluloses, pectins, lignin, and proteins (Carpita and McCann, 2000) (Fig. 1). Commelinoid monocots, such as the grasses, present type II cell walls in which the cellulose network is associated with arabinoxylans (AXs), xyloglucans (XGs), and mixed-linkage (1→3),(1→4)-β-D-glucans (MLGs). These polysaccharides are all embedded in a matrix containing branched AXs and pectins composed of homogalacturonan (HG),

rhamnogalacturonans and arabinogalactans (Carpita, 1996; Carpita and McCann, 2000). In this type of wall, there are also hydroxycinnamic acids, and *p*-coumaric and ferulic acids, which may form dimers interconnecting AX, XG, and pectin molecules. Cell-wall proteins such as extensins may also interfere with recalcitrance. However, no literature has been found reporting experiments with extensin for biofuel purposes. All of these cell-wall components are arranged as an entangled structure that confers important cell-wall properties such as porosity, mechanical strength, cell shape, and signalling, to name but a few (Fig. 1). Cell walls are therefore extremely complex composites that are precisely assembled and constantly modified in plant cells.

Besides cellulose, hemicelluloses, pectins, and proteins, lignin is also a relevant cell-wall component. It consists of a complex polymer of aromatic alcohols and has been traditionally addressed as one of the main contributors to cell-wall recalcitrance (Simmons *et al.*, 2010). However, recent studies have shown that recalcitrance does not rely on lignin alone but also on polysaccharides (Ray *et al.*, 2012; Torres *et al.*, 2013; Marriot *et al.*, 2014). For instance, in switchgrass, xylan is indicated as a relevant recalcitrance factor, since the removal of this polysaccharide resulted in almost 100% glucose yield after enzyme action (De Martini *et al.*, 2013). Similar results have been produced for *Miscanthus* biomass, in which the hemicellulose content was found to be the most relevant factor correlated to digestibility (Xu *et al.*, 2012). For maize, Penning *et al.* (2014) revealed that quantitative trait loci for lignin abundance do not overlap with those related to saccharification, but rather are associated with whole-cell-wall metabolism, vascularization, and fibre formation. Remarkably, qualitative (i.e. polysaccharide-polysaccharide interactions and polysaccharide fine structure) rather than quantitative effects seem to play a role in recalcitrance/hydrolysis in *Miscanthus sinensis* (De Souza *et al.*, 2015, this issue). In this sense, the interaction between polymers within the cell wall is also closely connected to its inherent resistance to hydrolysis. The assembly and range of possible cross-links between polymers defines the cell-wall architecture, divided into three domains (cellulose, hemicellulose, and pectins; Fig. 1a), which imposes several constraints on degradability.

The first barrier to hydrolysis is the selective cell-wall accessibility to GHs. This porosity feature is determined mainly by the pectin domain (Baron-Epel *et al.*, 1988) and the number and size of the enzymes that can penetrate the biomass are probably selected by the pore sizes of the wall matrix (Buckeridge *et al.*, 2015).

Another critical aspect for recalcitrance is the polymer-polymer interaction, which is closely associated with the hemicellulose domain. As these polysaccharides prevent access to cellulose, knowledge about their fine structure as well as their interactions with other polymers is mandatory. Differences in fine structure can be related to sugar (i.e. mono- or oligosaccharides) or non-sugar (i.e. acetylation, methylation, or feruloylation) branching decorations in the main chains of polysaccharides. These differences affect the interactions between polymers, which in turn are thought to display a

Fig. 1. Different views of the cell-wall assembly displaying some of its complexity. (A) Three cell-wall domains and their main features. (B) Assembly of cellulose micro- and macrofibrils as observed in grasses (maize and sugarcane). Microfibrils are shown covered with hemicelluloses and pectins. (C) Tentative architectural model of the grass cell wall, where the macrofibrillar domain is immersed in a pectin matrix. Adapted from Buckeridge *et al.* (2015), with permission.

complex structure that appears to determine the mode of action of GHs.

Every fine structure of each different polysaccharide in the wall of a given cell or tissue is encoded by the bio-synthetic system so that a given polysaccharide can, for instance, be ‘blocked’ or ‘open’ to hydrolytic enzyme action. In other words, the supramolecular structure formed by polysaccharides seems to derive a glycomic code, which is characterized by information properties that emerge from the assembly of cell-wall polymers (Buckeridge and De Souza, 2014; Barbieri, 2015). In this sense, an understanding of the glycomic code is one of the keys to open the way for efficient hydrolysis.

Another important aspect of the cell walls related to recalcitrance and architecture (the final expression of the glycomic code) is the way cellulose forms crystals, resulting in microfibrils that vary along the plant body and tissues (Driemeier *et al.*, 2012). In grasses, these microfibrils can form bundles, possibly tethered by hemicelluloses (i.e. xylans and/or XGs), generating macrofibrils (Ding and Himmel, 2006; Kozlova *et al.*, 2014; Langan *et al.*, 2014; Buckeridge *et al.*, 2015) (Fig. 1B).

Surpassing the recalcitrance issue has proven to be more challenging than previously foreseen. Solving this problem relies on the early production of GHs with lower diameter associated with knowledge related to the emergent properties of polymer–polymer interactions (e.g. the glycomic code). In this sense, the first set of enzymes (those allowed to cross the cell-wall pores) would open the access to further ones aimed at attacking key hemicelluloses in order to open accessibility to amorphous cellulose stretches. Following

these steps is crucial for gaining control on how to precisely attack polysaccharides, making cell walls as a whole more amenable to hydrolysis (Buckeridge and De Souza, 2014; Ong *et al.*, 2014). Furthermore, this would help to design rational enzyme cocktails able to facilitate the complete disassembly of the cell wall (Buckeridge and De Souza, 2014). As will be discussed in the following sections, the strategies currently employed to hydrolyse the cell wall often do not take into account all the features responsible for its recalcitrance.

Current approaches to increase cell-wall digestibility

Currently, cell-wall recalcitrance is overwhelmed with the use of biomass pre-treatments that are responsible for removing lignin and/or decreasing cellulose crystallinity. However, those procedures can interfere with macro- and microfibril associations, reducing the access to cellulose and also removing pectins and some soluble hemicelluloses from cell walls. For instance, Alonso-Simón *et al.* (2010) demonstrated that the use of liquid hot-water pre-treatment removes xylans and AX from wheat straw. Similarly, Li *et al.* (2014) showed loss of xylans, MLGs, XGs, and pectins in corn stover and goldenrod (*Solidago canadensis*) when they were pre-treated with alkaline hydrogen peroxide. Thus, considering the percentage of hemicelluloses in cell walls of the main biomass feedstocks (Pauly and Keegstra, 2008), pre-treatments might represent a loss of polysaccharide of about 12–33% that could potentially be used to produce 2G bioethanol.

In this sense, different approaches other than chemical pre-treatments have been attempted to improve biomass saccharification, such as the production of genetically modified plants in order to synthesize less-recalcitrant cell walls *in planta*. Although several efforts have aimed at decreasing lignin (Jung *et al.*, 2012; Vanholme *et al.*, 2012; Van Acker *et al.*, 2013) or overexpressing different microbial and plant GHs *in planta* (Brunecky *et al.*, 2011; Mahadevan *et al.*, 2011; Bartley *et al.*, 2013; Sumiyoshi *et al.*, 2013), those strategies had limitations. This is mainly due to undesirable traits in plants harbouring mutations in lignin biosynthesis, such as reduced biomass yields (Simmons *et al.*, 2010; Fu *et al.* 2011a, b), low germination frequency, lower seed viability, decreased height, and high pollen sterility (Goujon *et al.*, 2003; Mandal and Datta, 2014), besides the increased sensitivity to pathogens (Boerjan *et al.*, 2003). Moreover, some authors suggest that GH overexpression *in planta* increased the accessibility to polysaccharides other than the one targeted by the heterologous enzyme (Kaida *et al.*, 2009; Lionetti *et al.*, 2010; Mahadevan *et al.*, 2011; Biswal *et al.*, 2014). For instance, decreased cellulose crystallinity was observed after xyloglucanase expression in poplar (Kaida *et al.*, 2009). On the other hand, monosaccharide proportions other than glucose were kept constant when *Thermotoga maritima* cellulase was overexpressed in tobacco (Mahadevan *et al.*, 2011), emphasizing the need for a more complex set of enzymes to successfully attack a broader set of polysaccharides. Furthermore, cellulose hydrolysis in poplar xylem cells overexpressing xyloglucanase was higher than in xylem overexpressing xylanase or cellulase (Kaida *et al.*, 2009).

In a similar fashion, changes in pectin content and/or its substitution pattern led to increased saccharification. In poplar, tobacco, wheat, and *Arabidopsis*, the overexpression of plant pectinases such as endopolygalacturonase, pectate lyase, or pectin methylesterase led to increased release of simple sugars. It is noteworthy that not only pectins but also xylans were more easily extracted after pectate lyase overexpression (Lionetti *et al.*, 2010; Biswal *et al.*, 2014). Thus, pectic polymers, which have been widely neglected as an important component of cell walls of grasses used for bioenergy, can somehow influence the degradation of other cell-wall components. This may reduce the requirement for chemical pre-treatments.

Therefore, it seems that neither decreasing lignin content nor overexpressing a single enzyme would help in achieving major success on saccharification levels. Since the mobilization or removal of one carbohydrate leads to subsequent structural changes or rearrangements of other polymers, decreased recalcitrance may be obtained as a result of a more optimized degradation pathway. One way to address this issue is to understand the mechanistic aspects underlying some natural cell-wall degrading processes such as fruit ripening, cell-wall storage mobilization, aerenchyma formation, and abscission (Grandis *et al.*, 2014). As these processes involve a series of cell-wall modifications, they could help in designing optimized biotechnological strategies to engineer plants able to change their own cell walls in order to produce endogenous biological pre-treatments. This has the potential to diminish

the impact of the step of pre-treatment in industrial processes in the future.

Mechanisms of cell-wall degradation by plants

Several plant endogenous mechanisms of cell-wall disassembly share common features (Rose and Bennett, 1999; Roberts *et al.*, 2002; Grandis *et al.*, 2014). The step-by-step procedure undertaken by plants to modify their own cell walls commonly starts with cell expansion and separation, in which pectins are the main targets for hydrolysis. Some natural cell-wall degradation events seem to rely on cell targeting, from which the carbohydrate breakdown activation spreads through the surrounding cells (McManus *et al.*, 1998; Grandis *et al.*, 2014). Downstream signalling activates hydrolysis of cell components through programmed cell death (Bouranis *et al.*, 2007). Afterwards, first hemicelluloses and later on cellulose are degraded (Roberts *et al.*, 2002; Grandis *et al.*, 2014). These common steps required for cell-wall degradation in a broad spectrum of plant endogenous processes can be seen in a modular perspective. Thus, each module (i.e. cell separation, cell expansion, cell targeting, programmed cell death, hemicellulose hydrolysis, and cellulose hydrolysis) is a relatively conserved sequence of events taking place in fruit ripening, aerenchyma formation, leaf abscission, mobilization of cell-wall storage polysaccharides, and so on. On this basis, a modularity theory of hydrolysis has been proposed by Grandis *et al.* (2014) as a guide for a more successful strategy to achieve efficient hydrolysis *in planta*. In other words, plants could be induced to modify their cell walls by means of activation of key players from one or more modules so that cell walls could be swollen before harvesting. Thus, the biomass taken to industry could be more accessible to pre-treatment and hydrolysis by GHs.

During endogenous cell-wall hydrolysis, the plant GHs are produced and secreted to the wall in a temporal sequence compatible with the architectural feature of the cell wall of that specific cell and tissue (see Fig. 6 in De Souza *et al.*, 2013). In this way, expansins and pectinases should be recruited for the cell separation and expansion modules, whereas a whole set of xyloglucanases and cellulases would be used afterwards in the process in order to loosen the walls.

Considering key modifications and/or degradation of cell-wall polysaccharides found in these natural processes, in the following sections we will discuss how they can drive improvements in the technologies for 2G bioethanol production. Next, we will highlight how the modifications on each polysaccharide present in the cell walls of the main bioenergy feedstocks could lead to further changes potentially enhancing biomass hydrolysis.

Pectins: targets to change cell adhesion and porosity

Pectins are polysaccharides related to plant growth, development, and morphogenesis. They are also relevant in plant defence, interaction with ions, cell adhesion and cell

expansion, and cell-wall porosity (Baron-Epel *et al.*, 1988; Ridley *et al.*, 2001). The latter is specifically relevant to bio-energy, since the average pore size of wall matrices is likely to limit the traffic of GHs (Buckeridge *et al.*, 2015). Thus, in order to reduce the limitation of enzyme action in intact cell walls caused by the presence of the pectin domain, pectinases should be present in enzymatic cocktails (Buckeridge *et al.*, 2015) as part of the strategy to modify plants by engineering (Grandis *et al.*, 2014).

In nature, the attack on pectins seems to be an early step in distinct endogenous cell-wall degradation events. This reinforces the idea that it is necessary first to hydrolyse those polysaccharides in order to access other cell-wall polysaccharides. In natural processes such as lysigenous aerenchyma formation, which leads to the opening of gas spaces after cell death and cell-wall degradation (Evans, 2003), the first event seems to be cell separation and cell expansion (Yamauchi *et al.*, 2013; Grandis *et al.*, 2014). It is noteworthy that changes in the pattern of pectin esterification are observed in the middle lamella during the early stages of aerenchyma development (Gunawardena *et al.*, 2001). These changes are possibly related to the increase in polygalacturonase and pectin lyase expression during aerenchyma formation (Rajhi *et al.*, 2011).

Pectin hydrolysis also occurs during fruit softening (Brummell, 2006). As with aerenchyma formation, the first event observed during fruit ripening seems to be the attack on HG present in the middle lamella (Gunawardena *et al.*, 2001). Similarly, in kiwi fruits, the reduction in the molecular weight of polysaccharides present in soluble fractions strongly suggests HG depolymerization by endopolygalacturonase in the middle lamella (Redgwell *et al.*, 1992). Similar results have been obtained for papaya fruits, where an increase in the release of pectin-derived oligosaccharides was observed concomitant with an increase in polygalacturonase activity during fruit ripening (Sañudo-Barajas *et al.*, 2009).

Although high levels of endopolygalacturonase mRNA have proven not to be essential for tomato fruit softening (Kramer *et al.*, 1992), depolymerization of linear pectins was impaired when this gene was downregulated by antisense RNA (Smith *et al.*, 1990). The role of this enzyme in fruit softening was assessed further in papaya and apple fruits (Atkinson *et al.* 2012; Fabi *et al.*, 2014), leading to a decreased molecular mass of pectins. This confirmed the profound changes exerted by this enzyme in the cell-wall structure.

During tomato fruit ripening, pectin becomes gradually de-esterified due the action of pectin methylesterases. The latter enzymes are required to open access for polygalacturonases and consequently degrade the main chain of methyl-esterified pectin (Micheli, 2001). Together, these two enzymes are known to affect middle lamella integrity, controlling cell adhesion. Also, pectin dimethyl esterification favours other enzymes besides polygalacturonases by modifying cell-wall porosity (Baron-Epel *et al.*, 1988; Ehwald *et al.*, 1992; Rondeau-Mouro *et al.*, 2008). In terms of bioenergy, it has already been demonstrated that the reduction of demethyl-esterified HG in *Arabidopsis*, wheat, and tobacco increased saccharification efficiency by 35–40%. This suggests that

cellulose and hemicelluloses are more easily digested when methyl-esterified HGs are reduced in the wall, possibly facilitating the access of enzymes to the cell-wall network (Lionetti *et al.*, 2010).

Other endogenous cell-wall degradation events also require polygalacturonase and pectin methylesterase activities. Microscopic and spectroscopic approaches showed low levels of esterified pectin associated with high levels of polygalacturonase activity in the abscission zone of cells of oil palm (*Elaeis guineensis*) (Henderson *et al.*, 2001). In *Brassica napus*, a gene related to polygalacturonase seems to be expressed in the abscission zone from the base of the anther filaments, petals, and sepals (González-Carranza *et al.*, 2002). Once abscission starts, cell swelling is evident in the abscission zone where endo- β -glucanase and polygalacturonase mRNAs accumulate or have their activity increased (David *et al.*, 1999; Roberts *et al.*, 2002). As a result, middle lamella dissolution and abscission have been observed. Thus, the synergistic effect of pectin-degradation enzymes and cellulases is noticeable (González-Carranza *et al.*, 2007).

Additionally, fruit softening is significantly reduced by impaired β -galactosidase activity during the early steps in the ripening process (González-Carranza *et al.*, 2007), highlighting the relevance of this enzyme to cell-wall degradation. This GH is part of an array of enzymes required for pectin debranching. It is believed that removal of galactan substitutions from pectins will also be mandatory in other fruits (Ali *et al.*, 1995; Carrington and Pressey, 1996; Ali *et al.*, 2004; Gulzar and Amin, 2012; Kivi and Sartipnia, 2014), as well as for cotyledon cell-wall degradation (Buckeridge and Reid, 1994; Buckeridge *et al.*, 2005), allowing subsequent cell-wall changes to occur (Brummell and Harpster, 2001).

Arabinofuranosidases are thought to act synergistically with β -galactosidases on pectins (Hirano *et al.*, 1994), with transcripts for both enzymes showing a similar pattern (Sozzi *et al.*, 2002). Arabinofuranosidase catalyses the removal of α -L-arabinoyl branches from arabinogalactans and AXs (Yoshioka *et al.*, 1995). The upregulation of the corresponding gene is associated with peach fruit softening (Hayama *et al.*, 2006), and an increased loss of arabinose residues has been demonstrated to contribute to the ripening of several fruits (Yamaki *et al.*, 1979; Peña and Carpita, 2004; Rosli *et al.*, 2004; Tateishi *et al.*, 2005).

Therefore, the presence of β -galactosidases as well as arabinofuranosidases, endopolygalacturonases, and pectin methylesterases in plant tissues seems to be relevant for the early steps of cell-wall degradation. The use of such a set of enzymes for bioenergy purposes could favour not only 2G processes but 1G ones as well. By softening sugarcane stem with these enzymes, the cells would become easier to break, which could significantly reduce the energy necessary to extract sucrose.

Expansins

The process of cell expansion is an important step in natural systems that could be explored for improvement of bioethanol

production. During cell expansion, the polysaccharides that coat cellulose microfibrils should be rearranged (Cosgrove, 2000). This rearrangement involves expansins, which are proteins that provoke cell-wall relaxation, affecting the cellulose–XG network (McQueen-Mason and Cosgrove, 1995; Rose and Bennett, 1999). Expansins lack hydrolytic activity and are thought to ease the action of GHs on cellulose by destabilizing hydrogen bonds and promoting slippage of XGs on the surface of cellulose microfibrils (Sampedro and Cosgrove, 2005). For more details on the synergistic activity of microbial GHs and plant expansins, see Baker *et al.* (2000) and Arantes and Saddler (2010).

In natural systems such as fruit ripening, the overexpression of an expansin in ripening tomato resulted in softer fruits, which was correlated to extensive depolymerization of hemicelluloses and pectins (Brummell *et al.*, 1999; Brummell and Harpster, 2001). The same can be suggested for aerenchyma formation, where an increase in expansin transcripts was correlated to higher expression of pectin- and XG-active enzymes acting on cell-wall loosening in maize roots (Rajhi *et al.*, 2011). Thus, the synergistic action between hemicellulose- and pectin-degrading enzymes and expansins seems to be crucial for co-ordinated wall disassembly.

In this sense, the optimization of enzymatic cocktails to 2G bioethanol is already considering the inclusion of expansins as an optimization that would reduce biomass recalcitrance to hydrolysis (Mohanram *et al.*, 2013). As stated above for pectinases, expansins also have the potential to help 1G processing for the same reasons. Cells with loosened walls may become easy to break and facilitate the extraction of sucrose from sugarcane.

Hemicellulose removal: opening the doors towards cellulose

Hemicelluloses are the main domain present in cell walls of grass bioenergy feedstocks. Thus, the degradation of these polysaccharides is crucial to achieve efficient cell-wall disassembly. In dicots, hemicelluloses such as XGs, mannans, glucomannans, and galactomannans are involved in exceptionally efficient hydrolytic systems (Meier and Reid, 1982; De Mason *et al.*, 1983; Reid, 1985; Buckeridge *et al.*, 2000; Buckeridge, 2010). In cotyledons and endosperms of several species, cell-wall polysaccharides occurring in large proportions are completely hydrolysed in order to provide free sugars to be used for respiration and growth (Buckeridge, 2010). Although glucomannans are found in small amounts in cell walls of bioenergy crops, they can be cross-linked with cellulose microfibrils (Carpita and Gibeau, 1993), interfering with cellulose saccharification. The same could be inferred for galactomannans and mannans, which seem to be associated with increased biomass saccharification (De Souza *et al.*, 2015, this issue). In this way, enzymes such as endo- β -mannanases and α -galactosidases could improve the potential of enzymatic cocktails. Natural mechanisms that involve mannan degradation are related mainly to seed germination. In lettuce, for example, β -mannanase enzyme activity starts in the radicle tip and endosperm cap (Toorop

et al., 1998) and may be involved in cell-wall extensibility (Karssen *et al.* 1989). Some indication that these enzymes could improve cell-wall saccharification were given by Banerjee *et al.* (2010), who used a fungal mannanase and found an enhancement of hydrolysis of dried distillers' grains.

One of the main hemicelluloses in grasses is MLG (reviewed by Stone and Clarke, 1992). Due to its relatively simple structure, MLG requires only two enzymes (lichenase and β -glucosidase) to be completely hydrolysed (Fincher, 2009). Because of this, it has already been suggested that an increase in MLG content in feedstocks could increase glucose yields (McCann and Carpita, 2008).

Degradation of MLG was observed during grass and cereal seed germination (Buckeridge *et al.*, 2004) and mobilization in leaves of barley (Roulin *et al.*, 2002), as well as during aerenchyma formation in sugarcane roots (D. C. C. Leite, A. Grandis, E. Q. P. Tavares, S. Pattathil, U. Avici, A. P. De Souza, B. Mattei, M. J. Hahn, and M. S. Buckeridge, unpublished results). Crops like sugarcane and *Miscanthus* offer great opportunities for improvement of MLG degradation, making cell-wall composition more adapted for use as a bio-energy feedstock.

AX is another important hemicellulose in grasses. In contrast to MLGs, the structure of AX requires a broader set of enzymes to be degraded. In endogenous plant cell-wall degradation, hydrolysis of AX has rarely been observed. Some enzymes capable of degrading AX have been identified in germinating cereals, comprising mainly α -arabinofuranosidase and xylanase (Taiz and Honigman, 1976; Dashek and Chrispeels, 1977). Arabinofuranosidases act by removing arabinosyl side chains from AX, opening the way for degradation by xylanases (Preece and MacDougall, 1958). In fruit ripening, the presence of arabinofuranosidase mRNA and enzyme activity has been observed (Itai *et al.*, 2003). However, no report of the presence of xylanases has been made to date, suggesting that the complete degradation of this hemicellulosic polysaccharide during fruit ripening might not occur.

In the case of grasses, AX can be cross-linked to ferulic acids, so its degradation should involve feruloyl esterases. During aerenchyma formation in sugarcane roots, an increase in transcripts related to a precursor of a feruloyl esterase has been observed (E. Q. P. Tavares, A. P. De Souza, A. R. Piovezani, and M. S. Buckeridge, unpublished results), which might be indicative that this enzyme is acting during aerenchyma development. However, although endoxylanase transcripts are also found, there is no evidence of a reduction in AX in the cell wall during this same process (D. C. C. Leite, A. Grandis, E. Q. P. Tavares, S. Pattathil, U. Avici, A. P. De Souza, B. Mattei, M. J. Hahn, and M. S. Buckeridge, unpublished results). The AX present in cell walls of bioenergy crops such as sugarcane also contains decorations with acetyl groups that can interfere in enzymatic cleavage (De Souza *et al.*, 2013). Thus, in this case, an acetyl esterase should be necessary.

Concerning XG, in storage tissues XG transglycosylase/hydrolase (XTH) has been demonstrated to play a key role in

XG mobilization (Fanutti et al., 1993). Decreasing levels of storage XG have been reported during seedling development in cotyledons of the tropical tree *Hymenaea courbaril* (Tin et al., 2000), fruit ripening in avocado (O'Donoghue et al., 1994), and papaya fruit development (A. A. Cavaleri-Coreti, A. P. De Souza, and M. S. Buckeridge, unpublished results). In maize, increased expression of a gene coding for an XTH was observed during aerenchyma formation (Saab and Sachs, 1996; Rajhi et al., 2011). The XG depolymerization by endo- β -1,4-glucanases can be increased further by exoenzymes that will accomplish XG side-chain removal, making more sites available to the attack of endoenzymes (Rose and Bennett, 1999; Brummell and Harpster, 2001). In fact, Tin et al. (2000) showed that the complete degradation of XG requires α -xylosidases, β -galactosidases, and β -glucosidases to remove the XG side chains. As XG seems to be the hemi-cellulose that tethers cellulose, the mechanism by which it is hydrolysed in the plant cell can be further explored by plant scientists in order to reveal how to increase the access of enzymes to micro- and macrofibrils to cellulose (Buckeridge, 2010; Buckeridge et al., 2015).

Additionally, the presence of expansins (as discussed above) is critical in the process of detaching XG bound to the surface of cellulose micro/macrofibrils. This is due mainly to the reduced XG accessibility by XTH action when it is bound to cellulose (Pauly et al., 1999).

Cellulose

Most of the work on 2G bioethanol technologies regards cellulose degradation and cellulase activities as the main targets to solve the problem of recalcitrance. In this way, the comprehension of cellulose degradation in natural processes could help to understand how this polymer interacts with other cell-wall polysaccharides, improving knowledge on how to effectively access microfibrils.

Reports on cellulase activity are relatively common in descriptions of processes of cell-wall degradation in plants by fungi and bacteria. Usually, it is either a consequence of other polymer disassembly or overall modifications, but it rarely takes place as an isolated event. The detection of high cellulase activity in expanded cells during bean leaf abscission indicates that cell expansion must be coupled with cellulose hydrolysis. Both features are related to cell-wall loosening required for accomplishment of the abscission process (Ueda et al., 1996). In fact, cellulase action does occur in parallel with endopolygalacturonase activity in the abscission of orange calyx (Burns et al., 1998) and cotton leaves (Du et al., 2014).

As discussed previously, pectin modifications also seem to be relevant to cellulase action in lysigenous aerenchyma formation in maize roots (Gunawardena et al., 2001). In fact, for the cellulases to act on cellulose, an increase in porosity mediated by β -galactosidase and polygalacturonase is required for further disruption of hydrogen bridges between XG and cellulose by expansin (Rose and Bennett, 1999; Brummell, 2006). This raises an interesting interpretation: for cellulose to be targeted for enzymatic breakdown, cell metabolism

recruits other enzymes in order to remove polysaccharides that are coating cellulose microfibrils. This reinforces the idea that a whole series of 'accessory enzymes' and enzymes other than cellulases should be used in enzymatic cocktails for 2G bioethanol technologies.

From a bioenergy perspective, in some cases cellulase activity can be detrimental to saccharification. For instance, it has been observed that the action of cellulase on amorphous sites of cellulose microfibrils increased cellulose crystallinity proportion in avocado ripening. Cellulose and other cell-wall polymer contents did not change and also XG associated with cellulose was kept constant (O'Donoghue et al., 1994). These findings suggest that the rearrangements of microfibrils may interfere with cellulase activity, increasing crystallinity and possibly increasing recalcitrance.

Conclusions and perspectives

In this review, we have highlighted some of the main bottlenecks that need to be addressed in order to achieve better efficiency in the industrial processes for 2G bioethanol production. As the drawbacks of pre-treatments and enzymatic hydrolysis-limited efficiency are quite significant, we propose that the information about some natural cell-wall modification processes could be useful as guidelines to drive strategies for future research and technology development for bioenergy production.

Table 1 summarizes some features of the four main biological processes involving cell-wall modifications and how their key aspects could impact bioenergy production. In the four cases, the ability to hydrolyse cell walls *in vivo* could be manipulated to promote biological pre-treatment in order to provide less-recalcitrant biomass for hydrolysis in industry. This does not exclude the use of the pre-treatments already developed but rather would complement their application. The main goal is to decrease the energy input in this part of the process. Also, most of the enzymes that participate in these natural processes in plants could be characterized. In this sense, plant GHs would supplement the enzyme cocktails already in use in industry, in order to further optimize their performance.

If transgenic plants are to be largely used in bioenergy in the future, some important points should be considered. One of the major aspects is the cell adhesion involving changes in the pectins of the middle lamella. This can lead to cell separation and consequently increase the surface available to the GHs present in enzyme cocktails. As a result of some cell-wall modifications, recalcitrance could decrease due to changes in wall loosening *in vivo*.

Some challenges will have to be faced in order to achieve these goals from a practical viewpoint. In order to activate cell-wall modification modules in a given bioenergy crop, one will have to acquire control on how to activate that process at the right time. In this sense, plants holding inducible and/or tissue-specific promoters could be used to grow normally until the overexpression of the heterologous or endogenous enzyme is activated (Lionetti et al., 2010; Biswal et al., 2014). In addition, transcription factors controlling GH expression could be

Table 1. Biotechnological aspects related to bioenergy for possible uses of part of the four main biological processes in which cell-wall degradation takes place, in future programmes of synthetic biology

Bioenergy aspects	Biological process			
	CW storage mobilization	Fruit ripening	Aerenchyma formation	Abscission
Features to be used in synthetic biology	Increase one type of polysaccharide in the wall; regulation of timing of hydrolysis activation	Pectin hydrolysis; decrease cell adhesion; promote cell separation and cell expansion by changes in xyloglucan cellulose interactions; regulation of timing for process activation	Decrease cell adhesion; promote cell separation; modifications in pectins	Decrease cell adhesion; promote cell separation; regulation of timing of activation
CW domain attacked	Mainly the hemicellulose domain	Pectin, hemicellulose, and cellulose domains	Mainly the pectin domain	Pectin, hemicellulose and cellulose domains
Effect on pre-treatment	May make some pre-treatments easier due to decrease in recalcitrance	Loosening of walls and increase in substrate–enzymes contacts	Makes tissues loosen and increases effect	Little effect
Effect on saccharification	Increased facility for saccharification; avoidance of decrease of C5 impact	General increase in saccharification of walls as a whole	Limited increase in saccharification due to remaining recalcitrant walls	Limited increase in saccharification due to remaining recalcitrant walls
Impact on 1G	Little impact	Improvement of extraction of sucrose in sugarcane and sorghum	Limited effect as walls continue to be hard to break	Limited effect as walls continue to be hard to break
Main problems to be solved before application in synthetic biology	Gain control of polysaccharide biosynthesis mechanisms; find out which signalling mechanisms determine the partition of polysaccharides in the wall	Find what controls the timing of production of enzymes	Understand signalling so as to avoid total collapse of the wall	Understand how the system triggers only a specific group of cells instead of collapsing the whole tissue

relevant targets to understand how to activate cell-wall degradation processes.

The challenge is enormous and it will take some time before we can achieve a level of control that will be secure enough to be used in agriculture and profitable at the same time. However, some of the key basic knowledge is already available so that biotechnological strategies can now be devised. As a consequence, the efficiency of bioenergy production from plants can be even more improved in a scenario of intense demand for higher sustainability.

Acknowledgements

We thank the Instituto Nacional de Ciência e Tecnologia do Bioetanol – INCT do Bioetanol (FAPESP 2008/57908-6 and CNPq 574002/2008-1), the Centro de Processos Biológicos e Industriais para Biocombustíveis – CeProBIO (FAPESP 2009/52840-7 648 and CNPq 490022/2009-0), and the European Union consortium SUNLIBB (project ID 251132) for support.

References

- Aden A, Ruth M, Ibsen K, Jechura J, Neeves K, Sheehan J, Wallace B, Montague L, Slayton A, Lukas J. 2002. *Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover*. NREL Technical Report TP-510–32438.
- Ali ZM, Armugam S, Lazan H. 1995. β -Galactosidase and its significance in ripening mango fruit. *Phytochemistry* **38**, 1109–1114.
- Ali ZM, Chin LH, Lazan H. 2004. A comparative study on wall degrading enzymes, pectin modifications and softening during ripening of selected tropical fruits. *Plant Science* **167**, 317–327.

- Alonso-Simón A, Kristensen JB, Obro J, Felby C, Willats WG, Jørgensen H. 2010. High-throughput microarray profiling of cell wall polymers during hydrothermal pre-treatment of wheat straw. *Biotechnology and Bioengineering* **105**, 509–514.
- Arantes V, Saddler JN. 2010. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnology for Biofuels* **3**, 4.
- Atkinson RG, Sutherland PW, Johnston SL, Gunaseelan K, Hallett IC, Mitra D, Brummell DA, Schröder R, Johnston JW, Schaffer RJ. 2012. Down-regulation of *POLYGALACTURONASE1* alters firmness, tensile strength and water loss in apple (*Malus domestica*) fruit. *BMC Plant Biology* **12**, 129.
- Baker JO, King MR, Adney WS, et al. 2000. Investigation of the cell-wall loosening protein expansin as a possible additive in the enzymatic saccharification of lignocellulosic biomass. *Applied Biochemistry and Biotechnology* **84**, 217–223.
- Banerjee G, Car S, Scott-Craig JS, Borrusch MS, Walton JD. 2010. Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. *Biotechnol Biofuels* **3**, 22.
- Barbieri, M. 2015. *Code biology. A new science of life*. Switzerland: Springer International Publishing.
- Baron-Epel O, Gharyal PK, Schindler M. 1988. Pectins as mediators of wall porosity in soybean cells. *Planta* **175**, 389–395.
- Bartley LE, Peck ML, Kim SR, et al. 2013. Overexpression of a BAHD acyltransferase, OsAt10, alters rice cell wall hydroxycinnamic acid content and saccharification. *Plant Physiology* **161**, 1615–1633.
- Biswal AK, Soeno K, Gandia ML, et al. 2014. Aspen pectate lyase PtxtPL1–27 mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. *Biotechnology for Biofuels* **7**, 11.
- Boerjan W, Ralph J, Baucher M. 2003. Lignin biosynthesis. *Annual Review of Plant Biology* **54**, 519–546.
- Bouranis DL, Choriantopoulou SN, Siyiannis VF, Protonotarios VE, Hawkesford MJ. 2007. Lysigenous aerenchyma development in

roots—triggers and cross-talks for a cell elimination program. *International Journal of Plant Developmental Biology* **1**, 127–140

Brummell DA. 2006. Cell wall disassembly in ripening fruit. *Functional Plant Biology* **33**, 103–119.

Brummell DA, Harpster MH. 2001. Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Molecular Biology* **47**, 311–340.

Brummell DA, Harpster MH, Civello PM, Palys JM, Bennett AB, Dunsmuir P. 1999. Modification of expansin protein abundance in tomato fruit alters softening and cell wall polymer metabolism during ripening. *The Plant Cell* **11**, 2203–2216.

Brunecky R, Selig MJ, Vinzant TB, Himmel ME, Lee D, Blaylock MJ, Decker SR. 2011. *In planta* expression of *A. cellulolyticus* Cel5A endocellulase reduces cell wall recalcitrance in tobacco and maize. *Biotechnology for Biofuels* **4**, 1.

Buckeridge MS. 2010. Seed cell wall storage polysaccharides: models to understand cell wall biosynthesis and degradation. *Plant Physiology* **154**, 1017–1023.

Buckeridge MS, De Souza AP. 2014. Breaking the “glycomic code” of cell wall polysaccharides may improve second-generation bioenergy production from biomass. *BioEnergy Research* **7**, 1065–1073.

Buckeridge MS, dos Santos WD, Tiné MAS, De Souza AP. 2015. The cell wall architecture of sugarcane and its implication to cell wall recalcitrance. In: Lam E, ed. *Compendium of bioenergy crops for science publishers*. Boca Raton, FL: CRC Press (in press).

Buckeridge MS, Hutcheon I, Reid JSG. 2005. The role of exo-(1,4)- β -galactanase in the mobilisation of polysaccharides from the cotyledonary cell walls of *Lupinus angustifolius* L. following germination. *Annals of Botany* **96**, 435–444.

Buckeridge MS, Rayon C, Urbanowicz B, Tiné MAS, Carpita NC. 2004. Mixed linkage (1 $\rightarrow$ 3),(1 $\rightarrow$ 4)- β -D-glucans of grasses. *Cereal Chemistry* **81**, 115–127.

Buckeridge MS, Reid JSG. 1994. Purification and properties of a novel β -galactosidase or exo-(1/4)- β -D-galactanase from the cotyledons of germinated *Lupinus angustifolius* L. seeds. *Planta* **192**, 502–511.

Buckeridge MS, Santos HP, Tiné MAS. 2000. Mobilisation of storage cell wall polysaccharides in seeds. *Plant Physiology and Biochemistry* **38**, 141–156.

Burns JK, Lewandowski DJ, Nairn CJ, Brown GE. 1998. Endo-1,4- β -glucanase gene expression and cell wall hydrolase activities during abscission in Valencia orange. *Physiologia Plantarum* **102**, 217–225.

Carpita N, McCann M. 2000. The cell wall. In: Buchanan B, Gruissem W, Jones R, eds. *Biochemistry & molecular biology of plants*. Poole, UK: Orca Book Services and American Society of Plant Physiologists.

Carpita NC. 1996. Structure and biogenesis of the cell walls of grasses. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**, 445–476.

Carpita NC, Gibeaut DM. 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *The Plant Journal* **3**, 1–30.

Carrington CMS, Pressey R. 1996. β -Galactosidase II activity in relation to changes in cell wall galactosyl composition during tomato ripening. *Journal of the American Society for Horticultural Science* **121**, 132–136.

Carroll A, Somerville C. 2009. Cellulosic biofuels. *Annual Review of Plant Biology* **60**, 165–182.

Cosgrove DJ. 2000. Loosening of plant cell walls by expansins. *Nature* **407**, 321–326.

Dashek WV, Chrispeels MJ. 1977. Gibberellic-acid-induced synthesis and release of cell-wall-degrading endoxylanase by isolated aleurone layers of barley. *Planta* **134**, 251–256.

David A, Brummell I, Bradford D. 1999. Antisense suppression of tomato endo-1,4-glucanase Cel2 mRNA accumulation increases the force required to break fruit abscission zones but does not affect fruit softening. *Plant Molecular Biology* **40**, 615–622.

De Martini JD, Pattathil S, Miller JS, Li H, Hahn MG, Wyman CE. 2013. Investigating plant cell wall components that affect biomass recalcitrance in poplar and switchgrass. *Energy and Environmental Science* **6**, 898–909.

De Mason DA, Sexton R, Grant Reid JS. 1983. Structure, composition and physiological state of the endosperm of *Phoenix dactylifera* L. *Annals of Botany* **52**, 71–80.

De Souza AP, Grandis A, Leite DCC, Buckeridge MS. 2014. Sugarcane as a bioenergy source: history, performance, and perspectives for second-generation bioethanol. *BioEnergy Research* **7**, 24–35.

De Souza AP, Kamei CLA, Torres AF, Pattathil S, Hahn MG, Trindade LM, Buckeridge, MS. 2015. How cell wall complexity influences saccharification efficiency in *Miscanthus sinensis*. *Journal of Experimental Botany* **66**, XXX–XXX.

De Souza AP, Leite DCC, Pattathil S, Hahn MG, Buckeridge MS. 2013. Composition and structure of sugarcane cell wall polysaccharides: implications for second-generation bioethanol production. *Bioenergy Research* **6**, 564–579.

Ding SY, Himmel ME. 2006. The maize primary cell wall microfibril: a new model derived from direct visualization. *Journal of Agricultural and Food Chemistry* **54**, 597–606.

Driemeier C, dos Santos WD, Buckeridge MS. 2012. Cellulose crystals in fibrovascular bundles of sugarcane culms: orientation size distortion and variability. *Cellulose* **19**, 1507–1515.

Du M, Li Y, Tian X, Duan L, Zhang M, Tan W, Xu D, Li Z. 2014. The phytotoxin coronatine induces abscission-related gene expression and boll ripening during defoliation of cotton. *PLoS ONE* **9**, e97652.

Ehwald R, Woehlecke H, Titel C. 1992. Cell wall microcapsules with different porosity from suspension cultured *Chenopodium album*. *Phytochemistry* **31**, 3033–3038.

Evans, DE. 2003. Aerenchyma formation. *New Phytologist* **161**, 35–49.

Fabi JP, Broetto SG, Silva SLGL, Zhong S, Lajolo FM, Nascimento JRO. 2014. Analysis of papaya cell wall-related genes during fruit ripening indicates a central role of polygalacturonases during pulp softening. *PLoS ONE* **9**, e105685.

Fanutti C, Gidley MJ, Reid JS. 1993. Action of a pure xyloglucan endo-transglycosylase (formerly called xyloglucan-specific endo-(1 $\rightarrow$ 4)- β -D-glucanase) from the cotyledons of germinated nasturtium seeds. *The Plant Journal* **3**, 691–700.

Fincher GB. 2009. Exploring the evolution of (1,3;1,4)- β -D-glucans in plant cell walls: comparative genomics can help! *Current Opinion in Plant Biology* **12**, 140–147.

Fu C, Mielenz JR, Xiao X, et al. 2011a. Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proceedings of the National Academy of Sciences, USA* **108**, 3803–3808.

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

Gilbert HJ. 2014. Developing novel enzyme repertoires for the efficient deconstruction of plant biomass tailored for the bioenergy industry. In: McCann MC, Buckeridge MS, Carpita NC, eds. *Plants and bioenergy*. New York: Springer, 197–209.

González-Carranza ZH, Elliott KA, Roberts JA. 2007. Expression of polygalacturonases and evidence to support their role during cell separation processes in *Arabidopsis thaliana*. *Journal of Experimental Botany* **58**, 3719–3730.

González-Carranza ZH, Whitelaw CA, Swarup R, Roberts JA. 2002. Temporal and spatial expression of a polygalacturonase during leaf and flower abscission in oilseed rape and *Arabidopsis*. *Plant Physiology* **128**, 534–543.

Goujon T, Ferret V, Mila I, Pollet B, Ruel K, Burlat V, Joseleau JP, Barrière Y, Lapiere C, Jouanin L. 2003. Down-regulation of the *AtCCR1* gene in *Arabidopsis thaliana*: effects on phenotype, lignins and cell wall degradability. *Planta* **217**, 218–228.

Grandis A, De Souza AP, Tavares EQP, Buckeridge MS. 2014. Using natural plant cell wall degradation mechanisms to improve second generation bioethanol. In: McCann MC, Buckeridge MS, Carpita NC, eds. *Plants and bioenergy*. New York: Springer, 211–230.

Gulzar S, Amin S. 2012. Kinetic studies on β -galactosidase isolated from apricots (*Prunus armeniaca kaisa*). *American Journal of Plant Sciences* **3**, 636–645.

Gunawardena AHLAN, Pearce DME, Jackson MB, Hawes CR, Evans DE. 2001. Rapid changes in cell wall pectic polysaccharides are closely associated with early stages of aerenchyma formation, a spatially localized form of programmed cell death in roots of maize (*Zea mays* L.) promoted by ethylene. *Plant, Cell & Environment* **24**, 1369–1375.

- Hayama H, Shimada T, Fujii H, Ito A, Kashimura Y.** 2006. Ethylene-regulation of fruit softening and softening-related genes in peach. *Journal of Experimental Botany* **57**, 4071–4077.
- Henderson J, Davies HA, Heyes SJ, Osborne DJ.** 2001. The study of a monocotyledon abscission zone using microscopic, chemical, enzymatic and solid state ^{13}C CP/MAS NMR analyses. *Phytochemistry* **56**, 131–139.
- Hirano Y, Tsumuraya Y, Hashimoto Y.** 1994. Characterization of spinach leaf α -l-arabinofuranosidases and β -galactosidases and their synergistic action on an endogenous arabinogalactan-protein. *Physiologia Plantarum* **92**, 286–296.
- Itai A, Ishihara K, Bewley JD.** 2003. Characterization of expression, and cloning, of β -D-xylosidase and α -L-arabinofuranosidase in developing and ripening tomato (*Lycopersicon esculentum* Mill.) fruit. *Journal of Experimental Botany* **54**, 2615–2622.
- Jung JH, Fouad WM, Vermerris W, Gallo M, Altpeter F.** 2012. RNAi suppression of lignin biosynthesis in sugarcane reduces recalcitrance for biofuel production from lignocellulose biomass. *Plant Biotechnology Journal* **10**, 1067–1076.
- Kaida R, Kaku T, Baba K, Oyadomari M, Watanabe T, Nishida K, Kanaya T, Shani Z, Shoseyov O, Hayashi T.** 2009. Loosening xyloglucan accelerates the enzymatic degradation of cellulose in wood. *Molecular Plant* **2**, 904–909.
- Karssen CM, Zagorski S, Kepczynski J, Groot SPC.** 1989. Key role for endogenous gibberellins in the control of seed germination. *Annals of Botany* **63**, 71–80.
- Kim B, Dale BE.** 2004. Global potential bioethanol production from wasted crops and crop residues. *Biomass and Bioenergy* **26**, 361–375.
- Kivi AR, Sartipnia N.** 2014. Cellulase, polygalacturonase and β -galactosidase activity in ripening raspberry (*Rubus caesius* L.) fruit. *Scholars Academic Journal of Biosciences* **2**, 177–180.
- Kozlova LV, Ageeva MV, Ibragimova NN, Gorshkova TA.** 2014. Arrangement of mixed-linkage glucan and glucuronoarabinoxylan in the cell walls of growing maize roots. *Annals of Botany* **114**, 1135–1145.
- Kramer M, Sanders R, Bolkan H, Waters C, Sheehy RE, Hiatt WR.** 1992. Postharvest evaluation of transgenic tomatoes with reduced levels of polygalacturonase: processing, firmness and disease resistance. *Postharvest Biology and Technology* **1**, 241–255.
- Langan P, Petridis L, O'Neill HM.** 2014. Common processes drive the thermochemical pretreatment of lignocellulosic biomass. *Green Chemistry* **16**, 63–68.
- Li J, Pan L, Fletcher JN, *et al.*** 2014. In vitro evaluation of potential bitterness-masking terpenoids from the Canada goldenrod (*Solidago canadensis*). *Journal of Natural Products* **77**, 1739–1743.
- Lionetti V, Francocci F, Ferraria S, Volpi C, Bellincampi D, Galletti R, D'Ovidi R, De Lorenzo G, Cervone F.** 2010. Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. *Proceedings of the National Academy of Sciences, USA* **107**, 616–621.
- Lynd LR, Laser MS, Bransby D, *et al.*** 2008. How biotech can transform biofuels. *Nature Biotechnology* **26**, 169–172.
- Mahadevan SA, Wi SG, Kim YO, Lee KH, Bae HJ.** 2011. In planta differential targeting analysis of *Thermotoga maritima* Cel5A and CBM6-engineered Cel5A for autohydrolysis. *Transgenic Research* **20**, 877–886.
- Mandal A, Datta AK.** 2014. Stability analysis of a high fibre yield and low lignin content “thick stem” mutant in tossa jute (*Corchorus olitorius* L.). *BioMed Research International* doi: 10.1155/2014/539869.
- Marriot PE, Sibout R, Lapierre C, Fangel JU, Willats WGT, Hofte H, Gómez LD, McQueen-Mason SJ.** 2014. Range of cell-wall alterations enhance saccharification in *Brachypodium distachyon* mutants. *Proceedings of the National Academy of Sciences, USA* **40**, 14601–14606.
- McCann MC, Carpita NC.** 2008. Designing the deconstruction of plant cell walls. *Current Opinion in Plant Biology* **11**, 314–320.
- McManus MT, Thompson DS, Merriman C, Lyne L, Osborne DI.** 1998. Transdifferentiation of mature cortical cells to functional abscission cells in bean. *Plant Physiology* **116**, 891–899.
- McQueen-Mason SJ, Cosgrove DJ.** 1995. Expansin mode of action on cell walls. Analysis of wall hydrolysis, stress relaxation, and binding. *Plant Physiology* **107**, 87–100.
- Meier H, Reid JSG.** 1982. Reserve polysaccharides other than starch in higher plants. In: Loevis FA, Tanner W, eds. *Encyclopaedia of plant physiology*, vol. **13A**. Berlin: Springer-Verlag, 418–471.
- Micheli F.** 2001. Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *Trends in Plant Science* **6**, 414–419.
- Mohanram S, Amat D, Choudhary J, Arora A, Nain L.** 2013. Novel perspectives for evolving enzyme cocktails for lignocellulose hydrolysis in biorefineries. *Sustainable Chemical Processes* **1**, 15.
- O'Donoghue EM, Huber DJ, Timpa JD, Erdos GW, Brecht JK.** 1994. Influence of avocado (*Persea americana*) Cxcellulase on the structural features of avocado cellulose. *Planta* **194**, 573–584.
- Ong RG, Chundawat SPS, Hodge DB, Keskar S, Dale BE.** 2014. Linking plant biology and pretreatment: understanding the structure and organization of the plant cell wall and interactions with cellulosic biofuel production. In: McCann MC, Buckeridge MS, Carpita NC, eds. *Plants and bioenergy*. New York: Springer, 231–253.
- Pauly M, Albersheim P, Darvill A, York WS.** 1999. Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. *The Plant Journal* **20**, 629–639.
- Pauly M, Keegstra K.** 2008. Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant Journal* **54**, 559–568.
- Peña MJ, Carpita NC.** 2004. Loss of highly branched arabinans and debranching of rhamnogalacturonan I accompany loss of firm texture and cell separation during prolonged storage of apple. *Plant Physiology* **135**, 1305–1313.
- Penning BW, Sykes RW, Babcock NC, *et al.*** 2014. Genetic determinants for enzymatic digestion of lignocellulosic biomass are independent of those for lignin abundance in a maize recombinant inbred population. *Plant Physiology* **165**, 1475–1487.
- Preece IA, MacDougall M.** 1958. Enzymic degradation of cereal hemicelluloses II. Pattern of pentosan degradation. *Journal of the Institute of Brewing* **64**, 489–500.
- Rajhi I, Yamauchi T, Takahashi H, *et al.*** 2011. Identification of genes expressed in maize root cortical cells during lysigenous aerenchyma formation using laser microdissection and microarray analyses. *New Phytologist* **190**, 351–368.
- Ray MJ, Brereton NJB, Shield I, Karp A, Murphy RJ.** 2012. Variation in cell wall composition and accessibility in relation to biofuel potential of short rotation coppice willow. *Bioenergy Research* **5**, 685–698.
- Redgwell RJ, Melton LD, Brasch DJ.** 1992. Cell wall dissolution in ripening kiwifruit (*Actinidia deliciosa*): solubilization of the pectic polymers. *Plant Physiology* **98**, 71–81.
- Reid JSG.** 1985. Structure and function in legume-seed polysaccharides. In: Brett C, Hillman JR, eds. *Biochemistry of the plant cell walls*. Cambridge: Cambridge University Press, 259–268. Renewable Fuels Association. 2014. Available at <http://www.ethanolrfa.org/>, Accessed on 22 November 2014.
- Ridley BL, O'Neill MA, Mohnen D.** 2001. Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* **57**, 929–967.
- Roberts JA, Elliott KA, Gonzalez-Carranza ZH.** 2002. Abscission, dehiscence, and other cell separation processes. *Annual Review in Plant Biology* **53**, 131–158.
- Rondeau-Mouro C, Defer D, Leboeuf E, Lahaye M.** 2008. Assessment of cell wall porosity in *Arabidopsis thaliana* by NMR spectroscopy. *International Journal of Biological Macromolecules* **42**, 83–92.
- Rose JK, Bennett AB.** 1999. Cooperative disassembly of the cellulose-xyloglucan network of plant cell walls: parallels between cell expansion and fruit ripening. *Trends in Plant Science* **4**, 176–183.
- Rosli HG, Civello PM, Martínez GA.** 2004. Changes in cell wall composition of three *Fragaria* × *Ananassa* cultivars with different softening rate during ripening. *Plant Physiology and Biochemistry* **42**, 823–831.
- Roulin S, Buchala AJ, Fincher GB.** 2002. Induction of (1→3,1→4)- β -D-glucan hydrolases in leaves of dark-incubated barley seedlings. *Planta* **215**, 51–59.
- Saab IN, Sachs MM.** 1996. A flooding-induced xyloglucan endo-transglycosylase homolog in maize is responsive to ethylene and associated with aerenchyma. *Plant Physiology* **112**, 385–391.
- Sampedro J, Cosgrove DJ.** 2005. The expansin superfamily. *Genome Biology* **6**, 242.
- Sañudo-Barajas JA, Labavitch J, Greve C, Osuna-Enciso T, Muy-Rangel D, Siller-Cepeda J.** 2009. Cell wall disassembly during papaya

softening: role of ethylene in changes in composition, pectin-derived oligomers (PDOs) production and wall hydrolases. *Postharvest Biology and Technology* **51**, 158–167.

Simmons BA, Loqué D, Ralph J. 2010. Advances in modifying lignin for enhanced biofuel production. *Current Opinion in Plant Biology* **13**, 312.

Smith CJ, Watson CF, Morris PC, et al. 1990. Inheritance and effect on ripening of antisense polygalacturonase genes in transgenic tomatoes. *Plant Molecular Biology* **14**, 369–379.

Sozzi GO, Fraschina AA, Navarro AA, Cascone O, Greve LC, Labatovich JM. 2002. α -L-Arabinofuranosidase activity during development and ripening of normal and ACC synthase antisense tomato fruit. *Hortscience* **37**, 564–566.

Stone BA, Clarke AE, eds. 1992. *Chemistry and biology of (1→3)- β -glucans*. Melbourne: La Trobe University Press.

Sumiyoshi M, Nakamura A, Nakamura H, Hakata M, Ichikawa H, Hirochika H, Ishii T, Satoh S, Iwai H. 2013. Increase in cellulose accumulation and improvement of saccharification by overexpression of arabinofuranosidase in rice. *PLoS ONE* **8**, e78269.

Taiz L, Honigman WA. 1976. Production of cell wall hydrolyzing enzymes by barley aleurone layers in response to gibberellic acid. *Plant Physiology* **58**, 380–386.

Tateishi A, Mori H, Watari J, Nagashima K, Yamaki S, Inoue H. 2005. Isolation, characterization, and cloning of α -L-arabinofuranosidase expressed during fruit ripening of Japanese pear. *Plant Physiology* **138**, 1653–1664.

Tiné MA, Cortelazzo AL, Buckeridge MS. 2000. Xyloglucan mobilisation in cotyledons of developing plantlets of *Hymenaea courbaril* L. (Leguminosae-Caesalpinoideae). *Plant Science* **154**, 117–126.

Toorop PE, Aelst AC, Hilhorst HWM. 1998. Endosperm cap weakening and endo- β -mannanase activity during priming of tomato (*Lycopersicon esculentum* cv. MoneyMaker) seeds are initiated upon crossing a threshold water potential. *Seed Science Research* **8**, 483–492.

Torres AF, van der Weijde T, Dolstra O, Visser RGF, Trindade LM. 2013. Effect of maize biomass composition on the optimization of diluted-acid pretreatments and enzymatic saccharification. *Bioenergy Research* **6**, 1038–1051.

Ueda J, Miyamoto K, Hashimoto M. 1996. Jasmonates promote abscission in bean petiole explants: its relationship to the metabolism of cell wall polysaccharides and cellulase activity. *Journal of Plant Growth Regulation* **15**, 189–195.

Van Acker R, Vanholme R, Storme V, Mortimer JC, Dupree P, Boerjan W. 2013. Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnology for Biofuels* **6**, 46.

Vanholme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W. 2012. Metabolic engineering of novel lignin in biomass crops. *New Phytologist* **196**, 978–1000.

Xu N, Zhang W, Ren S, et al. 2012. Hemicelluloses negatively affect lignocellulose crystallinity for high biomass digestibility under NaOH and H₂SO₄ pretreatments in *Miscanthus*. *Biotechnology for Biofuels* **5**, 58.

Yamaki S, Machida Y, Kakiuchi N. 1979. Changes in cell wall polysaccharides and monosaccharides during development and ripening of Japanese pear fruit. *Plant Cell Physiology* **20**, 311–321.

Yamauchi T, Shimamura S, Nakazono M, Mochizukic, T. 2013. Aerenchyma formation in crop species: a review. *Field Crops Research* **152**, 8–16.

Yoshioka H, Kashimura H, Kaneko K. 1995. β -D-Galactosidase and α -L-arabinofuranosidase activities during the softening of apples. *Journal of the Japanese Society for Horticultural Science* **63**, 871–878.