



**Gustavo Pagotto Borin**

**ESTUDOS GENÔMICOS DA EXPRESSÃO GÊNICA GLOBAL  
DO FUNGO FILAMENTOSO *Trichoderma reesei*  
CRESCIDO EM BAGAÇO E COLMO DE CANA-DE-AÇÚCAR**

**GENOMIC STUDIES OF GLOBAL GENE EXPRESSION  
OF FILAMENTOUS FUNGUS *Trichoderma reesei*  
GROWN IN BAGASSE AND CULM OF SUGARCANE**

**CAMPINAS/SP**

**2015**





**UNIVERSIDADE ESTADUAL DE CAMPINAS**

**INSTITUTO DE BIOLOGIA**

**Gustavo Pagotto Borin**

**ESTUDOS GENÔMICOS DA EXPRESSÃO GÊNICA GLOBAL DO FUNGO  
FILAMENTOSO *Trichoderma reesei* CRESCIDO EM BAGAÇO E COLMO DE CANA-DE-AÇÚCAR**

**GENOMIC STUDIES OF GLOBAL GENE EXPRESSION OF FILAMENTOUS FUNGUS  
*Trichoderma reesei* GROWN IN BAGASSE AND CULM OF SUGARCANE**

DISSERTAÇÃO apresentada ao Instituto de Biologia da UNICAMP como parte dos requisitos exigidos para a obtenção do título de Mestre em Genética e Biologia Molecular, na Área de Microbiologia.

DISSERTATION presented to the Biology Institute of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Genetics and Molecular Biology, in the area of Microbiology.

**Orientador: Prof<sup>o</sup>. Dr<sup>o</sup>. Gustavo Henrique Goldman**

**Coorientadora: Dr<sup>a</sup>. Juliana Velasco de Castro Oliveira**

Este exemplar corresponde à versão final da dissertação defendida pelo aluno Gustavo Pagotto Borin e orientado pelo Prof<sup>o</sup>. Dr<sup>o</sup>. Gustavo Henrique Goldman.

**CAMPINAS/SP**

**2015**

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

B644e Borin, Gustavo Pagotto. 1991-  
Estudos genômicos da expressão gênica global do fungo filamentoso  
*Trichoderma reesei* crescido em bagaço e colmo de cana-de-açúcar / Gustavo  
Pagotto Borin. – Campinas, SP : [s.n.], 2015.

Orientador: Gustavo Henrique Goldman.  
Coorientador: Juliana Velasco de Castro Oliveira.  
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de  
Biologia.

1. *Trichoderma reesei*. 2. Etanol. 3. Biomassa vegetal. 4. RNA-Seq. 5.  
Hidrólise enzimática. I. Goldman, Gustavo Henrique. II. Oliveira, Juliana Velasco  
de Castro. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

Informações para Biblioteca Digital

**Título em outro idioma:** Genomic studies of global gene expression of filamentous fungus  
*Trichoderma reesei* grown in bagasse and culm of sugarcane

**Palavras-chave em inglês:**

*Trichoderma reesei*

Ethanol

Plant biomass

RNA-Seq

Enzymatic hydrolysis

**Área de concentração:** Microbiologia

**Titulação:** Mestre em Genética e Biologia Molecular

**Banca examinadora:**

Gustavo Henrique Goldman [Orientador]

André Ricardo de Lima Damásio

Fernando Segato

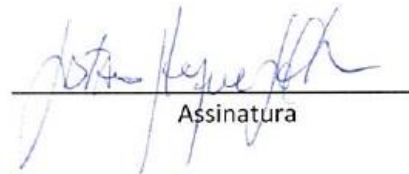
**Data de defesa:** 06-03-2015

**Programa de Pós-Graduação:** Genética e Biologia Molecular

Campinas, 06 de março de 2015.

**BANCA EXAMINADORA**

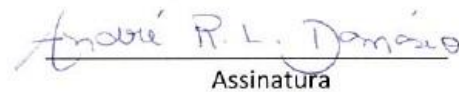
Profº. Drº. Gustavo Henrique Goldman (Orientador)



---

Assinatura

Profº. Drº. André Ricardo de Lima Damásio



---

Assinatura

Profº. Drº. Fernando Segato



---

Assinatura

Profº. Drº. Gonçalo Amarante Guimarães Pereira

---

Assinatura

Profº. Drº. Roberto do Nascimento Silva

---

Assinatura



## RESUMO

A parede celular vegetal é uma estrutura recalcitrante, composta por polissacarídeos complexos que podem ser quebrados em açúcares fermentáveis. A desconstrução desse material complexo pode ser feita por diversos tipos de enzimas hidrolíticas, que são produzidas naturalmente por uma variedade de microrganismos. Entre eles, o fungo *Trichoderma reesei* se destaca pela capacidade de produzir e secretar estas enzimas em grandes quantidades. Embora alguns trabalhos utilizando abordagens de proteômica e transcriptômica já tenham sido realizados com esse fungo, ainda não são conhecidos em detalhes os mecanismos moleculares responsáveis pela degradação da parede e a regulação gênica envolvida nesse sistema lignocelulolítico. O presente trabalho tem como objetivo principal a análise da expressão gênica global de *T. reesei*, crescido por 6, 12 e 24 horas em bagaço e colmo de cana-de-açúcar como fontes únicas de carbono, pela técnica de sequenciamento *high-throughput* de RNA (RNA-Seq). No transcriptoma de *T. reesei* foram identificadas sendo hiper-expressas as principais celulasas, hemicelulasas e proteínas acessórias relacionadas direta ou indiretamente com a desconstrução da parede vegetal. De modo geral, as celulasas e hemicelulasas apresentaram uma expressão maior do que outras enzimas, e o nível dos seus transcritos foi crescente ao longo do tempo tanto em colmo quanto no bagaço. A grande maioria dos genes de CAZymes e proteínas acessórias hiper-expressos foram compartilhados pelos dois substratos, o que demonstra que a estratégia usada por *T. reesei* para degradar a parede celular do colmo e do bagaço é similar. Adicionalmente, vários fatores de transcrição, proteínas de função desconhecida e transportadores supostamente envolvidos na assimilação dos açúcares liberados também foram hiper-expressos nas condições amostradas. Para validação do RNA-Seq, foi realizado PCR em tempo real de diversos genes hiper-expressos que codificam para enzimas hidrolíticas, reguladores transcricionais, proteínas acessórias e genes ainda não caracterizados. Para isso, a análise temporal foi ampliada para 30 minutos, 2, 4, 6, 12 e 24 horas de crescimento após o inóculo, o que permitiu uma análise mais detalhada da expressão desses genes. Como objetivo secundário, foi analisado o secretoma deste fungo e os açúcares concomitantemente liberados no sobrenadante. Estas análises indicaram que a

desconstrução da parede celular já se inicia dentro de 6 horas pós inoculo, com a liberação de monômeros (principalmente xilose e glicose) dos polissacarídeos e secreção de diversas CAZymes. Ensaio enzimáticos também foram realizados, mostrando atividades celulo e hemicelulolíticas. Assim, descrevemos pela primeira vez o arsenal de enzimas transcritas e secretadas por *T. reesei* RUT C30, desde pontos iniciais de crescimento, em bagaço explodido e colmo de cana-de-açúcar. Por fim, este trabalho também permitiu a identificação de vários genes, com função predita ou não, que podem abrir caminho para a descoberta de novos atuantes na resposta do fungo ao substrato lignocelulósico.



## ABSTRACT

Plant cell wall is a recalcitrant structure composed of complex polysaccharides which can be broken down into fermentable sugars. The deconstruction of this complex material can be made by a variety of hydrolytic enzymes which are naturally produced by a variety of microorganisms. Among them, stands out the fungus *Trichoderma reesei*, able to produce and secrete those enzymes in large quantities. Although some studies using transcriptomics and proteomics approaches have been performed with this fungus, the molecular mechanisms responsible for the degradation of the cell wall and gene regulation involved in this lignocellulosic system are not well known. This work has as main objective the analysis of global gene expression of *T. reesei* grown at 6, 12 and 24 hours in sugarcane bagasse and culm as sole carbon sources by high-throughput RNA sequencing technology (RNA-Seq). In the *T. reesei* transcriptome, it was identified the major cellulases, hemicellulases and accessory proteins directly or indirectly related to the deconstruction of plant cell wall. In general, cellulases and hemicellulases exhibited higher expression than other enzymes, and the level of their transcripts was increased over the time in both culm and bagasse cultures. Most of up-regulated CAZymes and accessory proteins were shared between the two substrates, which demonstrates the strategy used by *T. reesei* to degrade the bagasse and culm cell wall is similar. Furthermore, several transcription factors, proteins of unknown function and transporters supposedly involved in the assimilation of sugars were also up-regulated in the sampled conditions. To validate the RNA-Seq data, real-time PCR of several up-regulated genes encoding hydrolytic enzymes, transcriptional regulators, accessory proteins and proteins not yet characterized was carried out. The time points was extended to 30 min, 2, 4, 6, 12 and 24 hours of growth after inoculation, allowing a more detailed analysis of the expression of these genes. As a secondary objective, *T. reesei* secretome and the sugars released in the supernatant were analyzed. It was shown that the sugarcane cell wall deconstruction begins within the first 6 hours post inoculation, releasing sugar monomers (mainly xylose and glucose) from polysaccharides due to the secretion of several hydrolytic enzymes. Enzymatic assays were also performed, showing cellulosic and hemicellulosic activities. Finally, this is the first study showing the arsenal of enzymes transcribed and

secreted by *T. reesei* grown on steam exploded sugarcane bagasse and culm, at early time points. It was possible identify several genes, with predicted function or not, that can open new paths to discover novel players on the fungus response to lignocellulosic substrate.

## SUMÁRIO

Resumo.....	vii
Abstract.....	ix
<b>1. INTRODUÇÃO.....</b>	<b>1</b>
<b>2. REVISÃO BIBLIOGRÁFICA.....</b>	<b>5</b>
2.1 <i>Combustíveis fósseis e biocombustíveis.....</i>	5
2.2 <i>Etanol de primeira e segunda geração.....</i>	6
2.3 <i>Parede celular vegetal.....</i>	14
2.3.1 <i>Celulose.....</i>	15
2.3.2 <i>Hemicelulose.....</i>	16
2.3.3 <i>Lignina.....</i>	17
2.3.4 <i>Bagaço de cana-de-açúcar.....</i>	18
2.4 <i>Degradação da biomassa por enzimas e proteínas acessórias.....</i>	18
2.4.1 <i>Celulases.....</i>	20
2.4.2 <i>Proteínas acessórias e mecanismo oxidativo de degradação de                 celulose.....</i>	21
2.4.3 <i>Hemicelulases.....</i>	23
2.4.4 <i>Ligninases.....</i>	24
2.5 <i>Trichoderma reesei.....</i>	25
2.5.1 <i>Enzimas desconstrutoras da parede celular vegetal.....</i>	27
2.5.2 <i>Principais reguladores das enzimas lignocelulolíticas.....</i>	28
2.6 <i>Desafios biotecnológicos na produção do etanol 2G e biologia de sistemas.....</i>	31
<b>3. OBJETIVOS.....</b>	<b>35</b>
<b>4. MATERIAIS E MÉTODOS.....</b>	<b>37</b>
4.1 <i>Linhagens de fungo, meios de cultura e substratos de cana-de-açúcar.....</i>	37
4.2 <i>Indução do crescimento do fungo em bagaço e colmo de cana-de-açúcar.....</i>	38
4.3 <i>Extração do RNA total, tratamento com DNase e síntese de cDNA.....</i>	39
4.4 <i>Sequenciamento na plataforma Illumina® e análise das sequências.....</i>	40
4.5 <i>Validação da expressão dos genes por PCR em tempo real (qPCR).....</i>	41

4.6 Análise do secretoma.....	45
<b>5. RESULTADOS E DISCUSSÃO.....</b>	<b>47</b>
5.1 Transcriptômica de <i>T. reesei</i> crescido em biomassa da cana-de-açúcar.....	47
5.1.1 Perfil transcricional de CAZymes e proteínas acessórias.....	54
5.1.2 Perfil transcricional de transportadores hipotéticos de açúcar.....	71
5.1.3 Perfil transcricional de proteínas de função desconhecida.....	78
5.1.4 Perfil transcricional de fatores de transcrição.....	78
5.2 Análise do secretoma após sacarificação dos substratos.....	82
<b>6. CONCLUSÃO.....</b>	<b>85</b>
<b>7. REFERÊNCIAS.....</b>	<b>87</b>
<b>ANEXOS.....</b>	<b>109</b>
<b>ANEXO I.....</b>	<b>111</b>
<b>ANEXO II.....</b>	<b>147</b>
<b>ANEXO III.....</b>	<b>157</b>
<b>ANEXO IV.....</b>	<b>161</b>
<b>ANEXO V.....</b>	<b>165</b>
<b>ANEXO VI.....</b>	<b>171</b>

**Dedico este trabalho a meus pais e meus avós,  
pelo exemplo de luta e perseverança!**



## AGRADECIMENTOS

Ao Prof<sup>o</sup>. Dr<sup>o</sup>. Gustavo Henrique Goldman por ter me incentivado a fazer o mestrado, pelos ensinamentos e conhecimentos adquiridos, e pela oportunidade em ser seu orientado em um programa de pós-graduação muito bem conceituado.

À Dr<sup>a</sup>. Juliana Velasco de Castro Oliveira, pela orientação do meu projeto e direcionamento em todas as etapas do mestrado, desde a correção de relatórios até a realização dos experimentos. Agradeço pela paciência, atenção, conhecimentos compartilhados, exigência, discussões proveitosas e toda ajuda oferecida que me fez enfrentar e superar os desafios, que não foram poucos. Agradeço também por ter aprendido com você a ser um profissional com muito mais senso crítico e por ter sido seu orientando.

À Aline Tieppo, por todos os ensinamentos, dicas, instruções e ajuda na elaboração e realização dos experimentos. Agradeço ainda pelo companheirismo e por sempre deixar nossa ilha bem organizada, o que facilitou o bom andamento dos experimentos.

Agradeço ao Prof<sup>o</sup>. Dr<sup>o</sup>. André Ricardo de Lima Damásio, à Dr<sup>a</sup>. Cristiane Akemi e ao Prof<sup>o</sup>. Dr<sup>o</sup>. Gonçalo Guimarães, por todas as correções, direcionamento e discussão do meu projeto.

Ao Dr<sup>o</sup>. Diego Mauricio Riaño-Pachón, pela ajuda na análise e interpretação de dados, e também pelos ensinamentos em bioinformática.

Ao Prof<sup>o</sup>. Dr<sup>o</sup>. Marcos Silveira Buckeridge e à Dr<sup>a</sup>. Amanda Pereira, pela parceria na realização de experimentos e interpretação dos resultados.

À Prof<sup>a</sup>. Dr<sup>a</sup>. Camila Caldana, por toda ajuda na análise dos dados e conhecimentos compartilhados, e à sua técnica Juliana Aricetti pela ajuda na realização dos experimentos.

Ao Prof<sup>o</sup>. Dr<sup>o</sup>. Fabio Squina, pela colaboração e direcionamento nos experimentos.

À Dr<sup>a</sup>. Marina Camara Martins e à técnica Lúcia Wolf, pela colaboração e ajuda nos experimentos.

Agradeço ainda à Prof<sup>a</sup>. Dr<sup>a</sup>. Adriana Paes Leme e seu grupo do LNBio, ao laboratório CEFAP da USP, ao Dr<sup>o</sup>. Marcelo Carazzolle e à técnica Marcela Savoldi do laboratório do meu

orientador, por toda ajuda na realização nos experimentos. À secretária Lourdes da Pós por toda ajuda, paciência e atenção.

Agradeço à UNICAMP e ao CTBE pela minha formação acadêmica e por terem me proporcionado grande parte do conhecimento que tenho hoje. À FAPESP e ao CNPq, pelo suporte financeiro.

De modo especial, gostaria de agradecer a meus pais, Eduardo e Ângela, e meus irmãos, Tais e Felipe. Obrigado por terem me dado força, se interessarem pelo que eu faço e serem meu porto seguro. Pais, muito obrigado por me apoiarem. Vocês servem de exemplo para mim.

Quero agradecer também, de modo especial, à Camila Utsunomia e Thuanny Borges por todos os ensinamentos, experiências e conversas que tivemos e trocamos ao longo do tempo que trabalhamos juntos. Vocês foram muito importantes para minha formação.

Ao Renato Augusto, pela ajuda, parceria e amizade, mesmo trabalhando fora do laboratório, e à Rebeca Mariano, também pela amizade, ajuda nos experimentos e pela alegria contagiante. À Camila Sanchez e Eliane Santana, pela amizade, paciência, conversas, conselhos, ajuda nos experimentos, coletas à noite, e por deixarem o laboratório um ambiente muito mais alegre e prazeroso.

Aos meus amigos de laboratório, Roberta Dal’mas, Guilherme Zanini, Antônio Kaupert, Giovana Lopes, Aline Tieppo, Luciane Coerini, e a todos os outros amigos do CTBE, pela amizade, conselhos, conversas, ajuda nos experimentos e também pelos “barzinhos”. Vocês fazem e fizeram do laboratório uma segunda casa para mim. E apesar de não estarem mais trabalhando conosco, gostaria também de agradecer à Rafaella e Fabiana pelos ensinamentos e momentos bons vividos juntos. Agradeço também ao meu amigo Vitor Vidal com quem divido o apartamento, pela amizade, risadas, experiências compartilhadas e pela paciência em morar comigo.

Agradeço ainda a todos meus outros amigos que tornaram a realização do meu projeto de mestrado possível. Sou muito grato a vocês. Vocês são os irmãos que escolhemos ter. Agradeço ainda a Deus por ter cumprido mais uma etapa da minha vida e por Ele sempre ter me dado forças!



## LISTA DE ILUSTRAÇÕES

- Figura 1.** Fluxograma das etapas do processo de conversão de biomassa lignocelulolítica em etanol 2G. A hidrólise e fermentação podem ocorrer separadamente (SHF) ou de modo simultâneo (SSF e CBP). SHF: *separate hydrolysis and fermentation*; SSF: *simultaneous saccharification and fermentation*; CBP: *consolidated biomass bioprocessing* (Dashtban *et al.*, 2009).....12
- Figura 2.** Modelo simplificado da arquitetura da parede celular vegetal e os principais componentes da lignocelulose presente na parede: celulose, hemicelulose e lignina (Rubin, 2008).....15
- Figura 3.** Degradação de celulose por celulasas e por mecanismos oxidativos. As endoglucanases (EG) clivam as ligações glicosídicas na região menos organizada da celulose, conhecida como região amorfa. As extremidades redutoras e não-redutoras são então hidrolisadas progressivamente pelas celobiohidrolases (CBH1 e CBH2, respectivamente) com a participação dos domínios de ligação ao carboidrato (CBM). Dessa hidrólise, é formada como produto principal a celobiose, que é convertida em duas moléculas de glicose pelas  $\beta$ -glicosidases. Proteínas com atividade auxiliar da família AA9 (C<sub>1</sub>GH61 e C<sub>4</sub>GH61) também participam das desconstrução da celulose por um mecanismo oxidativo. As setas apontadas para a celulose indicam novos pontos de ataque para as celobiohidrolases (Horn *et al.*, 2012).....22
- Figura 4.** Mutagênese e seleção das cepas mutantes de *Trichoderma reesei*. Os agentes mutagênicos usados estão em negrito e os fenótipos para *screening* dos mutantes, em itálico. A cepa M7 está menos acentuada pois não se encontra mais disponível. UV: radiação por luz ultra-violeta; NTG: N-nitroguanidina; 2DG: 2-deoxiglicose (Le Crom *et al.*, 2009).....26
- Figura 5.** Curva de saturação das junções *exon-exon* de uma réplica da condição frutose 24 horas, mostrando a profundidade do sequenciamento do RNA-Seq. Observa-se que o sequenciamento atingiu a saturação, o que indica que o aumento de *reads* sequenciadas não implica em um maior número de junções *exon-exon*. O mesmo padrão foi obtido para as demais amostras.....49
- Figura 6.** Gráfico mostrando a correlação de Pearson ( $r^2$ ) obtida para as réplicas biológicas da condição frutose 24 horas. Uma alta correlação foi obtida entre as réplicas, o que demonstra a consistência dos dados do RNA-Seq. O mesmo também foi observado para as demais amostras (dados não mostrados).....50
- Figura 7.** Número de genes de *T. reesei* diferencialmente expressos em colmo (C) e bagaço (B) em relação à frutose, após 6, 12 e 24 horas (h) de crescimento.....51

**Figura 8.** Proporção de genes hiper-expressos identificados no transcriptoma de *T. reesei* crescido em colmo (C) e bagaço (B), após 6, 12 e 24 horas (h). Os genes foram agrupados em seis classes diferentes de acordo com sua função.....54

**Figura 9.** Hiper-expressão de genes de celulases (*cbh1* e *egl1*), hemicelulases (*xyn1*, *man1* e *axe1*), proteínas acessórias (*cip1* e *cel61a*), glicosil transferase (GT32, jgi|124397) e carboxilesterase (CE10, jgi|8573) no transcriptoma de *T. reesei* crescido em colmo (A) e bagaço (B) por 6, 12 e 24 horas (h). A expressão é mostrada em  $\log_2(\text{fold change})$ .....63

**Figura 10.** Expressão gênica da celobiohidrolase 2 (CBH2, jgi|122470) nas linhagens RUT C30 com o repressor *cre1* truncado e íntegro (*full*). A expressão foi normalizada pela expressão dos dois genes normalizadores *act* (actina) e *sar1* (GTPase). As amostras foram coletadas após 30 min ('), 2, 4, 6, 12 e 24 horas em colmo (C) e bagaço (B). Como condição controle, foi usado o crescimento em frutose após 24 horas. As barras de erro indicam o desvio padrão para as réplicas biológicas. As réplicas foram analisadas pelo *t test* não pareado (\*  $p < 0,05$ ; \*\*  $p < 0,01$ ;  $p < 0,001$ ).....71

**Figura 11.** Expressão gênica do transportador Crt1 (jgi|109243) nas linhagens RUT C30 com o repressor *cre1* truncado e íntegro (*full*). A expressão foi normalizada pela expressão dos dois genes normalizadores *act* (actina) e *sar1* (GTPase). As amostras foram coletadas após 30 min ('), 2, 4, 6, 12 e 24 horas em colmo (C) e bagaço (B). Como condição controle, foi usado o crescimento em frutose após 24 horas. As barras de erro indicam o desvio padrão para as réplicas biológicas. As réplicas foram analisadas pelo *t test* não pareado (\*  $p < 0,05$ ; \*\*  $p < 0,01$ ;  $p < 0,001$ ).....77

**Figura 12.** Expressão gênica da metiltransferase LAE1 (jgi|9778) nas linhagens RUT C30 com o repressor *cre1* truncado e íntegro (*full*). A expressão foi normalizada pela expressão dos dois genes normalizadores *act* (actina) e *sar1* (GTPase). As amostras foram coletadas após 30 min ('), 2, 4, 6, 12 e 24 horas em colmo (C) e bagaço (B). Como condição controle, foi usado o crescimento em frutose após 24 horas. As barras de erro indicam o desvio padrão para as réplicas biológicas. As réplicas foram analisadas pelo *t test* não pareado (\*  $p < 0,05$ ; \*\*  $p < 0,01$ ;  $p < 0,001$ ).....81

## LISTA DE TABELAS

- Tabela 1.** Relação dos oligonucleotídeos (oligos), temperatura de *melting* ( $T_m$ ) e tamanho esperado do produto de 20 genes analisados por PCR em tempo real. O ID JGI é referente ao número da proteína no banco de dados *Trichoderma reesei* RUT C30 v1.....43
- Tabela 2.** Número de *reads* geradas pelo sequenciamento do transcriptoma de *T. reesei*, porcentagem de *reads* mapeadas e cobertura do genoma. A porcentagem de *reads* que não tiveram um único mapeamento no genoma do fungo (*multi-mapeamento*) também é mostrada.....48
- Tabela 3.** Genes que codificam à CAZymes e proteínas acessórias hiper-expressos identificados no transcriptoma e secretoma de *T. reesei* crescido em bagaço (B) e colmo (C), após 6, 12 e 24 horas. As proteínas foram classificadas de acordo com o banco de dados CAZy, e a identificação (ID) de cada proteína é a disponível no banco de dados do JGI do *T. reesei* RUT C30 v1.0. O tipo de substrato preferencial de cada proteína, a presença de peptídeo sinal e se as proteínas foram encontradas no secretoma também são indicados (ver Anexo I).....56
- Tabela 4.** Transportadores de função conhecida e hipotéticos hiper-expressos no transcriptoma de *T. reesei* crescido em bagaço (B) e colmo (C), após 6, 12 e 24 horas (h). A identificação (ID) de cada proteína é a disponível no banco de dados do JGI do *T. reesei* RUT C30 v1.0. A expressão é apresentada como o  $\log_2$  do *fold change* e a presença dos segmentos transmembrânicos (TMH) também é mostrada. Em negrito, são mostrados dois genes com afinidade por xilose (ver discussão abaixo).....73



## LISTA DE ABREVIATURAS E SIGLAS

<b>AA</b>	<i>Auxiliary activity</i>
<b>BGL</b>	$\beta$ -glicosidase
<b>BLAST</b>	<i>Basic Local Alignment Search Tool</i>
<b>CE</b>	<i>Carbohydrate esterase</i>
<b>CAZy</b>	<i>Carbohydrate-Active Enzymes</i>
<b>CBM</b>	<i>Carbohydrate-binding module</i>
<b>CCR</b>	<i>Carbon catabolite repression</i>
<b>CDH</b>	<i>Cellobiose dehydrogenase</i>
<b>CBH</b>	Celobiohidrolase
<b>cm<sup>2</sup></b>	Centímetros quadrados
<b>CDS</b>	<i>Coding DNA sequence</i>
<b>CBP</b>	<i>Consolidated biomass processing</i>
<b>Ct</b>	<i>Cycle threshold</i>
<b>dbCAN</b>	<i>Database for Carbohydrate-active Enzyme Annotation</i>
<b>DP</b>	<i>Degree of polymerization</i>
<b>DNA</b>	<i>Deoxyribonucleic acid</i>
<b>dNTP</b>	<i>Deoxyribonucleotide triphosphate</i>
<b>cDNA</b>	DNA complementar
<b>DUF</b>	<i>Domain of unknown function</i>
<b>dsDNA</b>	<i>Double stranded DNA</i>
<b>EG</b>	Endoglucanase
<b>EC</b>	<i>Enzyme Commission</i>
<b>EDTA</b>	<i>Ethylenediamine tetraacetic acid</i>
<b>CO<sub>2</sub></b>	Gás carbônico
<b><i>cre1</i></b>	Gene do repressor catabólico de carbono 1
<b>GH</b>	<i>Glycosyl hydrolase</i>
<b>GT</b>	<i>Glycosyl transferase</i>
<b>g</b>	Gramma
<b>°C</b>	Graus centígrados
<b>GHG</b>	<i>Green house gases</i>
<b>ID</b>	<i>Identification number</i>
<b>ITS</b>	<i>Internal transcribed spacer</i>
<b>JGI</b>	<i>Joint Genome Institute</i>
<b>Kb</b>	Kilobases

<b>Kg</b>	Kilograma
<b>L</b>	Litro
<b>UV</b>	Luz ultravioleta
<b>LPMO</b>	<i>Lytic polysaccharide mono-oxygenase</i>
<b>MFS</b>	<i>Major Facilitator Superfamily</i>
<b>Mpa</b>	Megapascal
<b>MM</b>	Meio mínimo
<b>µL</b>	Microlitro
<b>µm</b>	Micrômetro
<b>mg</b>	Miligrama
<b>mL</b>	Mililitro
<b>mM</b>	Milimolar
<b>ng</b>	Nanograma
<b>NCBI</b>	<i>National Center for Biotechnology Information</i>
<b>NTG</b>	N-nitroguanidina
<b>Oligos</b>	Oligonucleotídeos
<b>PDA</b>	<i>Potato dextrose agar</i>
<b>pmol</b>	Picomol
<b>PCR</b>	<i>Polymerase chain reaction</i>
<b>PL</b>	<i>Polysaccharide lyase</i>
<b>pH</b>	Potencial hidrogeniônico
<b>1G</b>	Primeira geração
<b>qPCR</b>	<i>Quantitative PCR</i>
<b>RPKM</b>	<i>Reads per kilobase of transcript per million reads mapped</i>
<b>RNA</b>	<i>Ribonucleic acid</i>
<b>RNAi</b>	RNA de interferência
<b>RNA-Seq</b>	<i>RNA sequencing</i>
<b>rpm</b>	Rotações por minuto
<b>2G</b>	Segunda geração
<b>SHF</b>	<i>Separate hydrolysis and fermentation</i>
<b>SSF</b>	<i>Simultaneous saccharification and fermentation</i>
<b>ssDNA</b>	<i>Single stranded DNA</i>
<b>TMH</b>	<i>Transmembrane helice</i>
<b>TCDB</b>	<i>Transporter Classification Database</i>
<b>TAE</b>	<i>Tris-acetate-EDTA</i>

## 1. INTRODUÇÃO

Os combustíveis fósseis são a principal fonte de energia utilizada mundialmente, representando mais da metade do suprimento energético global (Goldemberg, 2007; World Energy Council, 2013). O aumento do preço dos combustíveis fósseis e a possibilidade de esgotamento das suas reservas, junto com o aumento da emissão de gás carbônico (CO<sub>2</sub>), aumentaram o interesse pelas fontes renováveis de energia, especialmente os biocombustíveis, como biogás, biodiesel e etanol (Leite *et al.*, 2009; Clark *et al.*, 2012; dos Santos Castro *et al.*, 2014). Os biocombustíveis são uma alternativa promissora para substituir os combustíveis fósseis por diversos motivos, dentre eles podemos citar: a redução da emissão de gases do efeito estufa, produção por tecnologias bem consolidadas e promoção de economias locais (Fukuda *et al.*, 2009; Banerjee *et al.*, 2010b; Hasunuma & Kondo, 2012). Em 2013, estima-se que foram produzidos mais de 87 bilhões de litros de etanol no mundo, sendo que os Estados Unidos e o Brasil foram os maiores produtores, utilizando o milho e a cana-de-açúcar (cana) como matérias-primas, respectivamente (REN21, 2014). O etanol produzido a partir dessas culturas é denominado de primeira geração (1G) e, no Brasil, é obtido através da fermentação do caldo extraído da cana pela levedura *Saccharomyces cerevisiae* (Amorim *et al.*, 2011). O etanol derivado da cana é mais vantajoso que o produzido utilizando o milho por ter um maior rendimento de combustível por hectare, apresentar um custo de produção mais baixo e ter um balanço energético mais positivo (Goldemberg, 2007; Amorim *et al.*, 2011). Diante da grande disponibilidade de matéria-prima, tecnologia de menor custo e condições de expandir a área de plantio de cana sem colocar em risco culturas voltadas para produção de alimentos, o Brasil tem potencial para ser o maior produtor mundial de etanol (Leite *et al.*, 2009; Triana, 2011). Adicionalmente, estudos sugerem que se o bagaço gerado como subproduto no processo de produção de etanol e açúcar fosse convertido em etanol lignocelulósico, haveria um aumento significativo no rendimento deste biocombustível por hectare (Macedo, 2005).

O etanol lignocelulósico ou de segunda geração (2G) é obtido a partir da fermentação dos açúcares liberados da degradação da parede celular vegetal (Ho *et al.*, 2014). No Brasil, o bagaço é utilizado principalmente para geração de energia nas usinas (Amorim *et al.*, 2011).

Por ser um resíduo lignocelulósico muito abundante (280 Kg/tonelada de cana), não baseado em uma cultura alimentar, o bagaço tem grande potencial para ser utilizado como matéria-prima para produção do etanol 2G (Lin & Tanaka, 2006; Ragauskas *et al.*, 2006), o que justifica a grande quantidade de estudos sendo realizados com essa biomassa (Wanderley *et al.*, 2013; Khuong *et al.*, 2014; Li *et al.*, 2014). A lignocelulose presente na parede vegetal é composta basicamente por celulose, hemicelulose e lignina (Ho *et al.*, 2014; Kricka *et al.*, 2014), que ficam intimamente interligados de modo a conferir recalcitrância à hidrólise da parede (Adsul *et al.*, 2011). Assim, faz-se necessário um pré-tratamento e a sacarificação da biomassa para que açúcares monoméricos sejam liberados e fermentados (Rubin, 2008). Existem diversos pré-tratamentos físicos, químicos e biológicos que podem ser utilizados de acordo com o tipo de biomassa (Sarkar *et al.*, 2012; Ho *et al.*, 2014), e a sacarificação dos polissacarídeos pode ser feita quimicamente ou pelo uso de enzimas produzidas por microrganismos (Lee & Lavoie, 2013). Essas duas etapas representam desafios biotecnológicos para a produção do etanol 2G, uma vez que devem ser eficientes e ter baixo custo para que este biocombustível seja economicamente viável (Klein-Marcuschamer *et al.*, 2012; Kricka *et al.*, 2014). Quanto às estratégias de fermentação, os açúcares podem ser fermentados separadamente da sacarificação da biomassa pré-tratada ou em um único reator com a fermentação ocorrendo à medida que os açúcares são liberados (Balat, 2011; Menon & Rao, 2012; Kricka *et al.*, 2014).

Fungos filamentosos, como as espécies *Trichoderma reesei* e *Aspergillus niger*, são ótimos desconstrutores de matéria orgânica e naturalmente possuem um arsenal de enzimas lignocelulolíticas capazes de quebrar a parede vegetal (Lynd *et al.*, 2002; Xie *et al.*, 2014). O *T. reesei* é utilizado amplamente na indústria para produção de celulasas e hemicelulasas nativas, mas também para produção de proteínas heterólogas, e tem servido como fungo modelo para estudos da degradação de lignocelulose (Shoemaker *et al.*, 1983; Penttilä *et al.*, 1986; Teeri *et al.*, 1987; Peterson & Nevalainen, 2012). A cepa hipercelulolítica *T. reesei* RUT C30, utilizada neste trabalho, foi obtida após três etapas de mutagênese (Peterson & Nevalainen, 2012) e se mostrou uma das melhores cepas produtoras de celulasas (Eveleigh & Montenecourt, 1979; dos Santos Castro *et al.*, 2014), dentre outras razões, por causa da



derrepressão catabólica por carbono conferida pelo truncamento do fator de transcrição *cre1* (Ilmen *et al.*, 1996).

A análise do genoma de *T. reesei* (Martinez *et al.*, 2008; Le Crom *et al.*, 2009) permitiu a identificação de 228 genes envolvidos direta ou indiretamente na sacarificação de biomassa (Häkkinen *et al.*, 2012), sendo que desse total 195 são genes que codificam a hidrolases (GHs, *glycosyl hydrolases*), 22 a esterases (CEs, *carbohydrate esterases*), 5 a liases (PLs, *polysaccharide lyases*) e 6 a proteínas com atividade auxiliar (AAs, *auxiliary activities*) (Häkkinen *et al.*, 2012; Levasseur *et al.*, 2013). Essas proteínas são classificadas e descritas no banco de dados CAZy (***Carbohydrate Active Enzyme***), de acordo com o mecanismo de ação, e a similaridade de sequência dos aminoácidos e estrutura entre as proteínas (Cantarel *et al.*, 2009). *T. reesei* produz pelo menos duas celobiohidrolases (CBH1/CEL7A e CBH2/CEL6A), cinco endoglucanases (EGs), duas  $\beta$ -glicosidases e outras celulases preditas (Häkkinen *et al.*, 2012). As CBHs e EGs compõem juntas até 90% do número total de enzimas secretadas, sendo que CBH1 e CBH2 constituem até 60% e 20% desse total, respectivamente (Rahman *et al.*, 2009). Além dessas celulases, outras proteínas envolvidas com a degradação de celulose também foram descobertas, como swolenina (Saloheimo *et al.*, 2002), hidrofobinas e as monooxigenases recentemente agrupadas na família AA9 (Häkkinen *et al.*, 2012). Apesar de ter um perfil mais celulolítico, *T. reesei* também produz hemicelulases, como  $\beta$ -xilanasas, mananase, acetilxilano esterase,  $\alpha$ -glucuronidase,  $\alpha$ -galactosidases e  $\beta$ -xilosidase (Martinez *et al.*, 2008; Häkkinen *et al.*, 2012). Assim, com o objetivo de hidrolisar biomassas lignocelulósicas, várias de suas enzimas estão sendo utilizadas para o desenvolvimento e comercialização de coquetéis enzimáticos, a exemplo do *Cellic CTec3*<sup>®</sup> (Novozyme A/S, Dinamarca) e *Accellerase*<sup>®</sup> 1500 (DuPont<sup>™</sup>, Estados Unidos) (Novozymes A/S, 2012; DuPont, 2013; Hassan *et al.*, 2013; Kolasa *et al.*, 2014).

A expressão destas enzimas lignocelulolíticas é induzida por, pelo menos, quatro moléculas diferentes (soforose, xilose, xilobiose e lactose), mas nenhuma delas tem capacidade de iniciar a expressão de todas as principais celulases e hemicelulases (Stricker *et al.*, 2008). A regulação a nível transcricional dos genes dessas enzimas é feita, principalmente, pelos ativadores Xyr1, Ace2, Ace3 e o complexo HAP2/3/5, e pelos

repressores Cre1 e Ace1, sendo Xyr1 e Cre1 os maiores reguladores positivo e negativo, respectivamente (Häkkinen *et al.*, 2014; Lichius *et al.*, 2014; Silva-Rocha *et al.*, 2014). Tendo em vista a grande importância dos fatores de transcrição na regulação gênica, vários estudos vêm sendo realizados para a identificação de novos reguladores (Häkkinen *et al.*, 2014; Karimi Aghcheh *et al.*, 2014), mas o entendimento global do processo de sacarificação realizados pelos fungos ainda não é totalmente conhecido (Silva-Rocha *et al.*, 2014). Nesse contexto, a união de abordagens “ômicas”, como transcriptômica e proteômica, pode ser uma ferramenta valiosa para auxiliar no entendimento básico do processo de degradação destas biomassas, o que poderá refletir em novas perspectivas para o desenvolvimento de coquetéis enzimáticos para a produção do etanol 2G, em escala industrial.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 Combustíveis fósseis e biocombustíveis

O crescimento da população mundial e da industrialização aumenta, a cada ano, a demanda por alimento, combustível e energia (Nass *et al.*, 2007; Liu *et al.*, 2012). Atualmente, os combustíveis fósseis como carvão, petróleo e gás, são as principais fontes de energia, representando mais de 80% do suprimento energético global (Goldemberg, 2007; World Energy Council, 2013). A possibilidade de esgotamento dessas reservas, o aumento do preço e alto grau de dependência econômica dos combustíveis fósseis, e as mudanças climáticas tornaram o desenvolvimento de fontes renováveis de energia, entre elas os biocombustíveis, uma prioridade global (Leite *et al.*, 2009; Clark *et al.*, 2012; dos Santos Castro *et al.*, 2014).

Em termos de bioenergia, ou seja, a energia contida em produtos derivados de materiais biológicos brutos (biomassa), etanol, biodiesel e biogás são os três produtos mais energéticos (Yuan *et al.*, 2008). O etanol pode ser obtido a partir de culturas voltadas para produção de alimentos com alto teor de açúcar e amido, como a cana-de-açúcar (cana) e o milho, respectivamente, ou a partir de resíduos lignocelulósicos (biomassa lignocelulósica). O biodiesel é produzido a partir da combinação do óleo diesel com o óleo extraído de sementes (soja, canola, palma, colza, girassol), algas ou resíduos alimentícios (gordura bovina e óleo de cozinha usado) que é transesterificado para remoção do glicerol (Yuan *et al.*, 2008). O biogás é o biocombustível produzido a partir de uma grande variedade de resíduos orgânicos, como restos de alimentos, cascas de frutas e gordura animal. Esses resíduos são então convertidos em metano, hidrogênio ou monóxido de carbono através do processo de gasificação (Yuan *et al.*, 2008; Ho *et al.*, 2014; Kiran *et al.*, 2014).

Atualmente, os biocombustíveis representam 10% do suprimento energético global (IEA, 2013; Ho *et al.*, 2014; The World Bank, 2014). Em um cenário otimista, Hasunuma e Kondo (2012) afirmaram que a substituição dos combustíveis fósseis por biocombustíveis pode diminuir a emissão global de dióxido de carbono (CO<sub>2</sub>) de 60 a 90%. Estudos sugerem que sua produção é capaz de promover economias locais sem entrar em conflito com culturas destinadas à produção de alimentos (Fukuda *et al.*, 2009; Banerjee *et al.*, 2010b).

Anualmente, são produzidos mais de 100 bilhões de litros de biocombustíveis, o que representa apenas 2,7% da energia total usada no setor de transporte (Kricka *et al.*, 2014). Assim, se um dos objetivos é tentar diminuir o consumo dos combustíveis fósseis, é de suma importância o aumento da produção desses combustíveis renováveis. Dentre eles, a produção do etanol a partir da cana detém a melhor tecnologia para produzir biocombustível de forma eficiente e econômica. O processo de absorção do CO<sub>2</sub> e conversão em biomassa vegetal pela fotossíntese juntamente com a fermentação alcoólica têm possibilitado a produção de um combustível limpo e de alta qualidade, com até 93% da energia original encontrada na sacarose da cana (Amorim *et al.*, 2011).

## **2.2 Etanol de primeira e segunda geração**

A produção de etanol pode ser classificada em dois tipos de acordo com a tecnologia e matéria-prima empregadas. O etanol 1G é obtido principalmente da sacarose presente na cana e do amido do milho, enquanto que o etanol 2G é obtido a partir de biomassa lignocelulósica (Ho *et al.*, 2014). Apesar do etanol 2G ainda não ser comercializado em larga escala devido a vários gargalos biotecnológicos que precisam ser superados (Banerjee *et al.*, 2010b), ele representa uma alternativa promissora para o desenvolvimento de um combustível sustentável para ser utilizado em escala global.

Em 2013, foram produzidos mais de 87 bilhões de litros de etanol no mundo todo, sendo que os Estados Unidos e o Brasil foram os maiores produtores, com 50,3 e 25,5 bilhões de litros, respectivamente (REN21, 2014). Hoje o Brasil responde por cerca de 30% da produção mundial de etanol, e neste contexto o programa nacional ProÁlcool teve participação importante para a consolidação da produção eficiente e econômica deste biocombustível. Criado em 1975, o programa pretendia estimular a produção brasileira de álcool a fim de atender o mercado interno de combustíveis e diminuir o consumo do petróleo que na época estava altamente valorizado (Leite *et al.*, 2009). Em 2003, com a introdução dos automóveis *flex* que podem rodar tanto com uma mistura de gasolina e etanol anidro quanto etanol hidratado, houve uma grande expansão do consumo de etanol 1G, chegando a superar o consumo da gasolina em 2008 (Amorim *et al.*, 2011). Considerando

aspectos ambientais e comparando o ciclo completo de produção da cana e do milho, há uma redução de 90 e 30% na emissão de gases do efeito estufa (GHG, do inglês *green house gases*) do etanol obtido a partir da cana e do milho, respectivamente, em relação à gasolina (Macedo *et al.*, 2008).

Devido a seu uso para produção de açúcar e etanol, o Brasil é hoje o principal produtor mundial de cana, com uma área estimada em 9.130,1 hectares de plantação (Conab, 2014), que devem produzir aproximadamente 671,69 milhões de toneladas na safra de 2014/2015 (Conab, 2014). Para produzir o etanol, a cana é colhida, lavada, moída e o caldo extraído é tratado para remoção de impurezas (Amorim *et al.*, 2011). Nesta etapa de moagem e extração do caldo é gerado o bagaço da cana, que é aproveitado pelas usinas para geração de energia (Goldemberg & Guardabassi, 2010) e também como ração animal (Banerjee *et al.*, 2010b). Dependendo da demanda do mercado, parte do caldo é aquecido, transformado em melaço e usado para produção de açúcar. O caldo restante junto com o melaço residual formam o mosto, que segue então para dornas de fermentação onde é fermentado por leveduras industriais tolerantes a condições de estresses presentes neste processo (baixo pH, altas concentrações de etanol, variação de temperatura, microrganismos contaminantes), como a espécie *Saccharomyces cerevisiae*. Após um ciclo de fermentação, as leveduras recebem um tratamento com ácido sulfúrico (H<sub>2</sub>SO<sub>4</sub>) para remoção da contaminação de bactérias e retornam para as dornas de fermentação onde iniciarão um novo ciclo. O etanol gerado como produto da fermentação é destilado e convertido em etanol anidro ou hidratado para ser comercializado, principalmente, como combustível ou aditivo para carros (Amorim *et al.*, 2011).

De forma similar ao etanol brasileiro, o etanol 1G obtido a partir do milho nos Estados Unidos também passa pelos processos de moagem, fermentação alcoólica e destilação, mas por este ser derivado do amido, é necessária uma etapa de conversão desse polímero em açúcares simples antes da fermentação. Para tanto, são adicionadas amilases que quebram as ligações  $\alpha$ -1,4 de glicose presentes no amido. Nesse processo de produção de etanol também são gerados produtos de alto valor agregado, como o glúten utilizado nos suplementos para animais (Bothast & Schlicher, 2005).

Em relação às vantagens do etanol produzido a partir da cana e do milho, o etanol brasileiro é mais competitivo que o dos Estados Unidos, uma vez que o Brasil ainda tem condições de expandir a área de plantio de cana sem colocar em risco culturas voltadas para produção de alimentos (Leite *et al.*, 2009). Além disso, estima-se que um hectare de área plantada possa produzir até 79,5 toneladas de cana, enquanto que o rendimento do milho nos Estados Unidos é de apenas 9,9 toneladas por hectare. Em termos de produção de combustível por hectare, isso representa 7.200 e 3.800 litros de etanol 1G da cana e do milho, respectivamente (Ho *et al.*, 2014). Assim, o etanol brasileiro é economicamente mais viável: enquanto o derivado do milho apresenta um custo de U\$0,45/L, entre outros motivos, por causa da etapa adicional da quebra do amido, o etanol da cana custa U\$0,28/L, um custo 40% menor em relação ao etanol americano (Andreoli & Souza, 2007). Por fim, o balanço energético do etanol da cana (9,3:1) também é mais positivo que o do etanol do milho (1,3:1), ou seja, para cada unidade de energia de combustível fóssil utilizada são produzidas 9,3 e 1,3 unidades de energia para o etanol da cana e do milho, respectivamente (Goldemberg, 2007; Amorim *et al.*, 2011).

Diante dessas vantagens da produção de etanol brasileiro a partir da cana, o Brasil se tornou uma referência mundial na produção deste combustível, e sendo um país que possui uma tecnologia mais barata e uma grande disponibilidade de matéria-prima (cana), tem potencial para ser o maior produtor mundial (Leite *et al.*, 2009; Triana, 2011). Em um trabalho encomendado pelo governo brasileiro, Leite e colaboradores (2009) sugeriram que para o Brasil suprir 5% do uso do petróleo mundial (em 2025, projetado para um total de 1,7 trilhão de litros), o país teria que produzir um volume em torno de 102 bilhões de litros de etanol. Para atingir esta produção fazendo uso da tecnologia atualmente disponível, haveria a necessidade de um aumento de 17 milhões de hectares de plantação de cana, que ocuparia então uma área cultivável próxima à da soja. Contudo, esta área cultivada pode ser diminuída consideravelmente (29%) se 8% do bagaço da cana for utilizado para a produção do etanol 2G (Leite *et al.*, 2009). A palha da cana gerada também como subproduto da produção sucro-energética poderia ser utilizada junto com o bagaço para o aumento do rendimento do etanol 2G (Leite *et al.*, 2009).

Os biocombustíveis 2G são uma alternativa em potencial aos combustíveis fósseis, uma vez que são derivados de fontes abundantes de energia e podem superar a limitação na produção do etanol 1G (Antizar-Ladislao & Turrion-Gomez, 2008) mantendo as vantagens energética, econômica e ambiental citadas acima. O etanol 2G pode ser derivado de biomassa vegetal formada por material lignocelulósico complexo, como resíduos florestais (casca de eucalipto) e da agricultura (bagaço e palha de cana, palha de milho e de trigo), resíduos sólidos industriais e municipais (papelão, resto de frutas e vegetais), além de culturas dedicadas à produção desse tipo de energia (como as gramíneas *switchgrass* e *Miscanthus sp.*, e as plantas lenhosas *willow*, *poplar* e pinheiro) (Yuan *et al.*, 2008; Ho *et al.*, 2014; Kricka *et al.*, 2014). A lignocelulose presente na parede celular vegetal é composta basicamente por celulose, hemicelulose e lignina (Ho *et al.*, 2014; Kricka *et al.*, 2014), e para a produção do etanol 2G é necessária a desconstrução destes polissacarídeos por processos de pré-tratamento e sacarificação, e liberação de açúcares que serão convertidos em biocombustível pela fermentação (Rubin, 2008).

Estes pré-tratamentos podem ser físicos, químicos, físico-químicos e/ou biológicos, e têm como objetivos: **i)** diminuir a recalcitrância da parede celular conferida principalmente pela lignina; **ii)** aumentar a hidrólise da hemicelulose e celulose; **iii)** produzir açúcares fermentáveis e evitar a perda/degradação destes açúcares; **iv)** limitar a formação de produtos inibitórios para a fermentação; e **v)** reduzir a demanda de energia e os custos (Sarkar *et al.*, 2012; Ho *et al.*, 2014). Para que o preço do etanol 2G seja economicamente competitivo é fundamental a escolha correta do pré-tratamento mais adequado ao tipo de biomassa a ser utilizada (Ho *et al.*, 2014). Dentre os pré-tratamentos físicos, pode-se citar a moagem seca ou úmida, moagem com bolas ou facas, pirólise (temperaturas acima de 300°C) e o uso de ultrassom; para os tratamentos químicos e físico-químicos, podem ser utilizados ácidos (H<sub>2</sub>SO<sub>4</sub>), bases (hidróxido de sódio, potássio, cálcio e amônio), explosão a vapor combinada ou não com amônia líquida, explosão por CO<sub>2</sub>, água quente pressurizada e oxidação. Para degradação da hemicelulose e lignina, fungos (do grupo *white rot*) podem ser empregados no pré-tratamento biológico para liberar a celulose, porém esse pré-tratamento

não é muito rentável e, por isso, pouco utilizado (Sarkar *et al.*, 2012; Ho *et al.*, 2014; Xu & Huang, 2014).

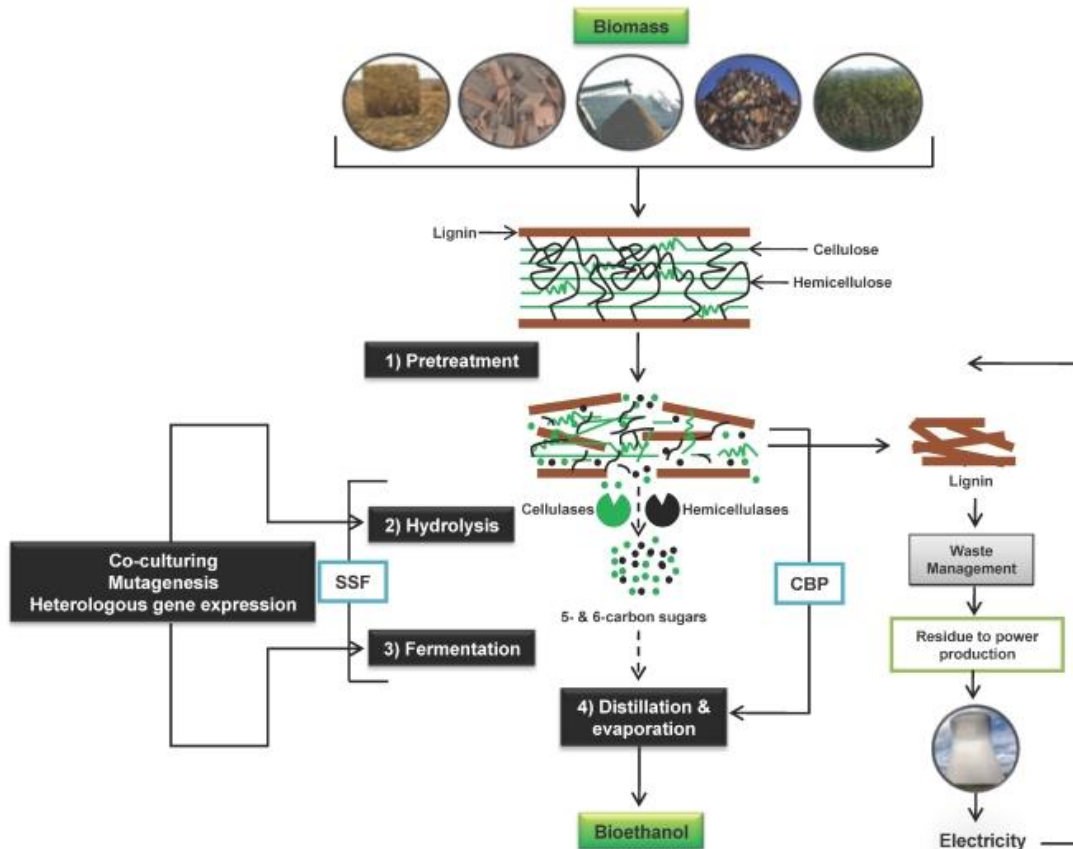
A explosão a vapor é um dos pré-tratamentos mais utilizados, e consiste em submeter a biomassa recalcitrante a um vapor de alta pressão (0,69–4,83 Mpa), em altas temperaturas (160-260°C), e diminuir rapidamente a pressão até a pressão atmosférica. Essa mudança drástica de pressão faz com que o vapor no interior da matriz lignocelulósica comece a expandir, separando as fibras em um processo de descompressão explosiva. Nesse pré-tratamento, o H<sub>2</sub>SO<sub>4</sub> ou o gás dióxido de enxofre (SO<sub>2</sub>) podem ser adicionados para maior hidrólise da hemicelulose. Além disso, as altas temperaturas do processo levam à hidrólise de grupos acetil da hemicelulose, e o ácido acético formado catalisa a hidrólise da hemicelulose, em um processo conhecido como auto-hidrólise (Sarkar *et al.*, 2012; Oliveira *et al.*, 2013; Xu & Huang, 2014). O alto rendimento de xilose (45-65%) e o uso limitado de químicos são algumas vantagens desse pré-tratamento. Como desvantagens, tem-se a formação de produtos inibitórios da fermentação e a destruição incompleta da lignocelulose, que pode precipitar os componentes solúveis da lignina e tornar a biomassa menos acessível à ação das enzimas (Sarkar *et al.*, 2012; Oliveira *et al.*, 2013; Ho *et al.*, 2014; Xu & Huang, 2014). Avaliando a influência do pré-tratamento da explosão a vapor no bagaço da cana, Sendelius (2005) observou que o maior rendimento de glicose (86,3%) e xilose (72%) ocorreu quando o bagaço era aquecido a 180°C, durante 5 minutos utilizando SO<sub>2</sub>. Posteriormente, o hidrolisado fermentado rendeu 80% de etanol com base em valores teóricos.

Após o pré-tratamento adequado, a celulose e hemicelulose estão mais acessíveis à hidrólise e passam pela etapa de sacarificação, onde são produzidos açúcares monoméricos (pentoses e hexoses) que podem ser fermentados. Essa etapa pode ser feita quimicamente pela adição de H<sub>2</sub>SO<sub>4</sub> ou bioquimicamente pelo uso de celulases e hemicelulases produzidas por microrganismos (Lee & Lavoie, 2013). Apesar da hidrólise ácida ser mais barata, o rendimento de produção de pentoses (xilose) e hexoses (glicose) é menor e seu uso não é visto como ambientalmente amigável. Por outro lado, o uso de enzimas hidrolíticas específicas resultam em um alto rendimento de açúcares, mas a necessidade de uma grande quantidade dessas enzimas para conversão da biomassa é um dos principais gargalos para a



produção do etanol 2G, uma vez que o torna economicamente inviável (Klein-Marcuschamer *et al.*, 2012). Alguns coquetéis enzimáticos já estão sendo comercializados, dentre eles *Cellic CTec3*<sup>®</sup> (Novozyme A/S, Dinamarca) e *Accellerase*<sup>®</sup> 1500 (DuPont<sup>™</sup>, Estados Unidos) (Novozymes A/S, 2012; DuPont, 2013; Hassan *et al.*, 2013; Kolasa *et al.*, 2014), mas esforços vêm sendo realizados para reduzir os custos de produção e aumentar a eficiência dos mesmos. Dentre eles podemos citar: o bioengenhamento de microrganismos capazes de produzir enzimas lignocelulolíticas em grandes quantidades; o uso de matérias-primas abundantes e de baixo custo; integração dos processos de segunda geração com a infraestrutura já existente e utilizada na primeira geração; além do aproveitamento de todos os componentes da biomassa e dos subprodutos gerados ao longo de toda cadeia produtiva do etanol (Marx *et al.*, 2013; Ho *et al.*, 2014).

A Figura 1 mostra um fluxograma do processo de produção do etanol 2G, com as etapas de conversão de biomassa lignocelulósica em biocombustível e as estratégias de hidrólise e fermentação dos açúcares. Para redução do custo de produção, a lignina extraída da lignocelulose também pode ser aproveitada para gerar calor e energia elétrica nas usinas (Dashtban *et al.*, 2009).



**Figura 1.** Fluxograma das etapas do processo de conversão de biomassa lignocelulolítica em etanol 2G. A hidrólise e fermentação podem ocorrer separadamente (SHF) ou de modo simultâneo (SSF e CBP). SHF: *separate hydrolysis and fermentation*; SSF: *simultaneous saccharification and fermentation*; CBP: *consolidated biomass bioprocessing* (Dashtban et al., 2009).

A biomassa lignocelulósica pré-tratada pode ser convertida em etanol 2G por três estratégias de fermentação: hidrólise separada da fermentação (SHF, *separate hydrolysis and fermentation*), sacarificação e fermentação simultâneas (SSF, *simultaneous saccharification and fermentation*) e processamento integrado de biomassa (CBP, *consolidated biomass processing*) (Figura 1). A primeira estratégia é um processo convencional de dois passos: no primeiro, a biomassa é hidrolisada por enzimas para formar açúcares redutores, e no segundo, os açúcares redutores são fermentados e convertidos em etanol por leveduras, como as do gênero *Saccharomyces* e *Pichia*. Os dois passos são realizados em compartimentos separados, o que permite ajustar as condições ótimas de temperatura e pH para cada etapa. A hidrólise de celulose pelas celulasas, por exemplo, é mais eficiente nas

temperaturas entre 50 e 60°C, enquanto que a fermentação é normalmente realizada entre 25 e 35°C devido ao metabolismo das leveduras. Uma das principais desvantagens desse processo é a inibição das enzimas hidrolíticas pela liberação de glicose e celobiose, o que requer a adição de mais enzima ( $\beta$ -glicosidase) para hidrólise dos açúcares inibitórios (Balat, 2011; Menon & Rao, 2012; Kricka *et al.*, 2014).

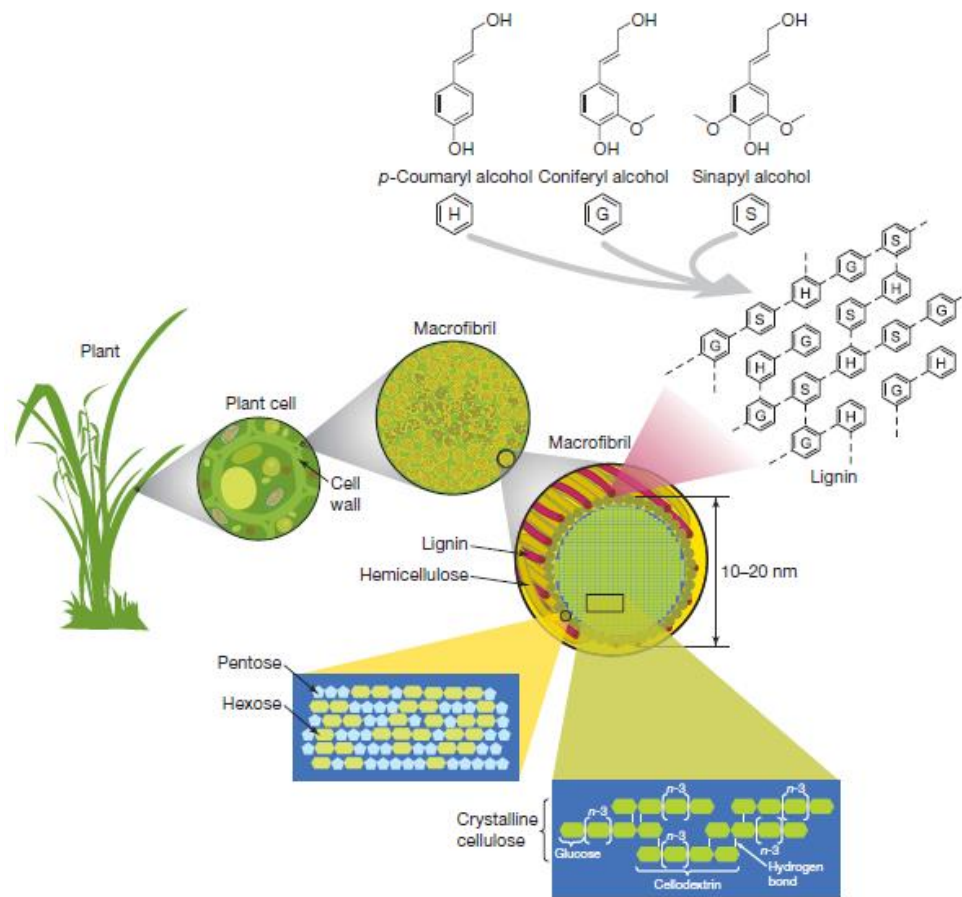
No processo SSF, a sacarificação da biomassa e fermentação dos açúcares acontecem no mesmo compartimento, com os açúcares liberados sendo imediatamente metabolizados e convertidos em etanol. Dessa forma, a concentração de açúcares inibidores é tão baixa que não causa uma inibição significativa das enzimas lignocelulolíticas, o que permite uma maior hidrólise da biomassa, adição de uma menor quantidade de enzimas, menor risco de contaminação biológica, maior rendimento e menor custo na produção do biocombustível. Assim, essa estratégia de hidrólise e fermentação simultânea é considerada promissora e mais vantajosa em relação à estratégia com os processos separados (SHF). A principal desvantagem é a necessidade de se usar uma temperatura que não é ótima para a hidrólise da biomassa e fermentação dos açúcares ao mesmo tempo. A fim de se contornar essa desvantagem e aumentar o rendimento deste processo, cepas de leveduras recombinantes estão sendo desenvolvidas para realizar a fermentação dos açúcares em temperaturas mais altas e próximas à temperatura ótima das enzimas hidrolíticas (Dashtban *et al.*, 2009; Limayen & Ricke, 2012; Menon & Rao, 2012; Kricka *et al.*, 2014).

Por fim, na estratégia CBP, as etapas de produção de enzimas lignocelulolíticas (celulases e hemicelulases), hidrólise da biomassa, e fermentação de pentose (xilose) e hexose (glicose) são realizadas em um único reator por um ou mais microrganismos. Essa combinação de etapas em um único processo é interessante uma vez que não há a necessidade de adição de enzimas produzidas de forma exógena. Apesar de vários microrganismos serem capazes de produzir álcool a partir de açúcar, eles são limitados em termos de co-fermentação de hexose e pentose, rendimento de etanol, e tolerância a inibidores (como o próprio etanol) e alta temperatura. Diante da inexistência de um microrganismo ideal, esforços vêm sendo feitos com o objetivo principal de engenheirar geneticamente microrganismos capazes de superar esses obstáculos e produzir o etanol 2G

de uma forma economicamente viável (Dashtban *et al.*, 2009; Nigam & Singh, 2011; Limayem & Ricke, 2012; Menon & Rao, 2012; Kricka *et al.*, 2014). Um dos candidatos é a levedura termotolerante *Kluyveromyces marxianus*, isolada originalmente de uvas em 1888. Essa levedura possui capacidade de crescer e co-fermentar hexose e pentose em temperaturas elevadas entre 42 e 45°C, sendo mais termotolerante que *Saccharomyces sp.* (Fonseca *et al.*, 2008). Expressando a endoglucanase EGII de *Trichoderma reesei* e a  $\beta$ -glicosidase BGL1 de *Aspergillus aculeatus*, Yanase e colaboradores (2010) desenvolveram uma linhagem de *K. marxianus* que converte o açúcar  $\beta$ -glicano diretamente em etanol a 48°C, com um rendimento de 0,47 g/g de etanol (92,2% do rendimento teórico). Dessa forma, essa levedura recombinante parece ser promissora para ser utilizada no CBP, apesar de ser menos tolerante ao etanol que *S. cerevisiae* (Limayem & Ricke, 2012; Kricka *et al.*, 2014).

### **2.3 Parede celular vegetal**

A lignocelulose é o material orgânico renovável mais abundante do mundo e o maior componente estrutural das plantas, correspondendo de 50 a 90% de toda biomassa vegetal (Sánchez & Cardona, 2008; Kricka *et al.*, 2014). Estima-se que a produção mundial desta biomassa seja de  $2 \times 10^{11}$  milhões de toneladas por ano, das quais de 8 a  $20 \times 10^9$  milhões são potencialmente acessíveis para serem processadas e convertidas em produtos de valor agregado (Kricka *et al.*, 2014). Como já mencionado, a lignocelulose é composta basicamente por celulose, hemicelulose e lignina, e a proporção de cada componente varia dependendo do tipo de vegetal, mas em média é 30-50% de celulose, 20-30% hemicelulose e 15-25% lignina (Kricka *et al.*, 2014). Pequenas quantidades de cinzas, lipídios, proteínas, açúcares solúveis, minerais e pectina também podem ser encontradas junto à lignocelulose em diferentes graus, de acordo com a fonte da biomassa (Pauly & Keegstra, 2008; Sanchez, 2009). A Figura 2 mostra simplificada um modelo da arquitetura da parede celular vegetal, com os principais elementos que a compõem.



**Figura 2.** Modelo simplificado da arquitetura da parede celular vegetal e os principais componentes da lignocelulose presente na parede: celulose, hemicelulose e lignina (Rubin, 2008).

### 2.3.1 Celulose

A celulose é o polissacarídeo mais abundante no mundo e o maior constituinte de todas as plantas (Percival Zhang *et al.*, 2006), correspondendo a um polímero linear composto por unidades de D-glicose ligadas umas às outras por ligações glicosídicas do tipo  $\beta$ -1,4 (Rubin, 2008). Cada molécula de glicose está posicionada de modo reverso ( $180^\circ$ ) à outra molécula adjacente (Kubicek, 2013), e o número de unidades presentes nas cadeias pode variar de centenas até 10.000 (Horn *et al.*, 2012). A ligação das cadeias de celulose adjacentes por pontes de hidrogênio, interações hidrofóbicas e forças de Van der Waal faz com que as cadeias fiquem alinhadas paralelamente, formando as microfibrilas de celulose

(Percival Zhang *et al.*, 2006). A celulose pode ser quebrada por ácidos ou celulases em monômeros de glicose, que são fermentados e convertidos em etanol, butanol, ácido lático ou ácido succínico (Adsul *et al.*, 2011). Quanto a sua morfologia, a celulose apresenta domínios cristalinos altamente compactados separados por regiões amorfas não organizadas (Adsul *et al.*, 2011). A celulose cristalina é altamente resistente à hidrólise porque sua estrutura compactada impede a entrada não somente de enzimas mas também de pequenas moléculas, como a água (Arantes & Saddler, 2010). Por outro lado, estudos mostraram que a estrutura amorfa da celulose é hidrolisada mais rapidamente que a celulose cristalina (Fan *et al.*, 1980; Lee *et al.*, 1983; Hall *et al.*, 2010), sendo prontamente degradada pelas celulases (Kumar *et al.*, 2008).

### **2.3.2 Hemicelulose**

A hemicelulose é o segundo polímero mais abundante da lignocelulose e sua composição é variável e dependente da biomassa vegetal (Peng *et al.*, 2012, Kricka *et al.*, 2014). Trata-se de um heteropolímero altamente ramificado composto por pentoses (xilose e arabinose), hexoses (manose, glicose e galactose) e por açúcares ácidos (ácido glucurônico, galacturônico e metilgalacturônico) (McMillan, 1993; Saha, 2003; Adsul *et al.*, 2011). Outros açúcares como ramosse e fucose também podem estar presentes em pequenas quantidades (Gírio *et al.*, 2010). As hemiceluloses são classificadas de acordo com a composição de seus polissacarídeos, e os mais comuns são xilano, xiloglucano e manano (Horn *et al.*, 2012). O xilano é o polímero mais comum na hemicelulose e é abundante em gramíneas e angiospermas (cana, trigo e *Miscanthus*). Esse polímero consiste em uma cadeia de xilose com ligações do tipo  $\beta$ -1,4 com cadeias laterais que se ligam e conferem ao xilano variados graus de ramificações (Horn *et al.*, 2012). Embora a maioria dos xilanos seja ramificada, polissacarídeos lineares já foram isolados (Kubicek, 2012). Monômeros de arabinose e galactose também podem estar ligados à xilose. O xiloglucano está presente em muitas angiospermas (Horn *et al.*, 2012) e é formado por uma cadeia de resíduos de glicose com ligações do tipo  $\beta$ -1,4 com uma alta porcentagem desses resíduos substituídos por xilose. Os xiloglucanos interagem com as microfibrilas de celulose pela formação de pontes de hidrogênio, e contribuem para a integridade da estrutura da celulose (Carpita & Gibeaut,

1993; de Vries & Visser, 2001; Gírio *et al.*, 2010). Tanto xilano quanto xiloglucano apresentam ainda grupos acetil ligados aos seus resíduos. A cadeia principal de manano é composta inteiramente por resíduos de manose, como em mananos e galactomananos, ou por resíduos de manose e glicose, como em glicomananos e galactoglicomananos (Scheller & Ulvskov, 2010). As pentoses e hexoses que são ligadas por ligações 1,3, 1,4 e 1,6 são muitas vezes acetiladas e formam uma estrutura hidrofílica que age como uma espécie de “cola” entre a celulose e a lignina (Adsul *et al.*, 2011). Mananos e glicomananos são normalmente poucos abundantes em gimnospermas e angiospermas, onde outras hemiceluloses parecem tê-los substituídos (Scheller & Ulvskov, 2010).

Devido à grande diversidade de açúcares que constituem a hemicelulose, são necessárias diversas enzimas para hidrolisá-la completamente em monômeros (Limayem & Rickie, 2012). As hemiceluloses são geralmente mais fáceis de serem quebradas que a celulose, mas alguns oligômeros conferem recalcitrância a esse material por causa do perfil complexo de ramificação e acetilação das cadeias (Agger *et al.*, 2010). Diferentemente da despolimerização da celulose que produz apenas glicose, a degradação de hemicelulose produz uma mistura de diferentes açúcares, em sua maioria pentoses que não podem ser utilizadas pela levedura industrial *Saccharomyces cerevisiae* (Banerjee *et al.*, 2010a; Horn *et al.*, 2012).

### **2.3.3 Lignina**

O terceiro componente da lignocelulose é a lignina, um polímero presente em todas as plantas vasculares e que representa a segunda fonte de carbono mais abundante do mundo (Dashtban *et al.*, 2010; Kubicek, 2012). A lignina é um polímero heterogêneo complexo formado por três álcoois aromáticos: coniferil, p-comaril e sinapil (Kubicek, 2012). Trata-se de um polímero hidrofóbico e altamente resistente à degradação química e biológica. Está presente na lamela média da parede vegetal e age como um cimento entre as células. Junto com as hemiceluloses, a lignina forma uma matriz amorfa na qual fibrilas de celulose ficam embebidas e protegidas da biodegradação (Adsul *et al.*, 2011). A lignina está ligada de forma cruzada com os carboidratos da parede vegetal por ligações do tipo éter ou éster (Takahashi & Koshijima, 1988), e esta associação entre a lignina e os demais componentes é que confere

a maior parte da recalcitrância da parede vegetal (Adsul *et al.*, 2011). Ao contrário da celulose e hemicelulose, a hidrólise da lignina não gera açúcares fermentáveis, ao invés disso produz compostos fenólicos que inibem a fermentação (Kricka *et al.*, 2014). Com o objetivo de reverter esse efeito negativo, Fu e colaboradores (2011) conseguiram reduzir significativamente os componentes da lignina em mutantes de *switchgrass* construídos por RNA de interferência (RNAi), e aumentaram assim o rendimento da produção de etanol 2G em até 38%.

#### **2.3.4 Bagaço de cana-de-açúcar**

Além da produção de açúcar e álcool, a cultura da cana gera como subprodutos o bagaço, palha, melaço, torta de filtro e vinhaça (Cortez *et al.*, 1992). O bagaço é o principal resíduo resultante da moagem do colmo da cana para extração do caldo (Basso *et al.*, 2013). Estima-se que a cada tonelada de cana moída são gerados cerca de 280 Kg de bagaço, que é constituído por aproximadamente 45% de celulose, 25% de hemicelulose, 20% de lignina e pequenas quantidades de outros compostos (Lin & Tanaka, 2006; Ragauskas *et al.*, 2006). Além de ser utilizado pelas próprias usinas para gerar calor e produzir eletricidade, o bagaço também pode ser utilizado como suplemento animal (Banerjee *et al.*, 2010b) e como matéria-prima para produção de compostos químicos (como por exemplo, furfural e cola de papel) (Cardona *et al.*, 2010). Atualmente cada vez mais estudos vêm sendo realizados para produção de etanol 2G a partir desta biomassa (Wanderley *et al.*, 2013; Khuong *et al.*, 2014; Li *et al.*, 2014). Como citado anteriormente, se 8% do bagaço fosse utilizado para produção deste etanol, haveria um aumento significativo no rendimento deste combustível por hectare (de 6.000 para 8.200 litros) (Leite *et al.*, 2009). Entretanto, para que estes números se tornem realidade, há desafios biotecnológicos que devem ser superados para que o etanol 2G seja economicamente viável, dentre eles destaca-se a quebra da parede celular vegetal.

#### **2.4 Degradação da biomassa por enzimas e proteínas acessórias**

A fim de tornar a produção dos biocombustíveis sustentável e rentável, há a necessidade de se reduzir os custos do pré-tratamento da biomassa, consolidar as etapas de seu processamento, desenvolver biocatalisadores eficientes, e alcançar a utilização completa



de todos os elementos da parede vegetal (Xie *et al.*, 2014). Notavelmente, muitos sistemas biológicos naturais evoluíram para superar o desafio da degradação da biomassa e utilizar eficientemente seus elementos devido à produção de enzimas que degradam a celulose, hemicelulose e lignina (Lynd *et al.*, 2002; Xie *et al.*, 2014). Estão inclusos nesses sistemas microrganismos unicelulares, como a bactéria *Clostridium*, fungos saprofitos, como as espécies dos gêneros *Trichoderma* e *Aspergillus*, consórcio de microrganismos reunidos no trato digestivo de cupins e ruminantes (Xie *et al.*, 2014), entre outros.

Os fungos estão entre os sistemas biológicos mais bem estudados para degradação de biomassa (Xie *et al.*, 2014), uma vez que evoluíram e desenvolveram um conjunto de enzimas hidrolíticas e oxidativas capazes de quebrar os polissacarídeos da parede vegetal, conferindo-lhes importante papel na degradação de materiais lignocelulósicos (Bourne & Henrissat, 2001; Adsul *et al.*, 2011; van den Brink & de Vries, 2011). Esse conjunto, entretanto, varia entre as espécies: *Trichoderma reesei*, por exemplo, tem enzimas altamente eficientes para degradação de celulose (Martinez *et al.*, 2008; Kubicek *et al.*, 2011), enquanto que as espécies de *Aspergillus* produzem muitas enzimas voltadas à degradação de pectina (Martens-Uzunova & Schaap, 2009) e hemicelulose (Polizeli *et al.*, 2005).

Essas enzimas são conhecidas como glicosil hidrolases (GHs, *glycosyl hydrolases*), carboidrato esterases (CEs, *carbohydrate esterases*) e polissacarídeo liases (PLs, *polysaccharide lyases*), e são classificadas e descritas no banco de dados CAZy (***Carbohydrate Active Enzyme***), de acordo com a similaridade das sequência dos aminoácidos e estrutura das proteínas (Cantarel *et al.*, 2009). As primeiras são responsáveis pela hidrólise e/ou transglicosilação de ligações glicosídicas e representam quase metade (47%) das enzimas classificadas no CAZy. As esterases removem grupamentos éster presentes nos mono, oligo e polissacarídeos, e facilitam a ação das GHs. As liases clivam as ligações glicosídicas de polissacarídeos contendo ácido urônico por um mecanismo de  $\beta$ -eliminação, que consiste na reação de quebra de uma ligação glicosídica 4-O do ácido urônico em três passos (Michaud *et al.*, 2003; Yip & Withers, 2006; Cantarel *et al.*, 2009; Lombard *et al.*, 2014).

### 2.4.1 Celulases

A natureza recalcitrante da celulose representa um desafio para os sistemas celulolíticos (Lynd *et al.*, 2002). Para que os microrganismos possam hidrolisar e metabolizar a celulose insolúvel, celulases devem ser produzidas e secretadas, podendo ficar no meio extracelular na forma livre ou associadas à membrana da célula (Lynd *et al.*, 2002). A degradação da celulose pela via clássica envolve a ação sinérgica de três classes de enzimas: 1) endo-1,4- $\beta$ -glucanases (EGLs) (EC 3.2.1.4); 2) exoglucanases, incluindo 1,4- $\beta$ -D-glicano glicanohidrolases (também conhecidas como celodextrinases, EC 3.2.1.74) e 1,4- $\beta$ -D-glicano celobiohidrolases (CBHs) (celobiohidrolase 1 CBH1, EC 3.2.1.176; celobiohidrolase 2 CBH2, EC 3.2.1.91); 3)  $\beta$ -glicosidases (BGLs) (EC 3.2.1.21) (Lynd *et al.*, 2002; Horn *et al.*, 2012). Todas essas celulases são hidrolases, pois clivam as ligações glicosídicas pela adição de uma molécula de água (Davies & Henrissat, 1995).

As endoglucanases (EGLs) iniciam a degradação da celulose clivando randomicamente ligações glicosídicas em sítios localizados na região amorfa ao longo da cadeia de celulose. Esse ataque promove uma rápida diminuição no grau de polimerização (DP, *degree of polymerization*) da celulose e expõe novas extremidades das cadeias de celulose (Kricka *et al.*, 2014). As endoglucanases estão presentes nas famílias das GHs 5, 6, 7, 8, 9, 12, 44 e 45 (van den Brink & de Vries, 2011; Kubicek *et al.*, 2014; CAZy database).

As exoglucanases atacam progressivamente as extremidades redutoras e não-redutoras da celulose, iniciando sua ação na extremidade da cadeia e continuando ao longo dessa até o seu final (Divne *et al.*, 1994; Teeri *et al.*, 1998; Horn *et al.*, 2012). Como produto, as CBHs liberam oligossacarídeos e, principalmente, celobiose, um dissacarídeo de glicose que inibe as celobiohidrolases (Sánchez & Cardona, 2008; Kricka *et al.*, 2014). As celobiohidrolases produzidas pelo fungo *T. reesei* são empregadas em coquetéis enzimáticos comerciais e correspondem em até 80% das proteínas totais (Rosgaard *et al.*, 2007). Quanto a sua distribuição, as celobiohidrolases estão presentes nas famílias GHs 5, 6, 7, 9 e 48 (van den Brink & de Vries, 2011; CAZy database).

Por fim, as  $\beta$ -glicosidases (BGLs) hidrolisam a ligação glicosídica  $\beta$ -1,4 da celobiose e de oligossacarídeos para liberar unidades de glicose (Kricka *et al.*, 2014), diminuindo assim o

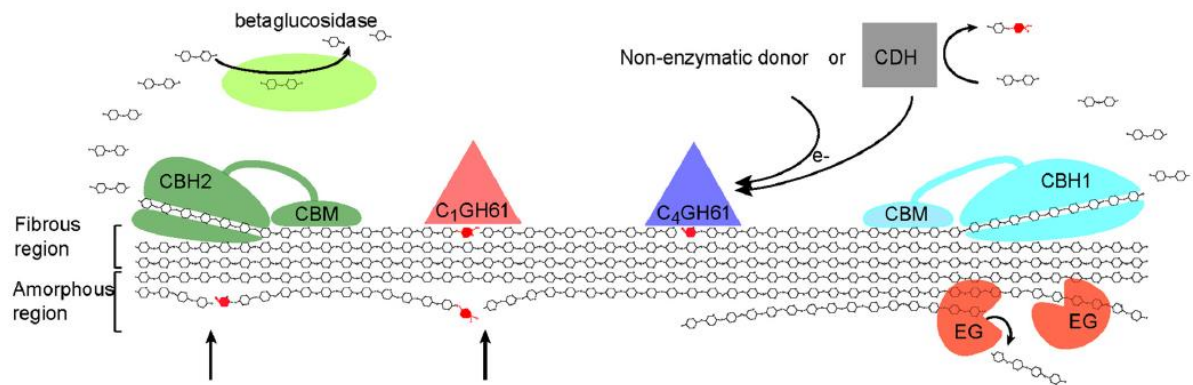
produto da inibição das celobiohidrolases (Gusakov & Sinitsyn, 1992). Sabe-se que conforme o DP aumenta, a taxa de hidrólise das  $\beta$ -glicosidases diminui consideravelmente (Zhang & Lynd, 2004). As  $\beta$ -glicosidases podem variar em termos de propriedade e localização celular, e embora sejam encontradas nas famílias das GHs 1, 3, 5, 9, 30 e 116 (Hakulinen *et al.*, 2000; Häkkinen *et al.*, 2012; CAZy database), a maioria das BGLs empregadas para hidrólise da celulose pertence à família GH3 (Hakulinen *et al.*, 2000).

Uma característica em comum da maioria das celulasas é uma estrutura modular conhecida como domínio de ligação ao carboidrato (CBM, *carbohydrate-binding module*). O CBM afeta a ligação da celulase à superfície da celulose, e provavelmente facilita a hidrólise desse polímero por aproximar o domínio catalítico ao substrato. A presença dos CBMs é particularmente importante para o início e processividade das celobiohidrolases (Teeri *et al.*, 1998), que é essencial para degradar as regiões cristalinas da celulose (Eijsink *et al.*, 2008).

#### **2.4.2 Proteínas acessórias e mecanismo oxidativo de degradação de celulose**

Em 2007, foi descoberta uma proteína do fungo *Thielavia terrestris*, denominada GH61, que aumentava a atividade de celulasas de *T. reesei* (Merino & Cherry, 2007). No ano seguinte, Moser e colaboradores (2008) também observaram o mesmo resultado com uma proteína com domínio CBM33 da bactéria *Thermobifida fusca*. Atualmente, essas enzimas são conhecidas como oxido-redutoras e pertencem à família AA (*auxiliary activity*), que corresponde às proteínas com atividade auxiliar do CAZy, que até pouco tempo atrás eram chamadas de LPMOs (*lytic polysaccharide mono-oxygenases*). Estudos recentes de proteínas de bactérias classificadas como CBM33 (Vaaje-Kolstad *et al.*, 2010; Forsberg *et al.*, 2011; Vaaje-Kolstad *et al.*, 2012) e de proteínas de fungos da antiga família GH61 (Harris *et al.*, 2010; Beeson *et al.*, 2012) mostraram que elas catalisam a celulose por um mecanismo oxidativo dependente da presença de cátions metálicos e de um doador de elétrons (Vaaje-Kolstad *et al.*, 2010). Alguns doadores de elétrons conhecidos por induzir essa atividade oxidativa são agentes redutores de baixo peso molecular, como ácido ascórbico, ácido gálico e glutatona reduzida (Quinlan *et al.*, 2011; Westereng *et al.*, 2011; Horn *et al.*, 2012). Essas proteínas também são CAZymes, mas devido ao seu modo de ação ser diferente das enzimas hidrolíticas, elas serão apresentadas neste trabalho como proteínas acessórias. A Figura 3

mostra simplificadamente o modo de ação das celulases e dessas proteínas acessórias na degradação de celulose.



**Figura 3.** Degradação de celulose por celulases e por mecanismos oxidativos. As endoglucanases (EG) clivam as ligações glicosídicas na região menos organizada da celulose, conhecida como região amorfa. As extremidades reductoras e não-reductoras são então hidrolisadas progressivamente pelas celobiohidrolases (CBH1 e CBH2, respectivamente) com a participação dos domínios de ligação ao carboidrato (CBM). Dessa hidrólise, é formada como produto principal a celobiose, que é convertida em duas moléculas de glicose pelas  $\beta$ -glicosidasas. Proteínas com atividade auxiliar da família AA9 ( $C_1GH61$  e  $C_4GH61$ ) também participam da desconstrução da celulose por um mecanismo oxidativo. As setas apontadas para a celulose indicam novos pontos de ataque para as celobiohidrolases (Horn *et al.*, 2012).

As celobiose desidrogenases (CDH, EC 1.1.99.18) também ajudam na despolimerização da celulose de modo sinérgico com as enzimas oxido-redutoras (Langston *et al.*, 2011; Sygmond *et al.*, 2012), catalisando a oxidação da extremidade reductora da celobiose e de oligossacarídeos. Nem todos os sistemas celulolíticos possuem CDHs e o exato papel que essas enzimas têm na degradação da lignocelulose ainda é desconhecido (Kallioinen, 2014).

Acredita-se que swoleninas também tenham uma participação na degradação de polissacarídeos, uma vez que a combinação destas com celulases e xilanas acelera a hidrólise de celulose e xilano (Kim *et al.*, 2009; Gourlay *et al.*, 2013). As swoleninas apresentam efeitos notáveis nos substratos de celulose insolúvel, como diminuição no tamanho e na cristalinidade das partículas, sem formar açúcares redutores (Saloheimo *et al.*, 2002; Chen *et al.*, 2010; Jager *et al.*, 2011). Gourlay e colaboradores (2013) demonstraram que as swoleninas são capazes de romper a fração hemicelulolósica de palha de milho pré-

tratado e que possuem sinergismo com xilanases. Recentemente, Andberg e colaboradores (2015) observaram atividade de SWO1 em ligações glicosídicas do tipo  $\beta$ -1,4, mostrando que ao contrário do que se pensava, ela possui atividade catalítica.

As hidrofobinas também merecem atenção pois são proteínas específicas de fungos filamentosos (Linder, 2009) que fazem a interface entre superfícies hidrofóbicas e hidrofílicas, mediando portanto a interação do fungo com seu ambiente (Linder *et al.*, 2005; Ohtaki *et al.*, 2006). A observação de que os genes dessas proteínas estavam hiper-expressos em *T. reesei* e *A. niger* crescidos em meio com palha de trigo sugere que o reconhecimento de superfícies sólidas pode ser um passo importante para a resposta fisiológica do fungo quando exposto à parede celular vegetal (Ries *et al.*, 2013). Assim, essas proteínas representam candidatos em potencial para o aumento da degradação da lignocelulose, embora isso ainda não tenha sido confirmado.

### **2.4.3 Hemicelulases**

Hemicelulases são enzimas que degradam a parte de hemicelulose de materiais lignocelulósicos e, devido a sua natureza complexa, várias enzimas são necessárias para a desconstrução desse polímero. As três cadeias principais de hemicelulose são hidrolisadas por um conjunto específico de enzimas:  $\beta$ -1,4-endoxilanase e  $\beta$ -1,4-xilosidase para xilano,  $\beta$ -1,4-endoglucanase e  $\beta$ -1,4-glicosidase para xiloglucano, e  $\beta$ -1,4-endomananase e  $\beta$ -1,4-manosidase para (galacto-) manano (de Vries & Visser, 2001; van den Brink & de Vries, 2011). Xilanases e mananases clivam as ligações  $\beta$ -1,4 do xilano e manano, e liberam oligossacarídeos que são convertidos em monômeros de xilose e manose por  $\beta$ -xilosidases e  $\beta$ -manosidases, respectivamente.  $\beta$ -xilosidases hidrolisam xilooligossacarídeos e xilobiose a partir da extremidade não-redutora da xilose, o que previne a inibição das xilanases por esses substratos (Knob *et al.*, 2010).  $\beta$ -manosidase é um tipo de enzima com ação do tipo exo que age na extremidade não-redutora de manano e manano-oligossacarídeos, liberando resíduos de manose (Coughlan *et al.*, 1993; Wong & Saddler, 1993). As endoglucanases ativas contra xiloglucano, também conhecidas como xiloglucanases, podem clivar resíduos ramificados de glicose e xiloglico-oligossacarídeos contendo mais que seis resíduos de glicose (Desmet *et al.*, 2007; Master *et al.*, 2008). De uma forma geral, as xilanases (EC 3.2.1.8) estão

presentes nas famílias GH5, 8, 10, 11, 30, 43 e 51, e as  $\beta$ -xilosidases (EC 3.2.1.37), nas famílias GH1, 3 (maioria), 30, 39, 43, 51, 52, 54, 116 e 120. As endomananases (EC 3.2.1.78) pertencem às famílias GH5, 26 e 113, e as manosidases (EC 3.2.1.25), à família GH1, 2 e 5. As xiloglucanases (EC 3.2.1.151) estão presentes nas famílias GH5, 9, 12, 16, 44 e 74 (van den Brink & de Vries, 2011; CAZy database).

A arabinose, um resíduo comum na hemicelulose, é clivado do xiloglucano e (arabino-) xilano pelas  $\alpha$ -arabinofuranosidases e arabinoxilano arabinohidrolases. As  $\alpha$ -arabinofuranosidases de fungos são principalmente encontradas nas famílias GH51 e 54, embora algumas enzimas bifuncionais das famílias GH3 e 43 também tenham sido descritas com essa atividade (Saha, 2000; Ravanal *et al.*, 2010). Entre as outras principais enzimas que atuam sobre a hemicelulose estão:  $\alpha$ -galactosidase, que libera galactose;  $\alpha$  e  $\beta$ -glucuronidases, que removem ácido glucurônico da xilose; acetil xilano esterases, feruloil e p-coumaroil esterases (Dashtban *et al.*, 2009; Dodd & Cann, 2009).

#### **2.4.4 Ligninases**

As ligninases são enzimas que degradam lignina, o componente mais recalcitrante da parede celular vegetal. Como a estrutura da lignina é complexa e variável, apenas poucos microrganismos são capazes de degradá-la. Dentre esses microrganismos estão bactérias do gênero *Streptomyces* sp. e *Nocardia* sp. e fungos basidiomicetos (*brown rot* e *white rot*), sendo os últimos os produtores mais eficientes de ligninases. Esses microrganismos produzem enzimas extracelulares não-específicas com ação oxidativa contra a lignina (Thevenot *et al.*, 2010), que são classificadas em duas famílias: as fenol oxidases (lacases) e peroxidases (peroxidase de lignina e manganês). A degradação se inicia com a produção de radicais livres de oxigênio pela reação de Fenton. Os radicais atacam então a lignina e açúcares associados à parede, permitindo que outras proteínas tenham acesso à estrutura lignocelulósica (Dashtban *et al.*, 2009). No genoma do ascomiceto *T. reesei* foi encontrada apenas uma lacase (jgi|88862) e nenhuma peroxidase. Essa lacase é secretada mas ainda não foi caracterizada (Ries, 2013).

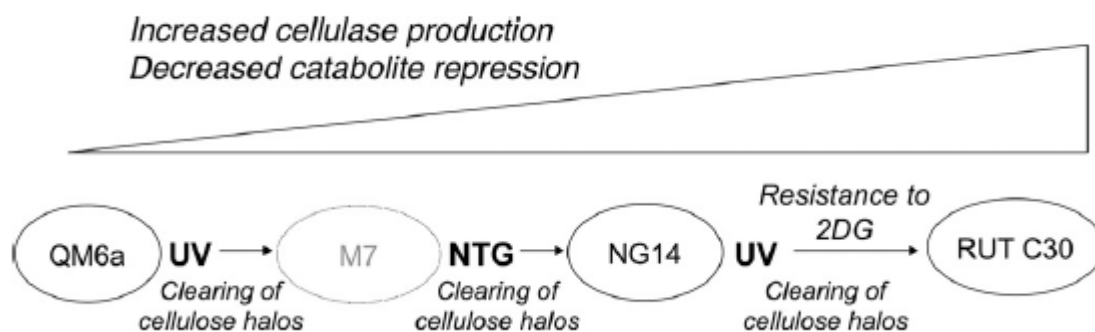
## **2.5 *Trichoderma reesei***

*T. reesei* é um fungo filamentosos, mesofílico e saprofítico que cresce aerobicamente sobre biomassa vegetal morta, entre outros substratos. Taxonomicamente é classificado como da divisão Ascomycota, classe Sordariomycetes, ordem Hypocreales, família Hypocreales e gênero *Trichoderma*. Até o momento, foram identificadas a nível molecular 104 espécies desse gênero, sendo a maioria saprofítica (Schuster *et al.*, 2010). Existem algumas espécies de patógenos de plantas e cogumelos (*T. koningii*), bem como espécies que se alimentam de outros fungos (*T. viride*, *T. harzianum* e *T. atroviride*) (Schuster *et al.*, 2010). São várias as aplicações dessas espécies, dentre elas podemos citar seu uso como agente de controle biológico na agricultura (*T. atroviride* e *T. virens*) e produção de enzimas industriais (*T. reesei*) (Schuster *et al.*, 2010).

O fungo filamentosos *T. reesei* é usado amplamente na indústria de enzimas e é o organismo celulolítico mais estudado (Foreman *et al.*, 2003; Martinez *et al.*, 2008; Kubicek, 2013). *T. reesei* é utilizado para produção de celulasas e hemicelulasas nativas, mas também para produção de proteínas heterólogas. Ele tem servido como fungo modelo para estudos da degradação de lignocelulose, sendo que diversas celulasas e hemicelulasas foram identificadas e caracterizadas pela primeira vez através dele (Shoemaker *et al.*, 1983; Penttilä *et al.*, 1986; Teeri *et al.*, 1987).

*T. reesei* foi isolado originalmente nas Ilhas Salomão, atual Papua-Nova Guiné, na Segunda Guerra Mundial, em 1944, colonizando barracas de algodão e roupas do exército americano (Seidl *et al.*, 2009). O isolado foi identificado inicialmente como *Trichoderma viride* e a cepa nomeada como QM6a (Mandels & Reese, 1957). Mais tarde, a cepa QM6a foi reconhecida como sendo diferente da espécie *T. viride*, sendo então renomeada como *T. reesei* em homenagem ao pesquisador Elwyn T. Reese (Simmons, 1977). Como *T. reesei* QM6a era um bom produtor de celulase e já havia a necessidade de se utilizar biomassa rica em celulose para produzir etanol e substituir as fontes de combustíveis fósseis, foram iniciados programas de desenvolvimento de mutantes de *T. reesei* capazes de produzir mais celulasas que a cepa parental (Mandels & Reese, 1957; Eveleigh, 1982; Montenecourt, 1983).

Uma das melhores cepas produtoras de celulases é a *T. reesei* RUT C30 (Eveleigh & Montenecourt, 1979; dos Santos Castro *et al.*, 2014). Essa cepa hipercelulolítica foi obtida depois de três etapas de mutagênese (Figura 4). Primeiro, foi feita a exposição da cepa QM6a à radiação de luz ultra-violeta (UV) e *screening* dos mutantes com capacidade de hidrolisar celulose mesmo em condições de repressão catabólica. Essa etapa levou à seleção da cepa M7 que foi, em seguida, submetida ao agente químico N-nitroguanidina (NTG), dando origem à cepa NG14 que é superior na quantidade de proteínas extracelulares e na atividade de celulase quando comparada com as cepas parentais (Eveleigh & Montenecourt, 1979). A cepa RUT C30 foi então obtida a partir da NG14 usando novamente mutagênese por UV (luz ultravioleta) e *screening* para alta atividade celulolítica e resistência ao metabólito 2-deoxiglicose (Wick *et al.*, 1957). O mutante obtido produz duas vezes mais proteína extracelular em relação à cepa NG14, chegando a produzir mais de 30 g/L de proteína em fermentação industrial (Eveleigh & Montenecourt, 1979).



**Figura 4.** Mutagênese e seleção das cepas mutantes de *Trichoderma reesei*. Os agentes mutagênicos usados estão em negrito e os fenótipos para *screening* dos mutantes, em itálico. A cepa M7 está menos acentuada pois não se encontra mais disponível. UV: radiação por luz ultra-violeta; NTG: N-nitroguanidina; 2DG: 2-deoxiglicose (Le Crom *et al.*, 2009).

Muitas enzimas de sacarificação não são produzidas em meio com açúcares monoméricos (por exemplo, glicose e galactose) (de Graaff *et al.*, 1994; Gielkens *et al.*, 1997; van Peij *et al.*, 1997; de Vries *et al.*, 1998). Isto porque a presença de carboidratos prontamente metabolizáveis reprime a síntese dessas enzimas, assegurando assim a utilização preferencial da fonte de carbono mais fácil de ser metabolizada. Para a célula isto é



benéfico porque a fonte de carbono mais favorável é utilizada e a energia não é desperdiçada na síntese de outros sistemas catabólicos. Este fenômeno é conhecido por “repressão por carbono” ou “repressão catabólica por carbono” (Ronne, 1995; Ruijter & Visser, 1997) e não ocorre na cepa RUT C30. Isto porque nesta cepa o gene *cre1* que atua como maior mediador da repressão está truncado, permitindo a maior produção de enzimas lignocelulolíticas (Ilmen *et al.*, 1996). Além dessa mutação, a cepa RUT C30 possui outra na matriz de leitura do gene *gls2 $\alpha$*  da subunidade alpha da glicosidase II (ID 121351) envolvida na glicosilação protéica (Geysens *et al.*, 2005) e uma deleção de 85 kb no genoma, que eliminou 29 genes em relação à QM6a, dentre eles transportadores, fatores de transcrição e enzimas do metabolismo primário (Seidl *et al.*, 2008). Por fim, a análise da glicômica do RUT C30 mostrou evidências de que as hidrolases extracelulares têm um perfil atípico de glicosilação não visto na glicômica do parental QM6a (Peterson & Nevalainen, 2012), causada pela mutação no gene *gls2 $\alpha$* , específica das cepas RUT C30 e NG14 (Stals *et al.*, 2004; Geysens *et al.*, 2005).

### **2.5.1 Enzimas desconstrutoras da parede celular vegetal**

Recentemente, as sequências dos genomas de *T. reesei* QM6a e da linhagem RUT C30 foram disponibilizadas e comparadas (Martinez *et al.*, 2008; Le Crom *et al.*, 2009; Koike *et al.*, 2013). Cada genoma contém 9.129 (Martinez *et al.*, 2008) e 9.852 genes (Koike *et al.*, 2013), respectivamente, e até o momento foram anotadas 235 enzimas desconstrutoras de polissacarídeos em *T. reesei* (Häkkinen *et al.*, 2012). Desse total, 195 genes são de GHs, 22 CEs, 5 PLs e 13 AAs (Häkkinen *et al.*, 2012; Levasseur *et al.*, 2013; Ries *et al.*, 2013). *T. reesei* produz pelo menos duas celobiohidrolases (CBH1/CEL7A e CBH2/CEL6A), cinco endoglucanases (EG1/CEL7B, EG2/CEL5A, CEL5B, EG3/CEL12A, EG45/CEL45A), duas  $\beta$ -glicosidases já caracterizadas (BGL1/CEL3A e BGL2/CEL1A) e mais cinco preditas (CEL3B, CEL3D, CEL1B, CEL3C, CEL3E) (Häkkinen *et al.*, 2012). As CBHs e EGs compõem juntas até 90% do número total de enzimas secretadas, sendo que CBH1 e CBH2 constituem até 60% e 20% desse total, respectivamente (Rahman *et al.*, 2009). As BGLs representam apenas 1% do total de enzimas secretadas (Margeot *et al.*, 2009). Além dessas celulases, estão presentes outras proteínas envolvidas com a degradação de celulose, como a swolenina (Saloheimo *et al.*,

2002), hidrofobinas (Kubicek *et al.*, 2008) e as monooxigenases da família AA9, citadas anteriormente (Häkkinen *et al.*, 2012).

O genoma de *T. reesei* possui ainda 16 genes de hemicelulases, incluindo quatro endo-1,4- $\beta$ -xilanasases (XYN1, 2, 3 e 4), uma mananase (MAN1), uma acetilxilano esterase caracterizada (AXEI) e outra predita (AXEII), uma  $\alpha$ -glucuronidase (GLRI), uma arabinofuranosidase caracterizada (ABFI) e duas preditas (ABFII e ABFIII), três  $\alpha$ -galactosidasases (AGLI, AGLII e AGLIII) e uma  $\beta$ -xilosidase (BXL1) (Martinez *et al.*, 2008; Häkkinen *et al.*, 2012). Uma acetil esterase (AES1) que remove grupos acetil da hemicelulose também foi identificada (Li *et al.*, 2008). Esterases de carboidratos da família CE15 também têm um papel na degradação da lignocelulose. *T. reesei* codifica uma glucuronoil esterase (CIP2) que ajuda na dissociação da lignina da hemicelulose pela clivagem de ligações éster entre esses dois elementos da parede, facilitando a remoção de lignina (Foreman *et al.*, 2003; Li *et al.*, 2007; Pokkuluri *et al.*, 2011).

### **2.5.2 Principais reguladores das enzimas lignocelulolíticas**

A regulação da expressão gênica das enzimas hidrolíticas ocorre a nível transcricional de modo coordenado e é dependente da presença de indutor (Kubicek, 2013). Foi demonstrado que uma variedade de açúcares induzem a produção de celulasas e hemicelulases, como sofrorose, celobiose, xilose, xilobiore, xilo-oligossacarídeos, L-arabitol, galactose e lactose (Saloheimo *et al.*, 1997; Nogawa *et al.*, 2001; Olsson *et al.*, 2003). Dentre estes açúcares, a sofrorose é o principal indutor de celulasas, sendo formada durante a hidrólise de celulose através de uma reação de transglicosilação (El-Gogary *et al.*, 1989; dos Santos Castro *et al.*, 2014). Na presença de indutores, a regulação é feita por fatores de transcrição específicos que se ligam à região promotora, ativando ou inibindo a expressão de genes que codifiquem às enzimas. Até o momento, são conhecidos quatro ativadores transcricionais XYR1, ACE2, ACE3 e o complexo HAP2/3/5, e dois repressores CRE1 e ACE1 envolvidos na regulação da expressão das enzimas (Kubicek *et al.*, 2009; Häkkinen *et al.*, 2014).

XYR1 (*xylanase regulator 1*) é o ativador transcricional mais importante dos genes que codificam a celulases e hemicelulases, dentre elas *xyn1*, *xyn2*, *bxl1*, *cbh1*, *cbh2*, *egl1*, *abf2* e *bgl1* (Portnoy *et al.*, 2011; dos Castro *et al.*, 2014). Trata-se de um fator de transcrição do tipo Zn<sub>2</sub>Cys<sub>6</sub> que se liga a *motifs* arranjados de forma repetida e invertida espaçados uns dos outros de 10 a 12 bases na região promotora dos seus genes alvos (Rauscher *et al.*, 2006). Homólogos dessa proteína desempenhando papel semelhante também foram encontrados em outros fungos desconstrutores de biomassa, como *A. niger* (XlnR) (van Peij *et al.*, 1998) e *Neurospora crassa* (Xlr-1) (Sun *et al.*, 2012). Foi demonstrado que a deleção de *xyr1* cessa a indução de celulases e hemicelulases na presença de celulose e sofrase, e prejudica a indução de hemicelulases envolvidas na degradação de xilano e arabinano, destacando seu papel essencial na indução desses genes (Stricker *et al.*, 2008; Akel *et al.*, 2009). A expressão de *xyr1* é regulada transcionalmente pelo repressor catabólico CRE1 (Portnoy *et al.*, 2011) e pelo fator de transcrição ACE1 na presença de glicose e xilose (Saloheimo *et al.*, 2000; Aro *et al.*, 2003; Mach-Aigner *et al.*, 2008). Apesar de análises feitas *in silico* predizendo sítios de ligação para esses reguladores (Ries, 2013), não foi confirmada experimentalmente a ligação deles no promotor de *xyr1*.

Analisando possíveis reguladores da transcrição de *cbh1*, Aro e colaboradores (2001) identificaram ACE2 como sendo uma proteína ativadora de celulases que se liga a um *motif* encontrado no promotor deste gene (Aro *et al.*, 2001). Esse *motif* também é reconhecido por XYR1 (Furukawa *et al.*, 2009), o que sugere uma interação desses dois fatores em mediar a expressão gênica de enzimas hidrolíticas. Além de *cbh1*, ACE2 também regula positivamente outras enzimas na presença de celulose, como *cbh2*, *egl1*, *egl2*, *bgl2* e *xyn2* (Aro *et al.*, 2003; Verbeke *et al.*, 2009). Foi demonstrado que a deleção de *ace2* diminuiu a transcrição das principais celulases CBH1, CBH2, EG1 e EG2, e reduziu de 30 a 70% a atividade celulolítica de *T. reesei* quando crescido em celulose (Aro *et al.*, 2001). Juntamente com o repressor ACE1, acredita-se que ACE2 faça a fina regulação da expressão de XYR1 em *T. reesei* (Mach-Aigner *et al.*, 2008; Stricker *et al.*, 2008).

Recentemente, ACE3 foi identificado como mais um regulador positivo da expressão de celulases e hemicelulases. Häkkinen e colaboradores (2014) verificaram que a atividade

celulolítica de CBH1 e EG1 de *T. reesei* aumentou significativamente quando superexpressaram *ace3*. De modo inverso, a deleção desse gene levou à drástica redução da atividade e expressão de celulases e hemicelulases, o que apoia o papel de ACE3 como indutor dessas enzimas (Häkkinen *et al.*, 2014).

O complexo HAP2/3/5 (*heme activator protein*) é formado pelas subunidades HAP2, HAP3 e HAP5, que aumentam a expressão de vários genes, incluindo celulases, xilanases, amilases e acetamidases (Narendja *et al.*, 1999; Tsukagoshi *et al.*, 2001). HAP2/3/5 liga-se à região promotora de vários genes, como *cbh1*, *cbh2*, *xyn1* e *xyn2* (Liu *et al.*, 2008), e acredita-se que esse complexo aumenta a transcrição desses genes pelo remodelamento da cromatina (Tsukagoshi *et al.*, 2001; Ries, 2013).

Por fim, LAE1 é outra proteína relacionada com o controle da expressão dos genes de degradação de lignocelulose, apesar de não ser um fator de transcrição, mas sim uma metiltransferase. Analisando a expressão gênica do mutante  $\Delta lae1$ , Seiboth e colaboradores (2012) verificaram uma forte hipo-expressão de várias celulases (celobiohidrolases, endoglucanases e  $\beta$ -glicosidases), proteínas acessórias (*cip1*, swoleninas, monooxigenases da família AA9), xilanases e várias outras hemicelulases (arabinofuranosidase, manosidase, galactosidase). Por outro lado, a superexpressão de *lae1* levou a um aumento significativo na transcrição de celulases (Seiboth *et al.*, 2012).

A regulação da expressão de celulases e hemicelulases é alvo da repressão catabólica por carbono (CCR). CRE1 é a principal proteína da CCR e reprime direta ou indiretamente a expressão de genes envolvidos na degradação de polissacarídeos, assim como de genes ligados à utilização de etanol e aminoácidos como fontes de carbono (Portnoy *et al.*, 2011). CRE1 é um fator de transcrição do tipo  $Cys_2His_2$  *zinc finger* que se liga aos promotores dos seus genes alvos por *motifs* identificados nos promotores dos genes *cbh1*, *cbh2*, *egl1*, *egl2* e *xyn1* em *T. reesei* (Ilmén *et al.*, 1997). Os sítios de ligação de CRE1 são compostos por dois *motifs* e acredita-se que a repressão direta só ocorra quando o fator se liga aos dois sítios (Cubero *et al.*, 1994; Mach *et al.*, 1996).

Apesar de ter sido demonstrado que CRE1 atua em vários níveis de regulação (Mello-de-Sousa *et al.*, 2014) e reprime diretamente a transcrição de celulases e hemicelulases

(Mach *et al.*, 1996; Ilmén *et al.*, 1998), Mello-de-Souza e colaboradores (2014) mostraram que a forma truncada de *cre1* é transcrita em mRNA e que seu produto de 96 aminoácidos ainda é capaz de ligar ao DNA, aumentando a produção de celulase em *T. reesei*. Os autores sugerem que isso pode ser explicado parcialmente pela abertura da cromatina mediada por CRE1 truncado, que torna os promotores acessíveis mesmo na presença de glicose (Mello-de-Souza *et al.*, 2014). A proteína homóloga à CRE1 em *A. nidulans*, CreA, está sujeita à troca de compartimentos celulares durante a CCR e é transportada para o núcleo na presença de glicose (Cziferszky *et al.*, 2002). Recentemente, Lichius e colaboradores (2014) realizaram ensaios de marcação de CRE1 e XYR1 fusionados com a proteína fluorescente GFP e analisaram a presença de cada proteína no núcleo e no citoplasma em condições de repressão (glicose) ou ativação (soforose) de genes de celulases e hemicelulases. De modo interessante, constataram que o repressor tende a ser recrutado pelo núcleo ao longo do pequeno intervalo de tempo analisado (5 a 25 minutos) na presença de glicose, e que a expressão de *cbh1* e *xyn2* diminuiu em relação à condição de crescimento em soforose. Na condição de indução por soforose, a situação se inverteu com uma perda da intensidade da fluorescência de CRE1 no núcleo e aumento na expressão de *cbh1* e *xyn2* (Lichius *et al.*, 2014).

Por último, ACE1 é outro regulador negativo de celulases e hemicelulases. ACE1 possui três *zinc fingers* do tipo Cys<sub>2</sub>His<sub>2</sub> e se liga ao *motif* encontrado na região promotora dos seus genes alvos. Em meio contendo celulose, a deleção de *ace1* resultou no melhor crescimento de *T. reesei* devido ao aumento na produção de celulase (Aro *et al.*, 2003). Na presença de celulose e soforose, a deleção de *ace1* mostrou uma maior expressão de celulases (*cbh1*, *cbh2*, *egl1* e *egl2*) e hemicelulases (*xyn1* e *xyn2*). O gene *ace1* é transcrito na presença de diferentes fontes de carbono e é independente de ACE2 e da CCR, embora 11 supostos sítios de ligação para CRE1 tenham sido encontrados dentro da sua região promotora (Aro *et al.*, 2003).

## **2.6 Desafios biotecnológicos na produção do etanol 2G e biologia de sistemas**

São vários os gargalos presentes no processo de produção do etanol 2G que precisam ser superados para que este combustível se torne cada vez mais competitivo e viável. Dentre

esses, estão: pré-tratamento e hidrólise mais eficazes e a um menor custo, maior produção de enzimas lignocelulolíticas a baixo custo, fermentação de xilose e glicose, diminuição da concentração de inibidores da fermentação, e uso de microrganismos mais robustos e tolerantes às condições adversas (Banerjee *et al.*, 2010a; Ho *et al.*, 2014). Nesse sentido, pesquisas básicas são essenciais para compreender as vias e os genes envolvidos na hidrólise da parede vegetal, sensoramento da fonte de carbono, produção de metabólitos e fermentação dos açúcares pelos microrganismos (Yuan *et al.*, 2008).

A biologia de sistemas integra as abordagens “ômicas” e tem como objetivo proporcionar uma descrição qualitativa e quantitativa dos processos celulares, mapear as interações da célula de modo global e, por último, prever como ela irá se comportar frente a uma condição (Mustacchi *et al.*, 2006). O crescimento das ferramentas de biologia de sistemas aumentou a habilidade de estudar rapidamente o comportamento celular (Nielson & Jewett, 2007). A transcriptômica é uma dessas ferramentas e permite a análise dos transcritos dos genes expressos em um dado estágio de desenvolvimento ou condição fisiológica, através da tecnologia de RNA-Seq, por exemplo. Essa tecnologia é considerada revolucionária por ser uma metodologia muito reprodutível e permitir a descoberta de novos genes e *splicing* alternativos (Wang *et al.*, 2009; Hiller *et al.*, 2009; Nagalakshmi *et al.*, 2010; Wang *et al.*, 2010; Oszolak & Milos, 2011). O RNA-Seq consiste basicamente em um sequenciamento massivo de moléculas de DNA derivadas dos RNA mensageiros presentes em uma dada amostra. Para isso, é necessário purificar as moléculas de RNAm contendo a cauda poli-A utilizando, por exemplo, beads com oligo-dT, fragmentar por hidrólise química ou enzimática (RNase III) o RNAm purificado e converter os fragmentos em bibliotecas de cDNA utilizando a transcriptase reversa. Em seguida, é feito o reparo das extremidades dos fragmentos de cDNA, ligação dos adaptadores e amplificação dos produtos por PCR em emulsão na superfície de beads (ion sphere particles) (Ion Torrent PGM™ System, Life Technologies™) ou em ponte (HiSeq™, Illumina®) (Wang *et al.*, 2009; Quail *et al.*, 2012). As bibliotecas de amplificadas obtidas são então sequenciadas por diferentes metodologias em sequenciadores de última geração com alto throughput, como as plataformas HiSeq™ (Illumina®), Ion Torrent PGM™ System (Life Technologies™) e PacBio RS System (Pacific

Biosciences®). Na primeira, o sequenciamento é realizado por síntese utilizando nucleotídeos marcados que emitem fluorescência à medida que são incorporados. No Ion Torrent, um sinal é detectado conforme prótons são liberados da incorporação dos nucleotídeos, e no PacBio, nucleotídeos ligados a um grupo fosfato marcado com fluoróforo são detectados à medida que são incorporados (Quail et al., 2012). Como resultado, são obtidas pequenas sequências de até 200 bases, denominadas reads, de cada cDNA amplificado a partir de uma extremidade (single-end sequencing) ou de ambas (pair-end). As reads são então pré-processadas para remoção das sequências dos adaptadores e seleção por tamanho mínimo e qualidade das bases sequenciadas. A seguir, o genoma de referência é indexado e utilizado para mapear as reads. Muitos estudos levam em consideração apenas as reads com um único alinhamento no genoma (Ries, et al., 2013; dos Santos Castro et al., 2014), mas para outras abordagens também é interessante analisar as reads com multi-mapeamento, como para estimar a expressão da isoforma de um gene (Turro et al., 2011) ou identificar small RNAs (Axtell, 2014; Otto et al., 2014). Após o mapeamento, as reads selecionadas de acordo com os parâmetros definidos são contabilizadas, normalizadas de acordo com o número total de reads, e então é calculada a expressão gênica diferencial de uma amostra, em uma determinada condição com base em diferentes testes estatísticos (Wang et al., 2009). Tanto o pré-processamento quanto a indexação do genoma, mapeamento das reads e análise da expressão diferencial dos genes pode ser feito por diversas ferramentas de bioinformática, dentre elas pode-se citar o Tophat, Cufflinks, Bowtie2, AlienTrimer, DESeq e edgeR (Anders & Ruber, 2010; Robinson et al., 2010; Langmead & Salzberg, 2012; Trapnell et al., 2012; Criscuolo & Brisse, 2013). Assim, a abordagem de RNA-Seq foi utilizada neste trabalho com o objetivo de analisar alterações na expressão gênica quando o fungo cresce utilizando biomassa recalcitrante como fonte de carbono, e também para tentar descobrir novos genes relacionados, direta ou indiretamente, com a hidrólise da biomassa, outras enzimas e proteínas acessórias, transportadores de açúcares e genes relacionados à regulação do sistema celulolítico de *T. reesei*, como os fatores de transcrição. A proteômica também tem se mostrado uma ferramenta valiosa por permitir a identificação global de proteínas sendo produzidas por uma célula em uma determinada condição (Kim et al., 2007). Dessa forma, o

sobrenadante do fungo *T. reesei* também foi investigado para identificar as proteínas secretadas na presença dos substratos, e várias proteínas identificadas por espectrometria de massas tiveram também seus transcritos detectados por RNA-Seq. Os resultados encontrados pela análise da transcriptômica e proteômica de *T. reesei* são congruentes e sustentam a resposta biológica de que esse fungo produz e secreta várias enzimas para quebrar a parede celular vegetal da cana. Os resultados obtidos da análise dos açúcares solúveis no sobrenadante apontam nessa mesma direção, o que confirma a importância de *T. reesei* como um ótimo fungo desconstrutor de biomassa e o interesse biotecnológico para produção do etanol 2G a partir do bagaço. Além disso, os resultados mostram que a integração das várias tecnologias existentes podem contribuir para uma maior compreensão de um fenômeno biológico a nível molecular e, assim, ajudar a superar os desafios biotecnológicos da produção do etanol 2G.



### 3. OBJETIVOS

O objetivo principal deste trabalho foi analisar a expressão gênica global de *T. reesei* RUT C30, crescido em bagaço e colmo de cana-de-açúcar, pela técnica de RNA-Seq. Como objetivo adicional, foi analisado o secretoma de *T. reesei*, crescido nas mesmas condições utilizadas para o transcriptoma. Como objetivos específicos, teve-se:

**(i)** Realizar e analisar o RNA-Seq de *T. reesei* crescido em bagaço e colmo de cana-de-açúcar por 6, 12 e 24 horas;

**(ii)** Identificar as enzimas hidrolíticas produzidas em diferentes *time points* e substratos, avaliando as diferenças e similaridades para a degradação de substratos distintos;

**(iii)** Identificar e estudar genes que apresentaram padrões de expressão interessantes, como por exemplo, sendo expressos unicamente na fonte de carbono recalcitrante e/ou que não tenha nenhuma função descrita relacionada (que possam estar envolvidos, direta ou indiretamente, na desconstrução da biomassa). O estudo pretendeu avaliar a expressão destes genes por PCR em tempo real, em diferentes substratos, na cepa RUT C30 e na cepa RUT C30 com o gene *cre1* íntegro (Ilmén *et al.*, 1996), para verificar diferenças na repressão catabólica.

**(iv)** Realizar e analisar o secretoma de *T. reesei* crescido em bagaço e colmo de cana-de-açúcar por 6,12 e 24 horas.



## 4. MATERIAIS E MÉTODOS

### 4.1 Linhagens de fungo, meios de cultura e substratos de cana-de-açúcar

Neste trabalho foram utilizadas as linhagens *T. reesei* RUT C30 (ATCC 56765) e a RUT C30 com o gene *cre1* íntegro (Ilmen *et al.*, 1996). O meio mínimo (MM) (pH 5,5) foi composto de 0,05% de extrato de levedura, 50 mL/L de solução de sais (6 g/L NaNO<sub>3</sub>, 1,5 g/L KH<sub>2</sub>PO<sub>4</sub>, 0,5 g/L KCl e 0,5 g/L MgSO<sub>4</sub>) e 200 µL/L de solução de oligoelementos (10 g/L ácido etilenodiamino tetra-acético (EDTA), 4,4 g/L ZnSO<sub>4</sub>·H<sub>2</sub>O, 1,0 g/L MnCl<sub>2</sub>·4H<sub>2</sub>O, 0,32 g/L CoCl<sub>2</sub>·6H<sub>2</sub>O, 0,315 g/L CuSO<sub>4</sub>·5H<sub>2</sub>O, 0,22 g/L (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>·4H<sub>2</sub>O, 1,47 g/L CaCl<sub>2</sub>·2H<sub>2</sub>O e 1 g/L FeSO<sub>4</sub>·H<sub>2</sub>O). De acordo com nossa condição experimental, foi adicionada uma concentração pré-definida da fonte de carbono. Para o cultivo em meio com bagaço explodido a vapor ou colmo de cana moído como fontes únicas de carbono, os micélios foram crescidos inicialmente em MM com 1% frutose e, em seguida, transferidos para MM sem extrato de levedura, adicionado de 0,5% de bagaço ou colmo de cana (pH 5,5) (vide 4.2). A frutose foi escolhida para o pré-inóculo por não induzir nem reprimir significativamente os genes das enzimas envolvidas na desconstrução de biomassa vegetal. Se o pré-inóculo fosse realizado com outras fontes de carbono, os genes das CAZymes poderiam ser induzidos (lactose, por exemplo) ou reprimidos (glicose) (Amore *et al.*, 2013), o que poderia influenciar a resposta inicial do fungo quando crescido em bagaço e colmo de cana. Todos os meios contendo o substrato recalcitrante ou frutose foram autoclavados a 121°C por 20 minutos para esterilização antes dos fungos serem inoculados ou transferidos.

O colmo *in natura* constituído por 32% de celulose e 48% de hemicelulose (Prof<sup>o</sup>. Dr<sup>o</sup>. Marcos Buckeridge, comunicação pessoal) foi picado, moído em um moinho de facas (modelo *pulverisette* 19, Fritsch), lavado seis vezes com etanol 80% (v/v), a 80°C, por 20 minutos, lavado com água destilada para remover o etanol remanescente e seco a 40°C. Embora o preparo deste material resulte no que também pode ser denominado bagaço, visualmente o colmo preparado desta forma é diferente do bagaço *in natura* resultante da usina. Além disso, ele apresenta uma composição de polissacarídeos diferentes, uma vez que no bagaço industrial os arabinoxilanos mais complexos (ramificados com arabinose e ácidos glucurônicos) são perdidos durante o processamento da cana para retirar a sacarose, ficando

principalmente os menos ramificados e uma proporção menor de glucanos de ligação mista, bem como de pectina (Prof<sup>o</sup>. Dr<sup>o</sup>. Marcos Buckeridge, comunicação pessoal). Por isso, ao longo do texto será utilizado o termo “colmo” ao invés de bagaço não tratado.

O bagaço *in natura* foi explodido a vapor por um tratamento com 14 kg/cm<sup>2</sup> de vapor de água durante 8 minutos, lavado exaustivamente com água destilada e seco a 40°C por vários dias. A proporção de celulose, hemicelulose e lignina para o bagaço explodido a vapor é de 47, 9 e 34%, respectivamente (Wanderley *et al.*, 2013). Depois de secos, tanto o bagaço explodido quanto o colmo moído foram peneirados em uma peneira industrial com poros de 600 µm. O bagaço explodido foi doado pelo Prof<sup>o</sup>. Dr<sup>o</sup>. João Atílio Jorge (USP – Ribeirão Preto) e o colmo *in natura* pelo Prof<sup>o</sup>. Dr<sup>o</sup>. Marcos Silveira Buckeridge (USP – São Paulo).

#### **4.2 Indução do crescimento do fungo em bagaço e colmo de cana-de-açúcar**

Primeiramente, o fungo foi inoculado em placa de petri com meio completo PDA (*potato dextrose agar*) de 5 a 10 dias a 29°C, exposto à luz artificial. Após este período, os esporos foram recolhidos das placas adicionando-se 1 mL de água destilada estéril, contados na câmara de Neubauer, e  $1 \times 10^6$  esporos foram inoculados em *erlemeyers* de 250 mL contendo 30 mL de MM com 1% de frutose como fonte de carbono, a 29°C por 48 horas, sob agitação de 200 rpm e expostos à luz. Os micélios foram então filtrados utilizando-se filtros Whatman de 11 µm (*GE Healthcare*), lavados com água estéril e transferidos para *erlenmeyers* de 250 mL contendo 30 mL dos meios com bagaço ou colmo de cana como fonte exclusiva de carbono, por 6, 12 e 24 horas para o experimento de RNA-Seq, e por 30 minutos, 2, 4, 6, 12 e 24 horas para as análises de PCR em tempo real. Como controle, ou condição “não induzida”, micélios crescidos em 30 mL de MM com 1% de frutose foram transferidos para 30 mL de novo meio MM + 1% frutose, por 24 horas. Para a coleta de cada ponto amostral, os micélios/substratos e sobrenadantes foram separados por filtração, lavados com água estéril e imediatamente congelados em nitrogênio líquido e armazenados a -80°C, para posterior extração de RNA (vide 4.3). Os sobrenadantes foram armazenados a -20°C para análise do secretoma, açúcares solúveis e atividade enzimática (Anexo I).

#### **4.3 Extração do RNA total, tratamento com DNase e síntese de cDNA**

O micélio congelado foi macerado até ficar na forma de pó. Os instrumentos utilizados nesta etapa (cadinho, pistilo e espátula) foram anteriormente tratados com o reagente RNaseZap (*Applied Biosystems, Life Technologies*) e aquecidos a 180°C por, no mínimo, quatro horas para eliminar contaminação com RNase. A extração do RNA total dos micélios foi realizada com o kit *RNeasy Plant Mini* (Qiagen), partindo de 100 mg de cada amostra, seguindo as recomendações do fabricante. Ao final da extração, o RNA foi aliquoteado para a quantificação no espectrofotômetro *NanoDrop 2000* (*Thermo Fisher Scientific*) e para verificar sua qualidade no *Bioanalyzer* (*Agilent 2100 Bioanalyzer, Agilent Technologies*). Após a confirmação da integridade do RNA, as amostras de RNA total para transcriptômica foram enviadas para o sequenciamento (RNA-Seq, vide item 4.4), enquanto que as amostras para validação da expressão gênica por PCR em tempo real (qPCR) foram tratadas com DNase, utilizando o kit *TURBO DNA-free* (*Applied Biosystems, Life Technologies*), seguindo as recomendações do fabricante. Após o tratamento, o RNA foi testado para verificar a presença de DNA genômico através de uma reação de PCR (*polymerase chain reaction*).

A reação de PCR foi composta de: 1,0 µL de RNA total ou tratado com DNase; 2,0 µL de *Taq Buffer* 10x, 1,6 µL de MgCl<sub>2</sub> (25 mM), 0,8 µL de dNTP (10 mM), 0,4 µL de oligonucleotídeos (oligos) *forward* e *reverse* (10 pmol/µL); 0,1 µL de *Taq DNA Polymerase* (5 U/µL) (Fermentas, Thermo Scientific) e 13,7 µL de água livre de nucleases para completar o volume final para 20 µL. Os oligos utilizados anelam-se na região ITS (*ribosomal internal transcribed spacer*) do DNA nuclear: ITS3 (5'-GCATCGATGAAGAACGCAGC-3') e ITS4 (5'-TCCTCCGCTTATTGATATGC-3'). Estes oligos são específicos para fungos e vêm sendo aceitos como um marcador padrão para a identificação destes microrganismos (Schoch *et al.*, 2012). Como controle positivo foi utilizado o DNA genômico extraído de *T. reesei*. As amostras foram amplificadas em um termociclador com a seguinte ciclagem: 94°C por 3 minutos (desnaturação do DNA); 35 ciclos de 94°C por 30 segundos; 56°C por 30 segundos (anelamento dos oligos); 72°C por 1 minuto (extensão); e uma extensão final de 72°C por 10 minutos. Em seguida, as amostras foram misturadas ao *loading buffer* 5x (50% glicerol, 0,125% azul de bromofenol, 0,125% xileno cianol e 1 mM EDTA) e ao fluoróforo GelRed™

10.000x (Biotium) diluído 1:500 em água destilada, e aplicadas em um gel de agarose 1% (m/v) imerso em tampão TAE 1x (Tris-Acetato-EDTA). O fluoróforo GelRed™ é um corante fluorescente de ácidos nucleicos, ambientalmente seguro, utilizado para coloração de dsDNA, ssDNA e RNA em géis de agarose (Biotium). A eficiência do tratamento com DNase foi avaliada pela ausência de bandas do produto ITS.

Uma vez confirmado que o tratamento com DNase foi eficiente, 400 ng do RNA foram convertidos em cDNA com o kit *SuperScript™ II Reverse Transcriptase* (Invitrogen, Life Technologies), seguindo as recomendações do fabricante. Para confirmação da síntese, o cDNA foi testado por PCR com os oligos ITS e a eficiência da síntese de cDNA foi avaliada pela presença do produto esperado.

#### **4.4 Sequenciamento na plataforma Illumina® e análise das sequências**

Após ter feito a extração, quantificação e análise da integridade do RNA total das amostras em duplicata, o material foi enviado para uma *facility* de sequenciamento *high-throughput*, localizada no USC *Epigenome Center*, Califórnia, Estados Unidos. As amostras foram enriquecidas para RNA mensageiro, as bibliotecas construídas com o kit *TruSeq RNA Sample Prep v2* (Illumina) e os transcritos sequenciados na plataforma Illumina® HiSeq2000, utilizando 7 *lanes* (4 amostras por *lane*) e gerando *single-end reads* de 50 pb. Os dados brutos gerados foram encaminhados e processados com a ajuda do biólogo computacional Dr<sup>o</sup>. Diego Maurício Riaño-Pachón (CTBE). As *reads* foram pré-processadas no AlienTrimer (Criscuolo & Brisse, 2013) para remoção dos adaptadores e triagem das *reads* por tamanho (mínimo de 40 pb) e qualidade das bases (Q>20). Em seguida, o genoma de *T. reesei* RUT C30, disponível no JGI ([http://genome.jgi.doe.gov/TrireRUTC30\\_1](http://genome.jgi.doe.gov/TrireRUTC30_1)), foi indexado pela ferramenta Bowtie2 (Langmead & Salzberg, 2012) e usado para mapear as *reads* com o TopHat2 (Kim *et al.*, 2013). A partir do mapeamento, foi possível analisar a profundidade do sequenciamento através da saturação das junções de *splicing* anotadas para o genoma pelo programa RSeQC (Wang *et al.*, 2012). Para contagem das *reads* mapeadas, a função *featureCounts* do pacote *Rsubread* v1.12.6 (<http://bioconductor.org/>) foi executada no R Project v3.0.2 (2013) (<http://www.r-project.org/>), utilizando como entrada os arquivos de extensão BAM gerados

no mapeamento. As *reads* atribuídas aos genes foram então normalizadas em RPKM (*reads per kilobase of transcript per million reads mapped*) (Mortazavi *et al.*, 2008), segundo a fórmula abaixo:

$$RPKM = \frac{R}{(T * B)}$$

Onde *R* é o número de *reads*, *T* o tamanho do gene em kilobases (kb) e *B* o tamanho total da biblioteca em megabases (Mb) (considerando tanto *assigned* quanto *unassigned reads*). Para análise da expressão diferencial, as *reads* normalizadas em RPKM das amostras crescidas em colmo e bagaço foram divididas então pelas *reads* normalizadas da frutose, e o *fold change* calculado foi convertido em  $\log_2$ . Os genes diferencialmente expressos foram classificados em hiper ( $\log_2(\textit{fold change}) \geq 1$ ,  $\textit{fold change} \geq 2$ ) e hipo-expressos ( $\log_2(\textit{fold change}) \leq -1$ ,  $\textit{fold change} \leq -2$ ), e somente aqueles com RPKM  $\geq 1$  em colmo e bagaço foram considerados. A seguir, os domínios funcionais das proteínas foram buscados manualmente no banco de dados específico para a cepa *T. reesei* RUT C30 v1.0 ([http://genome.jgi-psf.org/TrireRUTC30\\_1](http://genome.jgi-psf.org/TrireRUTC30_1)), no NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), InterPro (<http://www.ebi.ac.uk/interpro/>) e também no Pfam (<http://pfam.xfam.org/>) para verificar a provável função delas. Para verificar se o produto de um gene pode ser secretado, foi observada a presença de peptídeo sinal pelo programa de predição SignalP v4.1 (<http://www.cbs.dtu.dk/services/SignalP/>). A presença de domínios transmembrânicos (TMH) que prediz se uma proteína está ancorada ou não em membranas, como transportadores, foi verificada pelo programa TMHMM v2.0 (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>), e os genes de enzimas hidrolíticas ou de proteínas acessórias supostamente ligadas à desconstrução da parede vegetal foram classificados de acordo com o banco de dados do dbCAN (<http://csbl.bmb.uga.edu/dbCAN/>), que utiliza o banco específico do CAZy (<http://www.cazy.org>) para anotar as proteínas.

#### **4.5 Validação da expressão dos genes por PCR em tempo real (qPCR)**

Após as análises do resultado do RNA-Seq do *T. reesei* e levantamento bibliográfico de genes com função regulatória, foram selecionados 19 genes (Tabela 1) para validação por

PCR em tempo real (qPCR). Entre os genes alvo, encontram-se cinco fatores de transcrição, uma metiltransferase com papel regulatório, seis enzimas hidrolíticas da família das glicosil hidrolases (GHs), três proteínas acessórias e dois transportadores. Os genes de referência usados como controle interno foram o da  $\beta$ -actina (*act*) e o que codifica para uma GTPase pequena (*sar1*). Estes genes são aceitos como estáveis e o uso de ambos aumenta a confiabilidade na análise dos resultados (Steiger *et al.*, 2010). Os oligos para qPCR foram desenhados utilizando o *software Primer Express v3.0.1* (Applied Biosystems, Life Technologies) a partir da sequência CDS (*coding DNA sequence*) obtida de cada gene no banco de dados do *Trichoderma reesei* RUT C30 v1.0 ([http://genome.jgi-psf.org/TrireRUTC30\\_1](http://genome.jgi-psf.org/TrireRUTC30_1)). A sequência dos oligos utilizados, a temperatura de *melting* e tamanho do produto são descritos na Tabela 1.



**Tabela 1.** Relação dos oligonucleotídeos (oligos), temperatura de *melting* (*T<sub>m</sub>*) e tamanho esperado do produto de 19 genes analisados por PCR em tempo real. O ID JGI é referente ao número da proteína no banco de dados *Trichoderma reesei* RUT C30 v1.

Gene	Descrição/ Família	ID JGI	Oligo forward (5' → 3')	Oligo reverse (5' → 3')	<i>T<sub>m</sub></i> (°C)	Tamanho do produto (bp)
Regulador da xilanase I ( <i>xyr1</i> )	Fator de transcrição	98788	GCTTGTGACCAGTGCAACCA	CCAAGGCCGAATTCTATACAATG	57	79
Proteína de repressão catabólica por carbono ( <i>cre1</i> )	Fator de transcrição	23706	CGTGTGGGTGCGAATGTG	CGCGGTCACCGTCATCA	57	150
Fator de transcrição ( <i>ace1</i> )	Fator de transcrição	122363	AGTCCGCCGCTTATACCTT	GCGGTCGGTGAGATATTCGA	57	79
Fator de transcrição ( <i>ace2</i> )	Fator de transcrição	32395	AGATACGGGTGGGCACTTTTG	CGTTTGAAGCTCCACGACAA	56	79
FT não caracterizado	Fator de transcrição	26638	ATGACGGGAGTGCACTGCTT	CGTGTACGGCAGGTCGAGAT	58	79
Metiltransferase ( <i>lae1</i> )	Metiltransferase	9778	ATGCCGCGATGGGAGAAT	TGCATGGCACGTTGGAATAG	55	79
Celobiohidrolase I ( <i>cel7a/cbh1</i> )	Glicosil hidrolase (GH7)	125125	AGTGTCGCCGCGATCTGA	CGTGTCGCGTTGTTGGAT	56	79
Celobiohidrolase II ( <i>cel6/cbh2</i> )	Glicosil hidrolase (GH6)	122470	GCTGGTCCAACGCCTTCTT	CCAGTCTCCCCACTGTTGCT	58	79
B – glicosidase ( <i>cel3c</i> )	Glicosil hidrolase (GH3)	125268	CGTCGCTGCGTATTTTGATG	TGCTGTCGCTCACGATAACC	56	79
B – xilosidase ( <i>bx11</i> )	Glicosil hidrolase (GH3)	140746	CCCATATACGGCCATGCT	GAATCCGACGAGCCACTTGT	57	79
Xilanase 1 ( <i>xyn1</i> )	Glicosil hidrolase (GH11)	38418	CAGCCTTCGACAGCGACAA	GGGCACAGCGACCTTCAA	58	149
Xilanase 2 ( <i>xyn2</i> )	Glicosil hidrolase (GH11)	124931	GGTGACCTCGCCCAGCTT	ATCAACTTCTCGGGCAGCTACA	58	149
Hidrofobina 3 ( <i>hfb3</i> )	Proteína acessória	125146	CAACCCGCTGTGCTGTGA	ACAGACGCTGCCAAAGGACTT	57	102
Swolenina ( <i>swo1</i> )	Proteína acessória	104220	GGACCCTCGTGATAATGTTTG	GGTACCGCAGCCTGAGCAT	58	79
Proteína induzida por celulose ( <i>cip1</i> )	Proteína acessória	121449	CAATGGCTACTGTGGACACATCTT	GCCGAATCCAAGCTCTGACA	58	79
Transportador 1 (T1)	Transportador MFS	138519	ATCAATGGCACAAGCCTGAAC	GGTTGGTTGACGGCCATTC	56	149
Transportador 2 (T2)	Transportador MFS	109243	CAGCTCCAGCTCTTCGAGTATGA	AAGAGCGCCCGTAATCC	58	79
Actina ( <i>act</i> )	Controle endógeno	113342	TGAGAGCGGTGGTATCCACG	GGTACCACCAGACATGACAATGTTG	60	103
GTPase do tipo SAR/ARF ( <i>sar1</i> )	Controle endógeno	8999	TGGATCGTCAACTGGTCTACGA	GCATGTGTAGCAACGTGGTCTTT	58	115

Para a construção da curva padrão da PCR em tempo real, os fragmentos de cada gene foram primeiramente amplificados por PCR a partir do cDNA sintetizado de alguns tempos amostrais. Após a amplificação, as amostras foram aplicadas em um gel de agarose 2%, e após a eletroforese as bandas foram purificadas com o kit *PureLink Quick Gel Extraction* (Invitrogen, Life Technologies). Os produtos de PCR purificados foram sequenciados na plataforma *3500xL Genetic Analyzer* (Applied Biosystems, Life Technologies) para confirmar se eram os produtos de interesse, utilizando o kit *BigDye® Terminator v3.1 Cycle Sequencing* (Applied Biosystems, Life Technologies), seguindo as instruções do fabricante. Para confirmação das sequências de interesse, foram feitas buscas do resultado do sequenciamento com a ferramenta BLAST do NCBI (<http://blast.ncbi.nlm.nih.gov/>) ou do banco de dados do JGI (*Joint Genome Institute*). A curva padrão de cada gene foi construída a partir de diluições seriadas na base 10 ( $10^{-2}$  até  $10^{-8}$ ) do respectivo produto de PCR purificado. Foram aceitas curvas padrões com  $R^2 \geq 0,99$  e eficiência de amplificação dos oligos  $\geq 85\%$ .

Todos os experimentos de PCR em tempo real foram realizados no sistema *ViiA 7 Real Time PCR* (Applied Biosystems, Life Technologies). As reações de amplificação foram feitas em placas de *384-wells* em um volume final de 10  $\mu$ L utilizando o kit *SYBR® Green PCR Master Mix* (Applied Biosystems, Life Technologies), com 2 pmol de oligos *forward* e *reverse* (Tabela 1) e um volume diluído 10x do cDNA preparado anteriormente. O protocolo de amplificação consistiu de um passo inicial de ativação (10 minutos a 95°C), seguido de 40 ciclos de desnaturação (15 segundos a 95°C), anelamento e extensão (1 minuto a 60°C). Os ensaios foram feitos em triplicata. Os controles negativos sem o respectivo DNA foram submetidos ao mesmo procedimento para excluir ou detectar qualquer contaminação possível. A especificidade dos produtos foi analisada pela curva de dissociação gerada pelo *software ViiA™ 7* (Life Technologies). A normalização foi realizada dividindo-se a média da quantidade de transcrito de cada gene (obtidas nas réplicas) pela média das quantidades dos transcritos dos genes normalizadores (*act* e *sar1*). A quantificação da expressão gênica foi feita relacionando a quantidade de produto inicial cada gene com o valor de Ct (*cycle threshold*), baseando-se na curva padrão. As réplicas biológicas foram comparadas pelo teste estatístico

t *student* não-pareado, utilizando a ferramenta QuickCalcs do *software online* GraphPad (<http://graphpad.com/>) e considerando valores de  $p < 0,05$  como uma diferença estatisticamente significativa (mais de 95% de confiança).

#### **4.6 Análise do secretoma**

Além da análise do transcriptoma do fungo crescido em colmo e bagaço, foram analisados também o secretoma e os açúcares liberados da parede vegetal no sobrenadante. A metodologia utilizada está descrita no manuscrito submetido para a revista ***PLoS ONE*** (Anexo I). Sucintamente, o sobrenadante separado do micélio/substrato foi utilizado para análise das proteínas secretadas por espectrometria de massas (LTQ Velos Orbitrap), para ensaios enzimáticos com substratos específicos e para quantificar os açúcares solúveis liberados pela degradação da biomassa por cromatografia de troca iônica de alta resolução (HPAEC-PAD). Os resultados desta parte do trabalho também são apresentados no manuscrito (Anexo I).



## 5. RESULTADOS E DISCUSSÃO

### 5.1 Transcriptômica de *T. reesei* crescido em biomassa de cana-de-açúcar

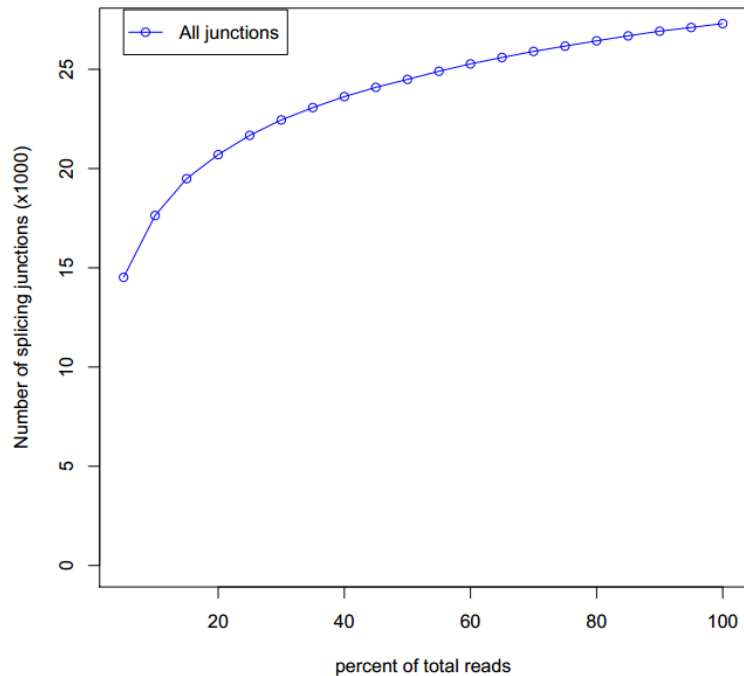
A fim de avaliar o perfil transcricional dos genes do fungo *T. reesei* crescido em bagaço e colmo de cana-de-açúcar, foi realizado o sequenciamento de seu transcriptoma por RNA-Seq após 6, 12 e 24 horas de crescimento. Esses pontos foram escolhidos pois a maioria dos trabalhos já realizados avalia o transcriptoma e secretoma de fungos desconstrutores de biomassa apenas após 24 horas de crescimento (Nitsche *et al.*, 2012; Ivanova *et al.*, 2013; Marx *et al.*, 2013), o que impede uma visão dos genes que estão sendo regulados e expressos em tempos mais iniciais e que fazem parte da resposta do fungo frente à presença do substrato recalcitrante. Como controle, ou condição não induzida, foi escolhida a condição de crescimento em frutose por 24 horas.

Para cada uma das réplicas e condição experimental foram obtidas, em média, mais de 30 milhões de *reads* (Tabela 2). Do número total de *reads* com qualidade que foram selecionadas segundo os parâmetros adotados (vide item 4.4), mais de 95% foram mapeadas unicamente contra o genoma de *T. reesei*, com exceção de uma das réplicas da condição colmo 6 horas (88,4%) (Tabela 2). Uma parcela insignificante foi mapeada contra os RNAs ribossômicos (dados não mostrados), o que mostra o sucesso na retirada destas moléculas para a construção das bibliotecas do sequenciamento. A porcentagem de *reads* que apresentaram mapeamento múltiplo, ou seja, em mais de um gene e que portanto foram desconsideradas da análise, foi insignificante (0,1-0,2%). Levando-se em consideração o tamanho das *reads* (50 pb), foram obtidos dados suficientes para cobrir, no mínimo, mais de 50 vezes o tamanho do genoma de *T. reesei* (32,69 Mpb) (Le Crom *et al.*, 2009), em todas as amostras (Tabela 2). Comparando esses resultados com os descritos em outros trabalhos que utilizaram a mesma técnica (RNA-seq), com o mesmo fungo e em materiais lignocelulósicos complexos (Ries *et al.*, 2013; Chen *et al.*, 2014), nota-se que os dados aqui obtidos foram superiores em número e cobertura, o que aumenta a confiabilidade dos resultados. O número de junções *exon-exon*, indicativo da profundidade do sequenciamento (Wang *et al.*, 2012), também foi avaliado (Figura 5). Observa-se que o número de junções detectadas não aumentaria substancialmente se houvesse um número maior de *reads* sequenciadas, o que

indica que a profundidade do sequenciamento está adequada para todas as amostras (Wang *et al.*, 2012) (Figura 5).

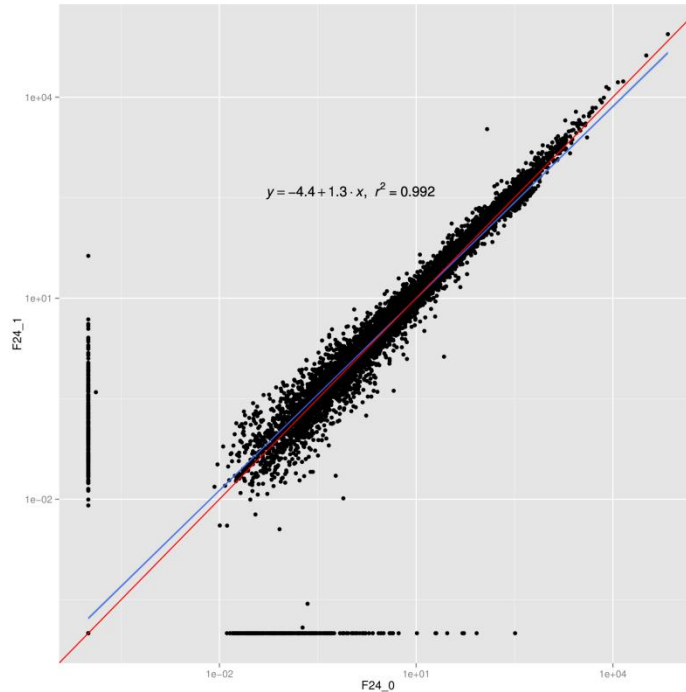
**Tabela 2.** Número de *reads* brutas e filtradas por qualidade geradas pelo sequenciamento do transcriptoma de *T. reesei*, porcentagem de *reads* mapeadas e cobertura do genoma. O transcriptoma foi analisado após crescimento de *T. reesei* em colmo (C) e bagaço (B), após 6, 12 e 24 horas (h). Como controle, foi utilizada a condição frutose (F) 24 horas.

Amostra	Réplica	<i>Reads</i> brutas	<i>Reads</i> filtradas	<i>Reads</i> mapeadas (%)	Cobertura do genoma (vezes)
B6h	1	49.367.513	46.421.218	96,7	75,51
B6h	2	48.653.781	45.731.691	96,3	74,42
B12h	1	46.933.558	44.288.183	97,2	71,79
B12h	2	53.590.571	49.947.954	97,1	81,97
B24h	1	49.305.453	46.428.111	95,5	75,41
B24h	2	50.560.691	47.556.369	95,2	77,33
C6h	1	50.563.846	47.901.115	97,9	77,34
C6h	2	53.473.629	50.900.402	88,4	81,79
C12h	1	34.915.261	33.200.447	97,2	53,40
C12h	2	34.615.720	32.956.899	97,4	52,95
C24h	1	33.029.352	31.083.522	97,1	50,52
C24h	2	32.347.500	30.362.215	97,1	49,48
F24h	1	47.185.370	44.387.600	97,0	72,17
F24h	2	50.655.259	47.771.315	96,9	77,48



**Figura 5.** Curva de saturação das junções *exon-exon* de uma réplica da condição frutose 24 horas, mostrando a profundidade do sequenciamento do RNA-Seq. Observa-se que o aumento de *reads* sequenciadas não implica em um maior número de junções *exon-exon*. O mesmo padrão foi obtido para as demais amostras.

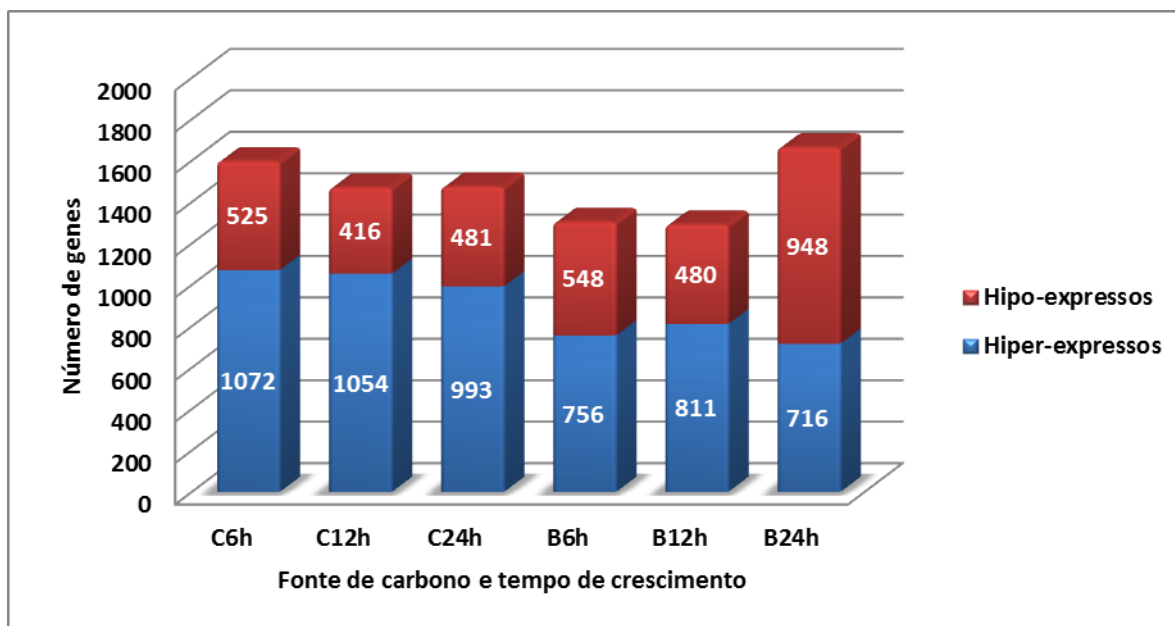
Com o objetivo de avaliar a reprodutibilidade e consistência das réplicas biológicas analisadas por RNA-Seq, foi calculada a correlação de Pearson para cada uma das condições estudadas (Figura 6). A correlação obtida ( $r^2$ ) foi alta em todas as condições ( $r^2 > 0,99$ ), sendo igual ou superior à correlação obtida por outros trabalhos utilizando RNA-Seq (dos Santos Castro *et al.*, 2014; Wu *et al.*, 2014), demonstrando, mais uma vez, a qualidade dos dados obtidos neste trabalho.



**Figura 6.** Gráfico mostrando a correlação de Pearson ( $r^2$ ) obtida para as réplicas biológicas da condição frutose 24 horas. Uma alta correlação foi obtida entre as réplicas, o que demonstra a consistência dos dados do RNA-Seq. O mesmo também foi observado para as demais amostras (dados não mostrados).

Após o mapeamento das *reads* contra o genoma, foi calculada a expressão diferencial dos genes na condição de crescimento em bagaço e colmo *versus* a em frutose. Os genes foram então classificados levando-se em consideração um *fold change* de no mínimo 2, em hiper ( $\log_2(\text{fold change}) \geq 1$ ) ou hipo-expressos ( $\log_2(\text{fold change}) \leq -1$ ), como representado na Figura 7. Com exceção da condição bagaço 24 horas, foi observada uma maior quantidade de genes sendo hiper-expressos do que hipo-expressos, sendo que em colmo esse número foi cerca de duas vezes maior nos três tempos analisados (Figura 7).





**Figura 7.** Número de genes de *T. reesei* diferencialmente expressos em colmo (C) e bagaço (B) em relação à frutose, após 6, 12 e 24 horas (h) de crescimento.

Como um dos objetivos é entender como o fungo altera seu metabolismo para degradar polissacarídeos de origem vegetal, inicialmente a análise do transcriptoma foi realizada dando ênfase para os genes hiper-expressos, uma vez que os substratos lignocelulósicos utilizados devem estar induzindo a expressão de enzimas desconstrutoras da parede celular vegetal, bem como transportadores e reguladores transcricionais envolvidos direta ou indiretamente com a degradação de biomassa. Por questão de espaço, no Anexo II são apresentados apenas os 50 genes mais hiper-expressos para cada condição de crescimento analisada. Desse total, pode-se perceber que as CAZymes (em **laranja**), proteínas acessórias (**azul**) e os transportadores hipotéticos de açúcar (**verde**) são os genes mais hiper-expressos em praticamente todas as condições de crescimento, sendo as CAZymes mais numerosas que os transportadores e proteínas acessórias em todas condições (Anexo II). De modo interessante, pode-se notar o aumento da expressão dos transportadores e CAZymes ao longo dos tempos analisados. Uma permease hipotética de xilose (jgi|33630), por exemplo, foi hiper-expressa com um *fold change* de 205, 376 e 1022 vezes ( $\log_2(\textit{fold change})$  igual a 7,7, 8,6 e 10) após 6, 12 e 24 horas de crescimento em colmo, respectivamente. Em bagaço, a endo- $\beta$ -1,4-xilanase 2 da família GH11 (*xyn2*, jgi|124931)

apresentou o mesmo perfil de expressão: começou com um *fold change* de 92 vezes ( $\log_2(\textit{fold change})$  igual a 6,5) no ponto 6 horas e terminou com um *fold change* de mais de 466 vezes ( $\log_2(\textit{fold change})$  igual a 8,9) após 24 horas de crescimento. Comparando os 50 genes mais hiper-expressos nos dois substratos, é possível notar que a grande maioria (40) foi compartilhada entre o colmo e o bagaço (Anexo II).

Para direcionar a análise de acordo com os objetivos, foi dada ênfase a algumas categorias de genes: (i) CAZymes, (ii) transportadores, (iii) proteínas acessórias, (iv) fatores de transcrição e (v) genes de função desconhecida. A primeira categoria é importante pois envolve os genes que codificam para proteínas capazes de degradar, direta ou indiretamente, a parede celular vegetal. De acordo com a sequência de aminoácidos, proteínas envolvidas na quebra, modificação ou síntese de ligações glicosídicas podem ser classificadas em uma das seis classes do banco de dados CAZy: glicosil hidrolases (GH), liases (PL), esterases (CE), glicosil transferases (GT), proteínas não-hidrolíticas com atividade auxiliar (AA) ou com domínio de ligação a carboidrato (CBM) (Cantarel *et al.*, 2009; Levasseur *et al.*, 2013). Contudo, nem todas estão envolvidas na degradação da biomassa vegetal, como, por exemplo, a classe das glicosil transferases (GT) que atuam no remodelamento da parede celular e as esterases da família 10 (CE10) que possuem atividade contra outros substratos que não carboidratos (Häkkinen *et al.*, 2012). Neste contexto, as proteínas acessórias (iii) também auxiliam na degradação da biomassa, mas sem apresentarem atividade hidrolítica.

Como produto da desconstrução da parede pelas CAZymes e proteínas acessórias, vários açúcares menores são liberados e então transportados por proteínas ancoradas na membrana plasmática do fungo para serem assimilados e gerar energia para manutenção das atividades celulares. Esses transportadores apresentam um potencial biotecnológico que pode ser explorado, por exemplo, para o melhoramento genético de cepas industriais utilizadas na produção de etanol. *Saccharomyces cerevisiae* não é uma boa assimiladora de pentose (Saloheimo *et al.*, 2007). Assim, a construção de cepas com transportadores eficientes de xilose, por exemplo, é de extrema importância para a fermentação industrial dos hidrolisados derivados do material lignocelulósico. Adicionalmente, além de atuarem

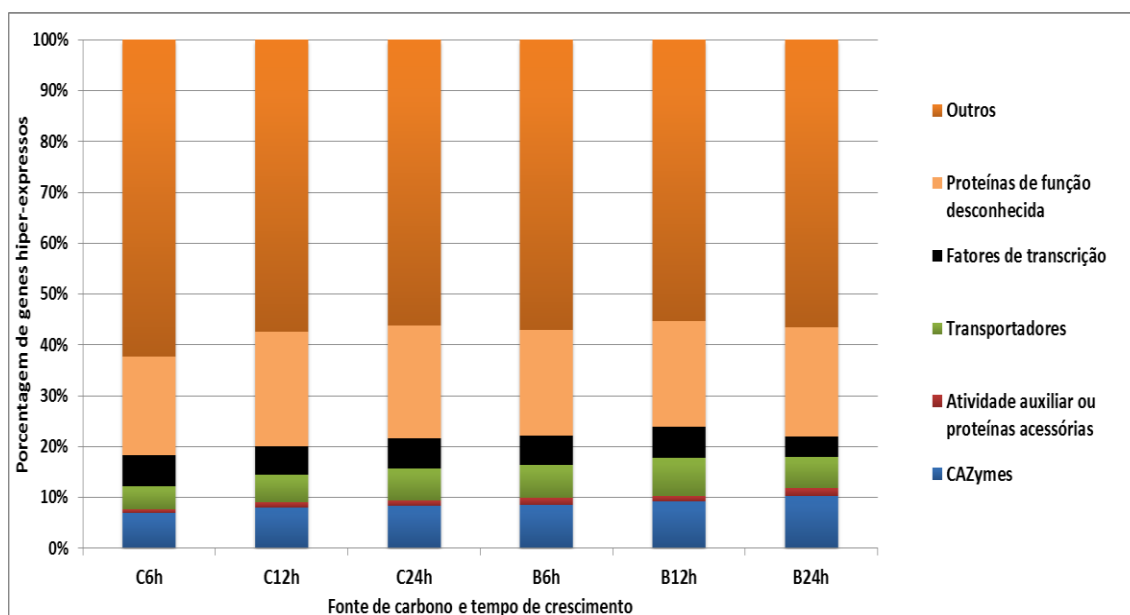
como transportadores de açúcares, foi mostrado recentemente que supostos transportadores podem atuar como sinalizadores dessas moléculas, transmitindo sinais externos para dentro da célula (Zhang *et al.*, 2013), o que ressalta ainda mais a importância de se estudar essas proteínas. Por essas razões, foram investigados transportadores hipotéticos de açúcar (*ii*) no transcriptoma de *T. reesei*.

Além disso, os açúcares transportados para o meio intracelular também atuam como indutores da expressão de genes de enzimas lignocelulolíticas, como xilose e celobiose (Amore *et al.*, 2013), pela ativação de fatores de transcrição e outros reguladores transcricionais. Em *T. reesei*, a regulação transcricional dos principais genes que codificam a enzimas hidrolíticas (xilanasas, celobiohidrolases e endoglucanases) é dependente do ativador XYR1 (Rauscher *et al.*, 2006) e está sujeita à repressão catabólica por carbono mediada pelo repressor CRE1 (Portnoy *et al.*, 2011). Outros fatores de transcrição fazem o *fine-tuning* dessa regulação, como por exemplo, ACE1, ACE2 e ACE3 (Aro *et al.*, 2001; Häkkinen *et al.*, 2014), e além destes, outros fatores vêm sendo identificados (Kubicek, 2012) e patenteados em processos para produção industrial destas enzimas (Pakula *et al.*, 2011a, b, c). Assim, dada a importância de *T. reesei*, principalmente como produtor de hidrolases, um melhor entendimento sobre a expressão dos genes que regulam esta produção é de extrema valia e, em razão disso, houve um interesse especial em identificar fatores de transcrição (*iv*) sendo hiper-expressos neste trabalho.

Por fim, os genes hiper-expressos de proteínas que não tinham função determinada quando buscadas contra o banco de dados geral do NCBI (<http://blast.ncbi.nlm.nih.gov>), ou que apresentaram um domínio de função desconhecida (como por exemplo, *DUF domain*) foram agrupados em proteínas de função desconhecida (*v*). Analisando suas sequências, foi possível notar que muitas dessas proteínas têm peptídeo sinal ou domínios transmembrana (TMH) e, portanto, podem ser secretadas para o meio extracelular ou estar ancoradas na membrana plasmática, respectivamente. Dessa forma, esses genes hiper-expressos são candidatos interessantes para novas proteínas envolvidas na degradação de substratos recalcitrantes ou no transporte de moléculas. Os genes não relacionados com nenhum dos processos biológicos descritos acima foram agrupados na categoria “outros”, inclusive

transportadores envolvidos com transporte de aminoácidos e outros solutos já caracterizados que não fossem açúcar, como, por exemplo, os transportadores de ferro (jgi|131826) e de oligopeptídeos (jgi|94554) hiper-expressos em colmo (Anexo II).

A Figura 8 mostra a proporção de genes hiper-expressos agrupados por função, em cada condição de crescimento. Em todas as condições, a maioria dos genes hiper-expressos pertence à categoria “outros”, que em geral são genes referentes a processos metabólicos gerais do fungo, alterados pela condição de crescimento imposta. Em seguida, a ordem decrescente de categorias é: proteínas de função desconhecida (19-22%), CAZymes (6-10%), transportadores hipotéticos de açúcar (4-7%), fatores de transcrição (4-6%) e, por último, proteínas acessórias (1-2%). Essa maior proporção de CAZymes é interessante uma vez que uma pequena parcela do genoma de *T. reesei* codifica a elas (314 genes) (Häkkinen *et al.*, 2012; dos Santos Castro *et al.*, 2014).



**Figura 8.** Proporção de genes hiper-expressos identificados no transcriptoma de *T. reesei* crescido em colmo (C) e bagaço (B), após 6, 12 e 24 horas (h). Os genes foram agrupados em seis classes diferentes de acordo com sua função.

### 5.1.1 Perfil transcricional de CAZymes e proteínas acessórias

Fazendo uma análise qualitativa do número de genes, de 7 a 12% dos genes hiper-expressos pertencem às CAZymes e proteínas acessórias em todas as condições analisadas (Figura 8, Tabela 3), existindo ainda a possibilidade desse número ser maior se forem

descobertas outras dentro das proteínas de função desconhecida (Anexo IV). No total, foram identificados 127 genes que codificam a CAZymes e proteínas acessórias hiper-expressos, sendo 82 GHs (32 famílias), 8 GTs (5 famílias), 19 CEs (sendo 9 da família CE10), 10 AAs (3 famílias), 2 PLs (2 famílias), 3 proteínas com atividade auxiliar hipotética e 3 proteínas acessórias (duas da família CBM1 e uma hidrofobina sem classificação no banco de dados CAZy mas que pode estar envolvida na degradação de biomassa (Tabela 3).

Considerando o genoma de *T. reesei*, 44% de todos os genes de CAZymes e proteínas acessórias (excluindo-se os genes que codificam para proteínas com atividade auxiliar hipotética, glicosil transferases e a família CE10) foram hiper-expressos no transcriptoma do fungo em pelo menos uma das condições avaliadas, sendo que as CEs e GHs foram as mais representativas em relação à quantidade de genes no genoma (45 e 42%, respectivamente) (Tabela 3) (Häkkinen *et al.*, 2012; Ries *et al.*, 2013; Wang *et al.*, 2015). Dos 127 genes, 91 (71,6%) foram hiper-expressos e compartilhados pelos dois substratos, 24 (18,9%) estavam presentes apenas no colmo e 12 (9,4%) no bagaço (Tabela 3). Exclusivamente em colmo, foram encontradas diversas GHs (manosidases, fucosidase, galactosidase, quitinase, endo-glucanase), proteínas AAs envolvidas na degradação de lignina (família AA3), várias GTs e esterases da família CE10 que não têm atividade contra carboidratos.

**Tabela 3.** Genes que codificam a CAZymes e proteínas acessórias hiper-expressos identificados no transcriptoma e secretoma de *T. reesei* crescido em bagaço (B) e colmo (C), após 6, 12 e 24 horas (h). As proteínas foram classificadas de acordo com o CAZy, e a identificação (ID) de cada proteína é a disponível no banco de dados do JGI do *T. reesei* RUT C30 v1.0. O tipo de substrato preferencial de cada proteína, a presença de peptídeo sinal e se as proteínas foram encontradas no secretoma também são indicados (ver Anexo I).

Protein ID	Annotation	CAZy family	Substrate	Log <sub>2</sub> (fold change)						SignalP	Secretome
				C6h	C12h	C24h	B6h	B12h	B24h		
70043	Candidate endo-1,3-β-glucanase	GH16	1,3-β-Glucan	-	1,3	1	-	1,1	-	-	-
124438	Endo-β-1,4-glucanase III, egl3/cel12a	GH12	1,4-β-Glucan	-	7,4	7,6	1,9	5,6	8,9	Yes	-
25940	Endo-β-1,4-glucanase, egl5/cel45a	GH45	1,4-β-Glucan	4	6,1	9,3	3,1	5,6	8,5	Yes	-
72489	Endo-β-1,4-glucanase II, egl2/cel5a	GH5	1,4-β-Glucan	2,7	5,3	8,9	3,4	5,9	8,5	Yes	Yes
5304	Endo-β-1,4-glucanase I, egl1	GH7	1,4-β-Glucan	4	5,2	9,3	2,8	5,2	8,3	Yes	Yes
136460	Candidate glucan endo-1,3(4)-β-D-glucosidase	GH16	1,3-1,4-β-Glucan	-	-	-	1,1	1,1	1,2	-	-
100583	Candidate glucan endo-1,3(4)-β-D-glucosidase	GH16	1,3-1,4-β-Glucan	1,3	1,7	1,4	1,9	1,1	2,7	Yes	-
56838	Candidate endo-beta-1,3(4)-β-D-glucanase	GH16	1,3-1,4-β-Glucan	1,2	1,2	-	-	-	-	Yes	-
139633	Lytic polysaccharide monoxygenase, cel61a/egl4	AA9 (GH61)	Cellulose	3,4	5,6	9,5	4	6,3	8,6	Yes	Yes
95463	Candidate β-glucanase	GH16	Cellulose	-	-	-	1,4	1,8	-	-	-
127115	β-glucosidase, bgl2/cel1a	GH1	Cellulose	3,2	4,3	5,6	3,9	4,8	5,9	-	Yes
77989	β-glucosidase, cel1b	GH1	Cellulose	1,7	2,2	3	1,9	2,6	2,8	-	-
130860	Candidate α-glucosidase/oligo α-glucosidase	GH13	Cellulose	2	2,5	1,7	2,5	2,2	1,8	-	-
136547	β-glucosidase, bgl1/cel3a	GH3	Cellulose	1,3	3,6	6,1	1,9	3,8	7,1	-	Yes
122639	Candidate β-glucosidase, cel3d	GH3	Cellulose	1,3	1,6	4,9	3,4	4,3	5,7	-	-
125268	Candidate β-glucosidase, cel3c	GH3	Cellulose	3,9	3,4	4,1	2,5	3,4	4,3	-	-
74305	Candidate β-glucosidase, cel3e	GH3	Cellulose	1,7	2,5	2	1,5	1,6	1,5	Yes	-
25095	Candidate β-glucosidase, cel3b	GH3	Cellulose	1,6	1	1,2	-	-	1,1	Yes	Yes
104546	Candidate α-glucosidase	GH31	Cellulose	2,8	2,2	2,1	1,5	1,4	1,8	Yes	Yes
76309	Candidate α-glucosidase	GH31	Cellulose	2	1,8	1,6	2	2	1,3	-	-
122470	Cellobiohydrolase II, cbh2/cel6a	GH6	Cellulose	3,2	5,4	8,7	2,9	5,4	8,2	Yes	Yes
125125	Exocellobiohydrolase I, cbh1/cel7a	GH7	Cellulose	3,9	5,9	9,3	3,6	6	8,7	Yes	Yes
122518	Lytic polysaccharide monoxygenase, cel61b	AA9 (GH61)	Cellulose	-	8,2	10,5	3,4	6,7	9	Yes	-
121449	Cellulose induced protein, cip1	CBM1	Cellulose (?)	3,3	5,8	9,6	4	6,4	9,2	Yes	-
104220	Swollenin, swo1	CBM1	Cellulose (?)	-	2,1	6,2	3,2	3,6	6,7	Yes	-

(Continuação)

Protein ID	Annotation	CAZy family	Substrate	Log <sub>2</sub> (fold change)						SignalP	Secretome
				C6h	C12h	C24h	B6h	B12h	B24h		
67432	Candidate $\beta$ -mannosidase	GH2	Mannan	1,7	2,5	4,8	1,2	1,6	1,6	-	-
93487	Candidate $\beta$ -mannosidase	GH2	Mannan	2,6	3,1	3	-	1,9	-	-	-
24645	Candidate $\alpha$ -1,2-mannosidase	GH47	Mannan	-	-	-	1,2	1,3	1,5	Yes	-
91275	Candidate $\alpha$ -1,2-mannosidase	GH47	Mannan	1,1	1,1	-	-	-	-	Yes	-
83445	Candidate $\alpha$ -1,2-mannosidase	GH47	Mannan	1,3	1,1	1,6	1,1	1,1	2,1	-	-
101105	Candidate $\alpha$ -1,2-mannosidase	GH47	Mannan	1,3	1,1	1,5	1,2	1,3	1,3	-	-
122377	$\beta$ -mannanase, <i>man1</i>	GH5	Mannan	4,3	6,5	11,7	-	5	8,7	Yes	Yes
70100	Candidate $\alpha$ -1,6-mannanase	GH76	Mannan	1,4	2,3	2,1	-	1,5	-	Yes	-
126869	Candidate $\alpha$ -1,6-mannanase	GH76	Mannan	-	1,1	2	-	-	2,4	-	-
112243	Candidate $\alpha$ -1,6-mannanase	GH76	Mannan	-	3	1,4	1,5	2	1,5	Yes	-
6566	Candidate $\alpha$ -1,2-mannosidase	GH92	Mannan	1,7	1,1	-	-	-	-	Yes	-
25615	Candidate $\alpha$ -1,2-mannosidase	GH92	Mannan	1,4	1,6	2	1,3	1,5	1,6	Yes	-
71164	Candidate $\alpha$ -1,2-mannosidase	GH92	Mannan	-	1,4	1,5	-	1,4	-	Yes	-
111943	Xyloglucanase, <i>cel74a</i>	GH74	Xyloglucan	2,1	3	5,8	2,2	2,4	3,8	Yes	Yes
38400	Candidate $\alpha$ -L-fucosidase	GH95	Xyloglucan	-	1,9	-	-	2,1	-	Yes	-
128297	Candidate $\alpha$ -L-fucosidase	GH95	Xyloglucan	-	-	4,9	3	3,9	5,1	Yes	-
133728	Candidate $\alpha$ -L-fucosidase	GH95	Xyloglucan	1,1	-	1,1	-	-	-	-	-
23616	Endo- $\beta$ -1,4-xylanase, <i>xyn3</i>	GH10	Xylan/ Arabinoxylan	-	6,3	6,4	-	4,9	8,5	Yes	-
134945	Candidate endo- $\beta$ -1,4-xylanase, <i>xyn5</i>	GH11	Xylan/ Arabinoxylan	-	-	12	5,7	6,3	6,8	-	-
124931	Endo- $\beta$ -1,4-xylanase, <i>xyn2</i>	GH11	Xylan/ Arabinoxylan	-	4,1	8,5	6,5	7,7	8,9	Yes	Yes
38418	Endo- $\beta$ -1,4-xylanase, <i>xyn1</i>	GH11	Xylan/ Arabinoxylan	-	-	8	2,1	3,2	2,7	Yes	-
25542	Candidate xylan- $\alpha$ -1,2-glucuronidase	GH115	Xylan/ Arabinoxylan	1,4	1,6	1,4	1	1,3	-	-	-
139631	Acetylxylan esterase, <i>axe1</i>	CE5	Xylan/ Arabinoxylan	1,1	5,1	9,5	4,6	6,6	9,1	Yes	Yes
88887	Candidate acetyl xylan esterase	CE5	Xylan/ Arabinoxylan	-	2,7	7,9	5	6,2	8,9	-	-
140746	$\beta$ -xylosidase, <i>bxl1</i>	GH3	Xylan/ Arabinoxylan	2,4	4,6	7,9	8	7,8	8,2	Yes	Yes
90847	Endo- $\beta$ -1,4-xylanase, <i>xyn4</i>	GH30	Xylan/ Arabinoxylan	-	1,1	6,8	5,3	5,1	6	Yes	Yes
77521	Candidate $\beta$ -xylosidase/ $\alpha$ -L-arabinofuranosidase	GH43	Xylan/ Arabinoxylan	3,1	3,4	5,2	4,4	5	5,3	-	-

(Continuação)

Protein ID	Annotation	CAZy family	Substrate	Log <sub>2</sub> (fold change)						SignalP	Secretome
				C6h	C12h	C24h	B6h	B12h	B24h		
128705	Candidate acetyl xylan esterase	CE3	Xylan/ Arabinoxylan	1,8	1,8	2,3	1,8	2	1,5	-	-
5643	Candidate acetyl esterase	CE16	Xylan/ Arabinoxylan	2	2	2,1	1,4	1,7	1	Yes	-
102517	$\alpha$ -L-arabinofuranosidase I, <i>abf1</i>	GH54	Xylan/ Arabinoxylan	2,8	4,2	5,7	2,7	1,7	-	Yes	-
118070	Candidate $\alpha$ -L-arabinofuranosidase, <i>abf2</i>	GH62	Xylan/ Arabinoxylan	6,2	7,6	10,3	6	5,1	7	Yes	-
90302	$\alpha$ -glucuronidase, <i>glr1</i>	GH67	Xylan/ Arabinoxylan	4	5,3	7,9	7,1	7,5	7,6	Yes	-
125575	Glucuronoyl esterase, <i>cip2</i>	CE15	Hemicellulose/Lignin	3,4	7,1	10,5	4,9	6,5	9,1	Yes	-
12151	Candidate chondroitin lyase	PL8	Hemicellulose/Pectin	-	-	1,4	-	2,2	5,3	Yes	-
136770	Acetyl esterase, <i>aes1</i>	CE16	Hemicellulose: side chain	3,4	6,3	9,7	6,5	7,7	7,7	Yes	-
134448	Candidate $\alpha$ -xylosidase/ $\alpha$ -glucosidase	GH31	Hemicellulose: side chain	6,9	7,8	11,9	8	7,5	8,3	-	-
133383	Candidate endo-polygalacturonase	GH28	Pectin	6,9	7,8	11,5	-	-	7,3	Yes	-
91667	Candidate exo-polygalacturonase	GH28	Pectin	-	-	9,4	-	6,5	8,8	Yes	-
11580	Candidate endo- $\beta$ -1,6-galactanase	GH30	Pectin	5,1	5,8	7,3	5,4	4,3	4	Yes	-
69609	Candidate $\beta$ -glucuronidase	GH79	Pectin	-	-	-	-	-	5	Yes	Yes
6699	Candidate rhamnogalacturonyl hydrolase	GH105	Pectin	-	-	8,1	-	5,8	7,4	-	-
98078	Putative D-galacturonic acid reductase	GH109	Pectin	-	-	-	-	-	2,2	-	-
140234	GDSL Lipase/Acylhydrolase family protein	CE12	Pectin?	1,1	-	-	-	-	-	-	-
75754	Candidate $\alpha$ -amylase	GH13	1,4- $\alpha$ -glucan	1,3	1,5	1,4	1,1	1,1	-	Yes	-
124526	Chitinase, <i>chi46</i>	GH18	Chitin	-	-	-	-	-	1,7	Yes	Yes
7503	Candidate chitinase, <i>chi18-15</i>	GH18	Chitin	-	-	-	-	2,9	3,5	Yes	-
79622	Candidate chitinase, <i>chi18-9</i>	GH18	Chitin	1,4	1,2	-	1,1	-	-	Yes	-
114979	Candidate chitinase, <i>chi18-3</i>	GH18	Chitin	-	1,1	-	-	-	-	Yes	-
142123	Candidate chitinase, <i>chi18-18</i>	GH18	Chitin	1,7	1,3	2,3	1	-	2,3	Yes	-
142298	Endo-N-acetyl- $\beta$ -D-glucosaminidase, <i>Endo T</i>	GH18	Chitin	2,1	2,1	2	2,3	2,3	2,7	Yes	-
24455	Exo- $\beta$ -D-glucosaminidase, <i>gls93</i>	GH2	Chitin	-	-	-	-	1,8	3,4	Yes	-
13025	Candidate chitosanase	GH75	Chitin	1,4	2,1	2	2,2	-	-	Yes	-
93070	Bifunctional catalase-peroxidase, <i>cat2</i>	AA1/AA2	Lignin	-	1,1	-	2	1,8	2	Yes	-
68144	Ferroxidase-like protein	AA1_2	Lignin	-	-	-	-	-	5,9	Yes	-
36679	FAD binding domain protein	AA7	Lignin	-	1,6	-	1,8	-	-	Yes	-



(Continuação)

Protein ID	Annotation	CAZy family	Substrate	Log <sub>2</sub> (fold change)						SignalP	Secretome
				C6h	C12h	C24h	B6h	B12h	B24h		
104503	Alcohol oxidase	AA3_2	Lignin	-	1,3	-	-	-	-	-	-
114133	GMC (glucose-methanol-choline) oxidoreductase	AA3_2	Lignin	4,3	4,1	5,5	3	-	3,8	-	-
7186	Alcohol oxidase	AA3	Lignin	2,2	2,6	3,5	-	2,6	2,1	-	-
77360	Alcohol oxidase	AA3_2	Lignin	2,3	2,1	2,4	1,8	1,6	1,5	-	-
93273	GMC oxidoreductase	AA3_2	Lignin	2	-	2,2	-	-	-	-	-
115811	Alcohol oxidase	AA3_2	Lignin	1,6	1,5	2,2	1,2	-	1,5	-	-
115868	Alcohol dehydrogenase	AA3_3	Lignin	3,4	3,2	2	4,4	4	4,6	-	-
114246	Alcohol oxidase	AA3_2	Lignin	-	-	1,2	-	-	-	-	-
7224	Putative trehalose phosphatase	GT20	Carbohydrate transport and metabolism	-	1,6	-	-	-	-	-	-
32557	Candidate $\alpha$ -galactosidase	GH27	Carbohydrate transport and metabolism	1	-	-	-	-	-	Yes	-
71638	Candidate $\alpha$ -galactosidase	GH27	Carbohydrate transport and metabolism	-	1,6	4	2	2,8	3,4	Yes	-
39277	$\alpha$ -galactosidase, agl3	GH27	Carbohydrate transport and metabolism	3	2,2	2,1	1,2	1,4	-	Yes	-
75420	Alpha-galactosidase 3	GH27	Carbohydrate transport and metabolism	2,9	2	1,9	1,7	2	1,4	Yes	Yes
6433	$\alpha$ -galactosidase, agl1	GH27	Carbohydrate transport and metabolism	1,2	1,9	1,3	-	-	-	Yes	-
93498	Candidate endo- $\beta$ -1,4-xylanase	GH30	Carbohydrate transport and metabolism	2,1	4,5	8,2	2,5	4,4	7,2	Yes	Yes
101346	$\beta$ -galactosidase, bga1	GH35	Carbohydrate transport and metabolism	2,8	3,3	4,5	3,2	1,8	1,2	Yes	-
12566	$\alpha$ -galactosidase, agl2	GH36	Carbohydrate transport and metabolism	2,3	2,7	2,4	1,9	1,8	1	Yes	-
102372	Candidate $\alpha,\alpha$ -trehalase	GH37	Carbohydrate transport and metabolism	1,7	2,4	2,1	-	1,1	-	Yes	Yes
83083	Candidate raffinose synthase domain protein	GH36	Carbohydrate transport and metabolism	1,4	1,2	1,3	1,1	-	-	-	-
39997	Candidate endo- $\beta$ -1,4-glucuronan lyase, glaA	PL20	Cell wall biosynthesis and morphogenesis	-	-	-	-	-	4	Yes	-

(Continuação)

Protein ID	Annotation	CAZy family	Substrate	Log <sub>2</sub> (fold change)						SignalP	Secretome
				C6h	C12h	C24h	B6h	B12h	B24h		
82054	Glycosyltransferase	GT32	Cell wall biosynthesis and morphogenesis	1,7	1,2	-	-	-	-	-	-
114628	Chitin synthase	GT2	Cell wall biosynthesis and morphogenesis	1	-	-	-	-	-	-	Yes
124397	Glycosyltransferase	GT32	Cell wall biosynthesis and morphogenesis	2,8	3	3,8	4,6	4,4	5,5	-	-
9074	Candidate $\beta$ -1,3-glucanosyltransferase	GH72	Cell wall biosynthesis and morphogenesis	1,2	-	-	1,3	-	1,9	Yes	Yes
98936	Candidate $\beta$ -1,3-glucanosyltransferase	GH72	Cell wall biosynthesis and morphogenesis	-	1,3	-	-	-	-	Yes	Yes
82061	Glycosyltransferase	GT2	Cell wall biosynthesis and morphogenesis	-	-	1,2	-	1,3	1,4	-	-
85693	Candidate $\alpha$ -glycosyltransferase	GT32	Cell wall biosynthesis and morphogenesis	1,1	-	1,1	1,9	1,4	2,1	-	-
100857	$\alpha$ / $\beta$ -hydrolase	CE10	Non-carbohydrate substrates	1	-	-	-	-	-	-	-
70716	$\alpha$ / $\beta$ -hydrolase	CE10	Non-carbohydrate substrates	5	5,8	5,5	-	-	-	-	-
129558	Putative triacylglycerol lipase	CE10	Non-carbohydrate substrates	2,3	2,9	4,5	2,1	2,6	2	Yes	-
77741	$\alpha$ / $\beta$ -hydrolase	CE10	Non-carbohydrate substrates	1,2	1,4	1,7	-	1,1	-	Yes	-
127881	$\alpha$ / $\beta$ -hydrolase	CE10	Non-carbohydrate substrates	2,1	1,7	1,5	-	-	-	Yes	-
8573	Carboxylesterase	CE10	Non-carbohydrate substrates	2,4	2	1,3	2,3	2,2	1,3	-	-
35023	Phospholipase/carboxylesterase	CE10	Non-carbohydrate substrates	-	1,2	1,3	-	1,1	-	-	-
133341	$\alpha$ / $\beta$ -hydrolase	CE10	Non-carbohydrate substrates	-	1,3	1,1	1,3	1,4	-	-	-
68899	Hypothetical protein	CE10	Non-carbohydrate substrates	1,2	1,1	1,1	-	-	-	-	-
125146	Hydrophobin 3, hfb3	-	-	-	-	-	-	-	1,2	Yes	-

(Continuação)

Protein ID	Annotation	CAZy family	Substrate	Log <sub>2</sub> (fold change)						SignalP	Secretome
				C6h	C12h	C24h	B6h	B12h	B24h		
77278	<i>α /β-hydrolase</i>	CE5	-	-	-	-	-	-	1,2	Yes	-
127981	<i>NAD-binding Rossmann fold oxidoreductase</i>	GH109	-	2,3	3,8	7,9	5	5,8	7,7	-	-
12995	<i>DUF1237 domain protein</i>	GH125	-	3,1	3,1	2,9	1,8	2	1,4	Yes	-
70331	<i>Hypothetical protein</i>	GH132	-	1,4	-	-	1,4	-	1,6	Yes	-
113024	<i>Putative triglyceride lipase-cholesterol esterase</i>	CE1	-	1,4	1,4	-	-	-	-	-	-
91131	<i>FAD binding domain protein</i>	GH16	-	2	1,3	-	-	-	-	-	-
141702	<i>Glycoside hydrolase</i>	GH17	-	1,1	-	-	-	-	-	-	-
26117	<i>UDP-Glycosyltransferase/glycogen phosphorylase</i>	GT1	-	-	-	2,6	-	-	-	-	-
130436	<i>Candidate S-formylglutathione hydrolase</i>	CE1	-	-	1,2	1,2	-	-	-	-	-
86661	<i>α-1,3-mannosyltransferase CMT1</i>	GT69	-	-	1,1	1,2	-	-	1,1	-	-

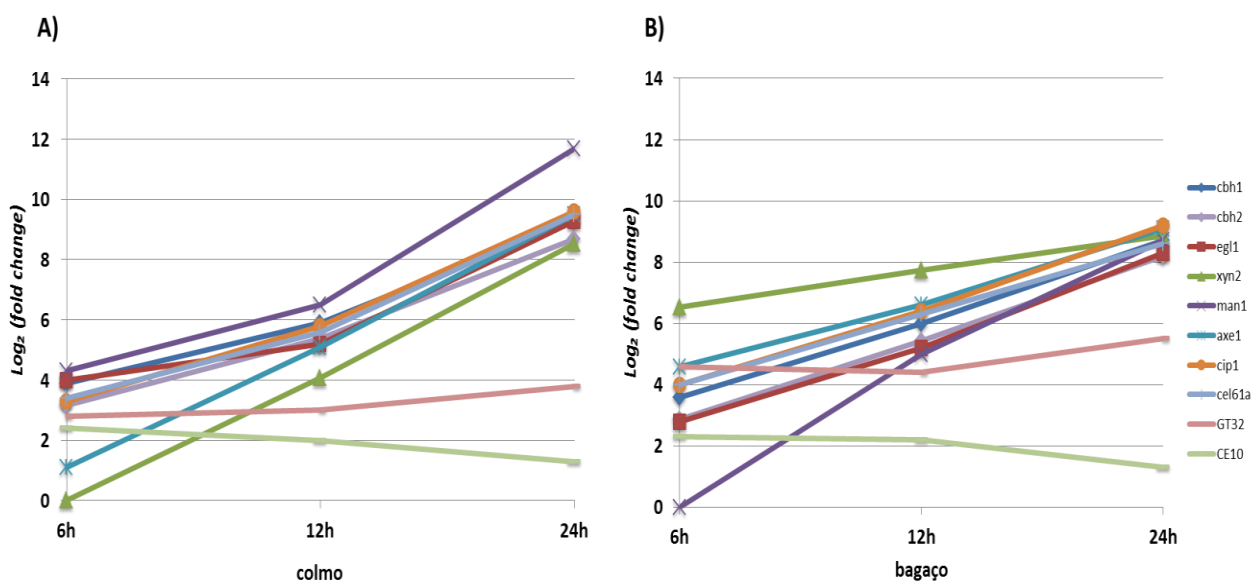
Em bagaço, foram encontradas exclusivamente 8 GHs (uma manosidase, pectinases, quitinases e duas glucanases hipotéticas da família GH16), uma proteína AA1, uma PL20 e uma CE5. Curiosamente, foi encontrado também o gene de uma hidrofobina (*hfb3*, jgi|125146) sendo hiper-expresso somente na condição bagaço 24 horas. Hidrofobinas são proteínas capazes de adsorver em superfícies hidrofóbicas e mediar a interação do fungo com o meio por meio de camadas anfipáticas (Linder *et al.*, 2005; Ohtaki *et al.*, 2006). Dessa forma, pode-se sugerir que a hiper-expressão desse gene em bagaço pode estar ajudando o fungo a reconhecer e/ou se ligar ao substrato. Estas diferenças podem ser explicadas pela composição diferente de polissacarídeos e dos açúcares liberados da degradação da parede do colmo e do bagaço. Apesar deste último substrato lignocelulósico ser derivado do colmo, ele foi explodido a vapor e teve sua composição alterada, contendo uma menor proporção de hemicelulose e pectina (Prof<sup>o</sup>. Dr<sup>o</sup>. Marcus Buckeridge, comunicação pessoal).

Contudo, nem todos os genes de CAZymes que tiveram seus transcritos detectados no RNA-Seq estão relacionados com a desconstrução da parede vegetal. Como já citado, a classe das glicosil transferases (GT) atua no remodelamento da parede celular e as esterases da família 10 (CE10) não possuem atividade contra carboidratos (Häkkinen *et al.*, 2012). Esses genes não foram hiper-expressos em várias condições e, de maneira geral, sua expressão foi menor em relação aos genes de celulasas e hemicelulasas (Tabela 3, Figura 9).

Em comparação com o secretoma (Anexo I), 23 genes hiper-expressos de CAZymes e proteínas acessórias também tiveram suas proteínas identificadas no sobrenadante (18% do total) (Tabela 3 e Anexo I). Essa baixa correlação de proteínas identificadas por proteômica e transcriptômica também foi vista em diversos trabalhos (Glaesener *et al.*, 2008; Ghazalpour *et al.*, 2011; Müller *et al.*, 2012), o que demonstra que essa relação é bastante complexa e que os níveis das proteínas são fortemente influenciados pelas modificações pós-traducionais e pela variação da estabilidade delas (Müller *et al.*, 2012). Além disso, outros fatores podem ter contribuído para esse resultado, como por exemplo: 1) a proteína produzida não é secretada para o meio extracelular; 2) problemas de *fold*ing das proteínas que foram produzidas, mas que não puderam ser secretadas, 3) ou ainda por causa da baixa

secreção da proteína no sobrenadante que não foi detectada por razões experimentais ou do espectrômetro de massas.

Embora *Trichoderma* seja estudado como um produtor de enzimas hidrolíticas, ele se destaca por sua capacidade de produzir celulases. Em cepas utilizadas industrialmente, 60% das proteínas secretadas referem-se à celobiohidrolase CBH1 (jgi|125125) e 20% à celobiohidrolase CBH2 (jgi|122470) (Schuster & Schmoll, 2010). No presente trabalho, os genes referentes a estas enzimas estão entre os mais hiper-expressos e com maiores valores de  $\log_2(\text{fold change})$  em relação à condição frutose (Tabela 3, Anexo II), sendo hiper-expressos em todas as condições analisadas. A expressão dos genes dessas duas celulases foi crescente ao longo do tempo e parece ser mais induzida em colmo (Tabela 3, Figura 9).



**Figura 9.** Hiper-expressão de genes de celulases (*cbh1*, *cbh2* e *egl1*), hemicelulases (*xyn2*, *man1* e *axe1*), proteínas acessórias (*cip1* e *cel61a*), glicosil transferase (GT32, jgi|124397) e carboxilesterase (CE10, jgi|8573) no transcriptoma de *T. reesei* crescido em colmo (A) e bagaço (B) por 6, 12 e 24 horas (h). A expressão é mostrada em  $\log_2(\text{fold change})$ .

No ponto 24 horas, por exemplo, foi detectado um *fold change* de 630 vezes para CBH1 em colmo ( $\log_2(\text{fold change})$  igual a 9,3) e 416 vezes em bagaço ( $\log_2(\text{fold change})$  igual a 8,7) (Tabela 3, Figura 9). Já CBH2 foi detectada com um *fold change* de 416 e 294 vezes em colmo e bagaço 24 horas, com valores de  $\log_2(\text{fold change})$  iguais a 8,7 e 8,2, respectivamente (Tabela 3). A hiper-expressão dessas duas celulases também foi verificada no secretoma de *T.*

*reesei* pela abundância de peptídeos (Tabela 3, Anexo I). CBH1 foi a enzima com maior número de peptídeos, e junto com CBH2 representaram 80% dos peptídeos de celulases (Anexo I), como descrito em trabalhos anteriores (Herpoël-Gimbert *et al.*, 2008; Adav *et al.*, 2012).

Em relação às hemicelulases, os principais genes que codificam às endoxilanases,  $\beta$ -xilosidases, arabinofuranosidases, galactosidases foram encontrados sendo hiper-expressos no transcriptoma de *T. reesei* (Tabela 3). Apesar da hemicelulose ter uma variedade muito maior de composição, estrutura e ramificação em relação a celulose (Souza *et al.*, 2012), de acordo com Martinez e colaboradores (2008), *T. reesei* possui apenas 16 genes que codificam para hemicelulases, a menor quantidade de genes entre os fungos *A. nidulans*, *A. oryzae*, *M. grisea*, *N. crassa* e *F. graminearum*. Dentre as hemicelulases já caracterizadas, foi possível identificar como sendo hiper-expressas quatro endo-xilanases (XYNI-IV), uma  $\beta$ -mananase (MAN1), duas acetil esterases (AXE1 e AES1), uma  $\alpha$ -glucoronidase (GLR1), uma arabinofuranosidase (ABF1), três  $\alpha$ -galactosidases (AGL1-3), uma  $\beta$ -xilosidase (BXL1), e uma glucuronoil esterase (CIP2) envolvida na quebra de ligações do tipo éster entre a hemicelulose e a lignina da parede vegetal (Häkkinen *et al.*, 2012) (Tabela 3). Além dessas hemicelulases, Häkkinen e colaboradores (2012) reanotaram os genes de CAZymes de *T. reesei* e encontraram vários deles sendo candidatos de hemicelulases, contudo eles precisam ser caracterizados para confirmação de sua atividade específica. Destes, analisando os genes hiper-expressos, foi possível encontrar duas endo-xilanases (93498 e XYNV, 134945), três  $\alpha$ -mananases (jgi|112243, 70100 e 126869) e uma arabinofuranosidase (ABF2) com atividade hipotética de hemicelulase, o que sugere que estas proteínas podem estar atuando na biomassa lignocelulósica. A maioria dessas hemicelulases caracterizadas ou não estava presente em quase todas as condições analisadas, e dentre os genes mais expressos em colmo e bagaço estavam *cip2*, *axe1*, *xyn2*, *man1*, *xyn3*, *aes1* e *abf2* (Tabela 3). Curiosamente, estes e outros genes hiper-expressos no transcriptoma de *T. reesei* crescido em colmo e bagaço também foram identificados como sendo altamente expressos após 24 horas de crescimento em palha de trigo (Ries *et al.*, 2013).

Vale ressaltar que a ação conjunta de hemicelulases com celulases é necessária para a completa degradação dos polissacarídeos encontrados na parede celular vegetal (Ilmen *et al.*, 1997; Foreman *et al.*, 2003). Assim, devido à alta expressão desses genes em colmo e bagaço (Tabela 3), pode-se afirmar que está havendo sinergismo das atividades celulo e hemicelulolíticas contra esses dois substratos, o que fica evidente com o aumento da expressão das celulases junto com o das hemicelulases ao longo do tempo (Figura 9, Tabela 3). Adicionalmente, diversos genes de proteínas acessórias apresentaram expressão considerável no transcriptoma de *T. reesei* desde 6 horas pós inóculo, como, por exemplo, *cip1*, *swo1* e os genes das famílias AA3 e AA9.

O gene da proteína CIP1 (jgi|121449) foi extremamente induzido nas condições colmo e bagaço 24 horas, sendo o gene mais expresso na condição bagaço, com valor de log do *fold change* igual a 9,2 (*fold change* igual a 588 vezes) (Tabela 3, Anexo II). Ries e colaboradores (2013) também observaram o gene *cip1* sendo induzido quando *T. reesei* foi crescido em palha de trigo, mas sua expressão nesse substrato (323 em RPKM) foi bem inferior à observada em colmo e bagaço 24 horas (1386 e 1038, respectivamente). Estudando a regulação de CAZymes de *T. reesei* em meios com lactose e soforose, Foreman e colaboradores (2003) observaram uma hiper-expressão de *cip1* na presença de soforose, e que o perfil transcricional desse gene era similar ao de endoglucanases e da celobiohidrolase CBHI. Em razão disso, esses genes foram agrupados no mesmo *cluster* de regulação gênica, sugerindo o potencial de CIP1 na degradação de celulose. Dessa forma, CIP1 pode estar auxiliando as enzimas hidrolíticas na desconstrução da parte celulolítica da parede dos substratos.

O gene *swo1* (jgi|104220) também mostrou uma hiper-expressão crescente tanto em colmo quanto bagaço, alcançando um *fold change* maior que 70 vezes no ponto 24 horas (Tabela 3). As swoleninas têm um papel no afrouxamento da parede vegetal, uma vez que quebram as pontes de hidrogênio entre os polissacarídeos, desorganizando as fibras de celulose e aumentando a entrada de água na matriz (McQueen-Mason & Cosgrove, 1994; Verma *et al.*, 2013). Em *T. asperellum*, foi observado que uma swolenina aumentou significativamente a capacidade de colonização do fungo na raiz de planta (Brotman *et al.*,

2008). Essa proteína pode estar aumentando então o ataque das celulases por facilitar a exposição das fibras de celulose presentes na parede. Além disso, recentemente Andberg e colaboradores (2015) observaram atividade de SWO1 em ligações glicosídicas do tipo  $\beta$ -1,4, um modo de ação característico tanto de celobiohidrolases quanto endoglucanases. Assim, SWO1 parece participar do sistema celulolítico de *T. reesei* ao lado de enzimas hidrolíticas e, possivelmente, facilitando a interação com o substrato recalcitrante.

Segundo uma revisão feita por Levasseur e colaboradores (2013), o fungo *T. reesei* possui 13 genes pertencentes à família AA e outros distantemente relacionados baseado na similaridade de sequências entre as proteínas, totalizando 27 genes. Essas famílias são compostas por ferroxidases, lacases, oxidases e reductases envolvidas na degradação de lignina (Levasseur *et al.*, 2013), e no transcriptoma de *T. reesei* foram identificados 10 genes hiper-expressos dessas proteínas. Até recentemente, acreditava-se que as proteínas da família AA9 (também chamadas de LPMO, *copper-dependent lytic polysaccharide monoxygenases*) eram celulases da família GH61, porque foi observada uma fraca atividade de endoglucanase para a proteína CEL61a codificada por *egl4/cel61a* de *T. reesei* (Saloheimo *et al.*, 1997; Karlsson *et al.*, 2001; Morgenstern *et al.*, 2014). Entretanto, viu-se que essa atividade era muito menor que aquela medida para outras endoglucanases e que era improvável que um fungo com um arsenal celulolítico tão eficiente dependesse de uma endoglucanase com atividade catalítica fraca (Morgenstern *et al.*, 2014). Assim, por aumentarem a degradação de celulose junto com as celulases clássicas através de um mecanismo oxidativo que cliva a celulose usando agentes redutores, como ascorbato, glutatona e até mesmo lignina (Quinlan *et al.*, 2011; Beeson *et al.*, 2012; Levasseur *et al.*, 2013), as LPMOs foram classificadas dentro de uma família denominada AA9.

Outros genes das famílias com atividade auxiliar também foram hiper-expressos em várias condições, mas sua expressão foi menor em relação aos genes das LPMOs (AA9). Curiosamente, 3 genes hiper-expressos (jgi|93070, 7186 e 36679) que não foram anotados pelo trabalho de Levasseur e colaboradores (2013), tiveram suas proteínas preditas pelo banco de dados dbCAN (<http://csbl.bmb.uga.edu/dbCAN/>) como tendo atividade auxiliar das famílias AA2, AA3 e AA7, respectivamente. Assim, esses genes são bons candidatos para



estudos mais detalhados sobre as proteínas que eles codificam e se elas atuam sobre a biomassa vegetal, sendo inclusive de interesse biotecnológico na suplementação de coquetéis enzimáticos.

Além do log do *fold change* em relação à condição não induzida, pode-se avaliar os dados do RNA-Seq em função do RPKM, o que reflete a porcentagem dos transcritos detectados. Em colmo 24 horas, é possível notar que do total de RPKM, um terço (33%) (21.860 de 66.292, em valor de RPKM) pertence à hiper-expressão de genes de CAZymes e proteínas acessórias. Usando palha de trigo, Ries e colaboradores (2013) encontraram uma proporção bem menor de CAZymes e proteínas acessórias no ponto 24 horas (13,76% em valor total de RPKM para *T. reesei* QM6a). Além disso, já se nota a hiper-expressão de vários genes de CAZymes e proteínas acessórias no ponto 6 horas de crescimento em bagaço ou em colmo, com genes de glicosil hidrolases expressos com um *fold change* de até 120 vezes em colmo (GH31, jgi|134448) e 256 vezes em bagaço (*bx1*, jgi|140746) em relação à frutose (Tabela 3). Assim, foi possível observar que *T. reesei* é capaz de responder à fonte recalcitrante presente no meio já nas primeiras horas de crescimento, e que ao longo do tempo essa resposta tende a aumentar a expressão das CAZymes e proteínas acessórias envolvidas na desconstrução de biomassa vegetal (Tabela 3, Figura 9).

A proporção de CAZymes e proteínas acessórias (excluindo-se as glicosil transferases e esterases da família CE10) muda sutilmente entre os substratos em termos de RPKM: enquanto em bagaço a proporção de transcritos é de 6, 10 e 28,4% do total de RPKM para os tempos 6, 12 e 24 horas, respectivamente, a proporção em colmo é de 4,6, 8,5 e 33% (dados não mostrados). Foi possível observar também que uma pequena quantidade de genes corresponde a mais da metade do RPKM total dessa classe de proteínas. Em colmo 24 horas, por exemplo, os transcritos de apenas 3 famílias (GH6, GH7 e CBM1) de CAZymes são os responsáveis por 56,8% (12.439 de 21.860 em RPKM) do RPKM total das CAZymes e proteínas acessórias. Curiosamente, essa proporção é ainda maior para bagaço 24 horas onde estas mesmas 3 famílias (GH6, GH7 e CBM1) e uma hidrofobina sem classificação no banco de dados CAZy compõem 65% (17.591 de 26.809 em RPKM) do RPKM total das CAZymes e proteínas acessórias. Dentre esses genes, estão as celobiohidrolases 1 e 2 (CBH1 e

CBH2), uma endoglucanase já caracterizada (EGI) e outra hipotética (CEL1B), e as proteínas acessórias CIP1, SWO1 e HFB3. Notavelmente, o gene *hfb3* representa quase 37% dos transcritos totais (9.844 de 26.809 em RPKM) de CAZymes e proteínas acessórias da condição bagaço 24 horas. Apesar da grande quantidade de transcritos, *hfb3* teve um baixo valor de  $\log_2(\text{fold change})$  (1,2) (Tabela 3), o que indica uma grande quantidade de transcritos também na condição frutose. Hidrofobinas são proteínas com regiões hidrofílicas e hidrofóbicas que interagem e cobrem superfícies, diminuindo a tensão superficial e ajudando o fungo a sobreviver e se adaptar ao ambiente no qual está crescendo (Linder *et al.*, 2005). Em trabalhos anteriores que analisaram o perfil transcricional de *T. reesei* crescido em palha de trigo (Ries *et al.*, 2013) e avicel celulose (Chen *et al.*, 2014), também foram detectados níveis altos dos transcritos dessa proteína (182 e 407, respectivamente, em termos de RPKM) e de outra hidrofobina (*hfb2*, jgi|119989), o que mostra a importância dessas proteínas no crescimento do fungo em substratos celulósicos.

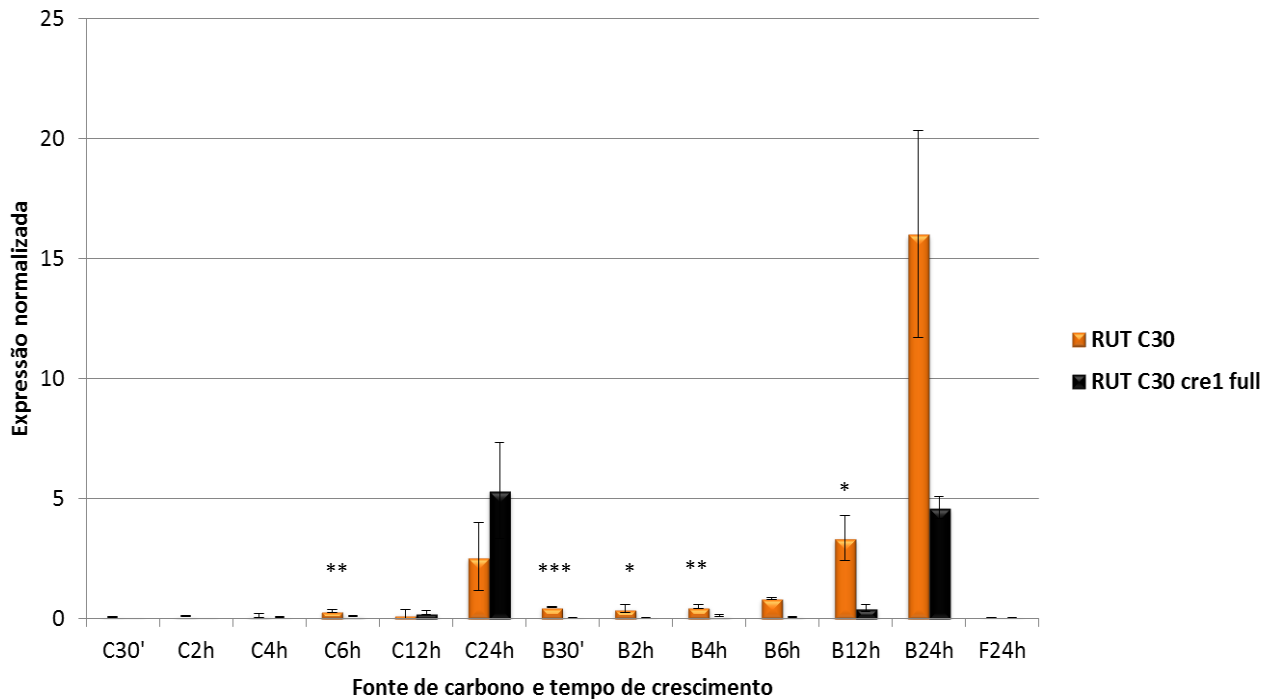
Essa resposta fisiológica do fungo em hiper-expressar uma grande diversidade de CAZymes e proteínas acessórias em várias condições (Tabela 3) e o perfil transcricional crescente dessas enzimas ao longo do tempo (Figura 9) são apoiados por um modelo de indução de enzimas em fungos (Carle-Urioste *et al.*, 1997; Tian *et al.*, 2009). O modelo propõe uma expressão basal de hidrolases que iniciam a degradação de polissacarídeos complexos e liberam componentes indutores que desencadeiam a completa transcrição dos genes ligados a desconstrução daqueles açúcares. Analisando o crescimento de *A. niger* em palha de trigo, Delmas e colaboradores (2012) puderam observar que havia uma expressão temporal de hidrolases e proteínas acessórias, sendo que algumas delas já começavam a ser expressas nos tempos mais iniciais (3 e 6 horas) e outras nos tardios (12 e 24 horas). Com os resultados obtidos, Delmas e colaboradores sugeriram que nas primeiras horas de exposição ao substrato lignocelulósico, não há nenhum açúcar solúvel disponível que possa ser utilizado pelo fungo, e que portanto a repressão por carbono realizada por CreA (homólogo do CRE1 em *T. reesei*) estaria diminuída. Nessa condição de *starvation*, algumas enzimas seriam então ativadas, quebrariam a parede vegetal e liberariam açúcares solúveis. Nas horas subsequentes, os reguladores transcricionais, dentre eles principalmente o XlnR (homólogo

do XYR1 em *T. reesei*), seriam induzidos pelos açúcares que foram assimilados e ativariam a expressão de outras enzimas lignolulíticas, aumentando a resposta hidrolítica do fungo contra a biomassa vegetal. Apesar desse modelo de indução ter sido proposto para *A. niger* crescido em outro tipo de substrato lignocelulósico, pode-se aplicar a mesma explicação para a resposta do *T. reesei* crescido em colmo e bagaço de cana. As enzimas também foram expressas temporalmente e de forma crescente (Anexo II), e açúcares solúveis foram liberados como resultado da degradação da parede (Anexo I). Embora colmo e bagaço induzam de forma diferente a expressão desses genes, o que é notado não somente pela diferença nas enzimas hiper-expressas, como também no nível em que são transcritas, o compartilhamento da maioria das CAZymes e proteínas acessórias entre os dois substratos revela que a estratégia usada por *T. reesei* para degradar a parede vegetal do colmo e do bagaço é similar.

A expressão dos genes de algumas CAZymes e proteínas acessórias juntamente com outros genes (Anexo III) foi validada por qPCR, e, como esperado, o resultado encontrado confirmou os dados de RNA-Seq (Tabela 3, Anexo III). E a fim de avaliar a repressão catabólica por carbono mediada por CRE1 na expressão de genes supostamente envolvidos com a desconstrução da parede do bagaço e colmo da cana, foi feita a validação da expressão gênica também com a cepa RUT C30 com o *cre1* íntegro. De maneira geral, viu-se que a expressão das celulasas (CBH1, CBH2 e CEL3c), hemicelulasas (BXL1, XYN1 e XYN2) e das proteínas acessórias (SWO1 e CIP1) foi maior nos tempos mais tardios (12 e 24 horas), tanto em colmo quanto no bagaço (Figura 10, Anexo III). Estes mesmos genes apresentaram uma expressão desprezível na condição não induzida, frutose 24 horas, mostrando que os dois substratos lignocelulósicos induziram a expressão dessas CAZymes e proteínas acessórias. Em relação à repressão catabólica, foi visto que alguns genes tiveram sua expressão reprimida por CRE1 em algumas condições (*cel3c* em bagaço 24 horas, por exemplo), mas outros mostraram uma maior expressão na cepa com o repressor íntegro (*swo1* em bagaço 12 horas e *xyn1* em colmo 24 horas, por exemplo) (Anexo III). Isso levanta a hipótese de que existem outros elementos regulatórios ainda não conhecidos capazes de superar a repressão catabólica conferida por *cre1* e induzir a expressão gênica. Além disso,

outras modificações encontradas no genoma da cepa RUT C30 podem ter auxiliado no aumento de seu poder celulolítico (Le crom *et al.*, 2009). Foi observada ainda uma alta expressão de *hfb3* (jgi|125146) em várias condições de crescimento em relação aos normalizadores *act* e *sar1* (vide item 4.5), mas a diferença de expressão entre os substratos lignocelulósicos e a frutose foi pequena (Anexo III), como também verificado nos dados de RNA-Seq (Tabela 3).

A Figura 10 abaixo mostra a expressão do gene *cbh2* (jgi|122470) em ambas as cepas de *T. reesei* RUT C30. Essa celobiohidrolase foi uma das celulasas mais expressas em colmo e bagaço (Tabela 3) e mostrou ter um perfil de expressão crescente ao longo do tempo, como também verificado por qPCR. Contudo, nos dados de RNA-Seq, sua expressão foi maior em colmo do que em bagaço, o contrário do observado por qPCR (Figura 10). Principalmente nos pontos mais iniciais do colmo (30', 2h e 4h) e na condição frutose 24 horas, sua expressão foi desprezível, mostrando que *cbh2* foi induzida somente nos tempos tardios do colmo. Em bagaço, *cbh2* atingiu sua maior expressão também nos pontos mais tardios, mas curiosamente já começou a ser expressa após 30 minutos, o que mostra uma resposta rápida do fungo quando exposto à este substrato lignocelulósico, possivelmente devido a uma liberação maior e/ou mais rápida dos açúcares indutores. Em relação à repressão por *cre1*, é notável a menor expressão de *cbh2* com o *cre1* íntegro em bagaço 12 e 24 horas. Isso indica que o gene está sendo reprimido mesmo na presença de um substrato lignocelulósico que induz a transcrição de hidrolases, e que é portanto regulado por *cre1*.



**Figura 10.** Expressão gênica da celobiohidrolase 2 (CBH2, jgi|122470) nas linhagens RUT C30 com o repressor *cre1* truncado e íntegro (*full*). A expressão foi normalizada pela expressão dos dois genes normalizadores *act* (actina) e *sar1* (GTPase). As amostras foram coletadas após 30 min ('), 2, 4, 6, 12 e 24 horas em colmo (C) bagaço (B). Como condição controle, foi usado o crescimento em frutose após 24 horas. As barras de erro indicam o desvio padrão para as réplicas biológicas. As réplicas foram analisadas pelo *t test* não pareado (\*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$ ).

### 5.1.2 Perfil transcricional de transportadores hipotéticos de açúcar

Em número de genes, a proporção de transportadores hipotéticos de açúcar foi similar a das CAZymes e proteínas de função desconhecida, variando de 4 a 7% entre as condições de crescimento (Figura 8). Como já mencionado anteriormente, foram considerados apenas os transportadores hipotéticos de açúcar ou os que não estavam anotados para outros tipos de solutos. Assim, foram identificados 83 genes de transportadores hipotéticos da família MFS (*Major Facilitator Superfamily*) sendo hiper-expressos no transcriptoma de *T. reesei* (Tabela 4). Esse número representa 40% dos transportadores MFS (207 genes) e 11% de todos os transportadores (816) encontrados no genoma de *T. reesei* (Wang *et al.*, 2015). Mais da metade (58 genes) foi hiper-expressa em ambos os substratos, 16 somente em colmo e 9

em bagaço (Tabela 4). A maioria deles não foi funcionalmente caracterizada e nem seus homólogos em outros fungos filamentosos, o que mostra a importância de serem validados funcionalmente.

Esta superfamília MFS é a maior família de transportadores secundários conhecida e está presente no genoma de bactérias até humanos. Os transportadores MFS transportam seletivamente uma variedade de substratos através da membrana, dentre eles mono e oligossacarídeos, aminoácidos, peptídeos e proteínas, vitaminas, cofatores de enzimas, drogas etc (Reddy *et al.*, 2012), e são divididos em 76 subfamílias de acordo com o banco de dados TCDB (*Transporter Classification Database*) (<http://www.tcdb.org/>). Segundo análises de bioinformática, a maioria dos transportadores MFS tem 12 domínios transmembrânicos (TMHs) (Yan, 2013), embora já tenham sido preditas proteínas com um número maior ou menor de domínios (Reddy *et al.*, 2012). O número de segmentos transmembrânicos também foi predito pelo programa TMHMM (vide item 4.4) a partir da sequência de aminoácidos para confirmar a função de transportadores, e a grande maioria apresentou 12 TMHs (Tabela 4).

**Tabela 4.** Transportadores de função conhecida e hipotéticos hiper-expressos no transcriptoma de *T. reesei* crescido em bagaço (B) e colmo (C), após 6, 12 e 24 horas (h). A identificação (ID) de cada proteína é a disponível no banco de dados do JGI do *T. reesei* RUT C30 v1.0. A expressão é apresentada como o  $\log_2$  do *fold change* e a presença dos segmentos transmembrânicos (TMH) também é mostrada. Em negrito, são mostrados dois genes com afinidade por xilose (ver discussão abaixo).

ID JGI	Protein	$\log_2(\text{fold change})$						TMH
		C6h	C12h	C24h	B6h	B12h	B24h	
<b>33630</b>	<b>Putative xylose permease</b>	<b>7,7</b>	<b>8,6</b>	<b>10</b>	<b>8,2</b>	<b>6,4</b>	<b>5</b>	<b>10</b>
91594	Putative maltose permease	3,6	4,8	9,1	4,3	3,9	5,5	12
109243	Cellulose response transporter, crt1	4,7	6	7,4	5,6	6,7	7,5	12
127980	MFS lactose permease	2	2,8	6,2	3,9	4,3	5,8	12
97259	General substrate transporter	1,4	2,1	6	3,7	5,1	6,9	12
138519	General substrate transporter	1,5	2,1	6	6,4	6,4	6,1	11
137795	Putative MFS transporter	2,8	3,4	5,3	2,8	3,3	5,3	12
69330	MFS general substrate transporter	1,1	2,2	5,3	1,3	3	4,3	12
25063	Putative MFS transporter	3,8	3,7	5,3	3,6	3,2	3,8	12
90719	Transmembrane transporter	3,6	4,8	4,6	5,1	5,5	5,9	11
77945	Putative MFS transporter	5,1	2,9	4,1	3,8	5,4	4,2	12
79592	General substrate transporter	1,4	1,7	3,6	1,9	2,1	2	10
38765	General substrate transporter	3,3	3	3,5	1,1	3,5	4,3	12
91127	Putative MFS transporter	5	3,1	3,3	2,5	2,5	1,5	10
140464	Major facilitator superfamily protein	2,8	2,6	3	2,6	2,7	1,3	12
128552	Maltose permease	1,9	2,2	2,9	2,8	3,5	3,9	10
136711	Major facilitator superfamily transporter	2	3	2,9	2,6	4	3,5	10
85254	General substrate transporter	2,2	2,3	2,6	1,9	2,1	1,7	10
84089	MSF superfamily transporter	2,9	2,5	2,4	1,8	1,7	1,1	10
<b>136712</b>	<b>D-xylose proton symporter</b>	<b>4,2</b>	<b>3,5</b>	<b>2,3</b>	<b>2,5</b>	<b>2,7</b>	<b>1,5</b>	<b>12</b>
11734	General substrate transporter	4,8	2,1	2,1	1,8	1,8	1,2	12
26008	Major facilitator superfamily transporter	1,8	1,7	2	1,5	1,7	1,8	10
7811	General substrate transporter	1,2	1,4	1,9	1,2	2	3,6	12
12083	General substrate transporter	2,4	2	1,9	1,6	1,8	1,4	11
121436	MFS general substrate transporter	1,9	1,5	1,7	1,8	2	1,9	10
94445	MFS transporter-like protein	2,7	1,9	1,7	1,7	2,5	1,6	12
35773	MFS general substrate transporter	1,7	1,7	1,5	1,7	1,4	1,6	12
13268	General substrate transporter	1,4	1,4	1,3	1,1	1	1	12
84399	Major facilitator superfamily transporter	2,2	1,8	1,2	1,3	2,5	2	12
26932	Putative MFS monosaccharide transporter	4,7	5,7	5,2	5,9	3,6	-	10
13305	MFS general substrate transporter	-	1	2	1,6	2	1,5	14
73108	MFS general substrate transporter	-	1,5	1,8	1,2	1,9	1,3	12

**(Continuação)**

ID JGI	Protein	Log <sub>2</sub> (fold change)						TMH
		C6h	C12h	C24h	B6h	B12h	B24h	
34717	<i>MFS transporter</i>	1,9	2,3	1,7	1,8	3,3	-	12
52104	<i>MFS general substrate transporter</i>	2,4	1,7	1,5	1,3	1,2	-	10
106261	<i>MFS general substrate transporter</i>	1,8	1,2	1,4	1,5	1,5	-	12
67815	<i>MFS general substrate transporter</i>	-	1	1,4	1,5	1,5	1,5	13
87338	<i>Fucose permease</i>	1,3	1,5	1,2	1,2	1,1	-	15
24422	<i>MFS general substrate transporter</i>	2,1	1,1	1	1	2,2	-	10
73367	<i>MFS general substrate transporter</i>	-	5,3	8,1	-	5,9	6,4	12
132743	<i>MFS general substrate transporter</i>	2	2,9	3	-	1,4	-	10
137001	<i>Putative MFS hexose transporter</i>	-	1	1,6	-	1,3	1,4	10
106322	<i>General substrate transporter</i>	-	1,1	1,6	-	1,1	1,1	12
70522	<i>MSF superfamily transporter</i>	1,6	1,4	1,5	-	1,1	-	9
23545	<i>Major facilitator superfamily transporter</i>	-	1,1	1,3	-	1,3	1,3	12
88479	<i>MFS transporter</i>	1,3	-	1,2	1,7	1	-	11
109034	<i>MFS general substrate transporter</i>	1,1	1	-	1,1	1,1	-	14
102186	<i>Predicted transporter</i>	-	1,4	-	1,6	1,2	1,4	10
84853	<i>MFS general substrate transporter</i>	-	-	1,9	-	1,2	1,2	13
89291	<i>Putative maltose permease</i>	1,9	1,6	1,6	-	-	-	12
137229	<i>MFS general substrate transporter</i>	-	1,1	1,4	-	1,1	-	10
7623	<i>Hexose transporter-like protein</i>	1,7	1,6	1,3	-	-	-	12
65065	<i>MFS general substrate transporter, partial</i>	1,8	1,1	1,2	-	-	-	9
86755	<i>MFS general substrate transporter</i>	-	-	1,2	1,1	1,3	-	12
26434	<i>Putative MFS monosaccharide transporter</i>	-	-	1,1	1,1	-	1,1	11
124396	<i>Glucose transporter</i>	2	2,8	-	1,1	-	-	8
97622	<i>MFS general substrate transporter</i>	1,5	-	-	1,6	1,8	-	10
79984	<i>Hexose transport-related protein</i>	-	1,5	-	1,1	1,3	-	12
116350	<i>MFS general substrate transporter</i>	-	-	-	1,4	1,1	1,8	12
93149	<i>MFS general substrate transporter</i>	-	-	1,3	-	-	1,2	11
135589	<i>Maltose permease</i>	-	-	1,2	-	-	1,3	12
74036	<i>MFS general substrate transporter</i>	-	1,5	1,2	-	-	-	12
39450	<i>MFS general substrate transporter</i>	-	-	1,1	-	1	-	12
134530	<i>MFS general substrate transporter</i>	-	-	1,1	1,2	-	-	12
12422	<i>Putative MFS transporter</i>	1,2	1	-	-	-	-	12
115870	<i>MFS general substrate transporter</i>	-	-	-	1,6	1,2	-	13
85306	<i>MFS general substrate transporter</i>	-	-	3,7	-	-	-	11
97401	<i>Sugar transporter</i>	-	-	1,2	-	-	-	12
76604	<i>MFS general substrate transporter</i>	-	-	1,2	-	-	-	12
93519	<i>MFS general substrate transporter</i>	-	-	1,1	-	-	-	10
87089	<i>MFS general substrate transporter</i>	-	-	1	-	-	-	11
136988	<i>Sugar transporter 1, stp1</i>	1,1	-	-	-	-	-	12
90843	<i>Major facilitator superfamily transporter</i>	1,7	-	-	-	-	-	12



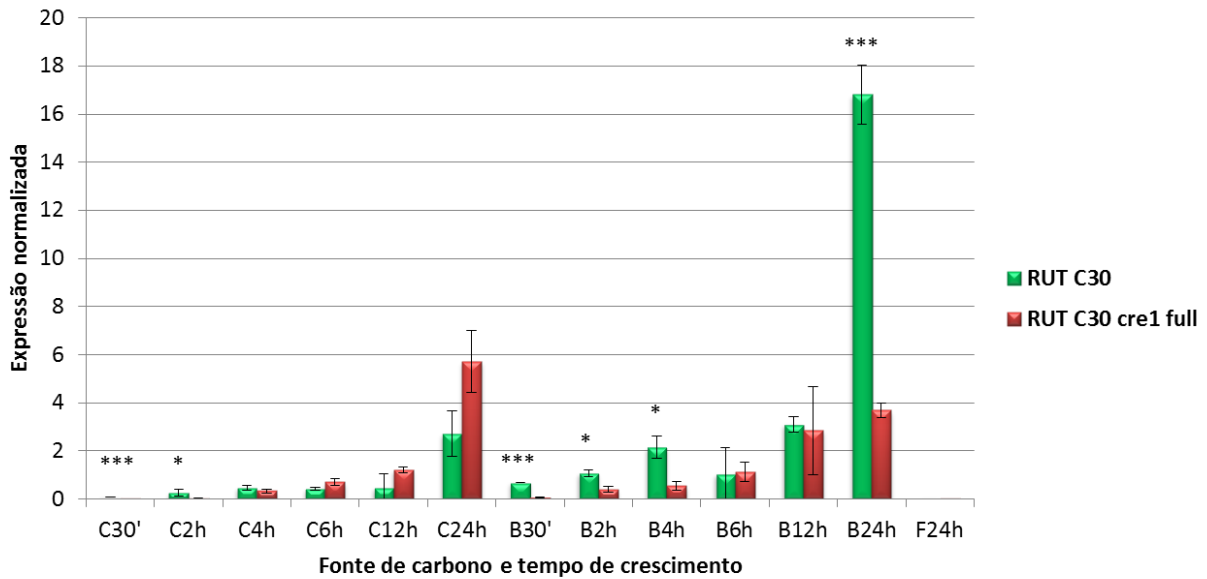
(Continuação)

ID JGI	Protein	Log <sub>2</sub> (fold change)						TMH
		C6h	C12h	C24h	B6h	B12h	B24h	
68089	General substrate transporter	1,3	-	-	-	-	-	11
55710	MFS general substrate transporter	1,3	-	-	-	-	-	10
35634	MFS general substrate transporter	-	1,1	-	-	-	-	11
88679	Putative MFS transporter	-	2,8	-	-	-	-	7
24994	MFS general substrate transporter	-	-	-	-	-	1	11
77618	Putative sugar transporter	-	-	-	-	2,5	-	12
77979	MFS transporter	-	-	-	-	1,2	-	11
78278	MFS general substrate transporter, partial	-	-	-	-	1,2	-	9
87685	Sugar transporter	-	-	-	-	1,2	-	10
70312	Predicted hexose transporter (strong similarity with <i>Beauveria bassiana</i> )	-	-	-	-	1,1	-	9
11899	MFS general substrate transporter	-	-	-	-	1	-	12

Entre os genes mais hiper-expressos em colmo estão uma permease hipotética de xilose (jgi|33630) e uma permease hipotética de maltose (jgi|91594). Esses genes foram hiper-expressos em todas as condições e foram mais induzidos em colmo do que em bagaço (Tabela 4). Curiosamente, a permease hipotética de xilose mostrou um perfil de expressão crescente em colmo e decrescente em bagaço ao longo do tempo (Tabela 4). Por ser um transportador hipotético de xilose, talvez essa diferença de perfil esteja relacionada com o influxo e consumo desse açúcar na célula. Além dessa permease hipotética de xilose, outro transportador de xilose (*D-xylose proton symporter*, jgi|136712) foi hiper-expresso em todas as condições analisadas (Tabela 4), mas a quantidade de transcritos detectados foi bem menor (dados não mostrados). Ries e colaboradores (2013) também observaram a permease hipotética de xilose (jgi|33630) sendo hiper-expressa em palha de trigo (324 em palha de trigo 24 horas contra 0 em glicose 48 horas, em valores de RPKM), mas não notaram o mesmo para o outro transportador (jgi|136712). A hiper-expressão desses dois transportadores hipotéticos de xilose é interessante do ponto de vista biotecnológico, uma vez que seus genes podem ser validados funcionalmente e engenheirados em leveduras industriais para assimilar xilose no processo de produção de etanol 2G.

Em bagaço, os dois genes mais expressos foram do transportador hipotético (jgi|97259) e Crt1 (jgi|109243). Ambos os genes foram expressos em todas as condições de

crescimento e apresentaram um perfil transcricional crescente ao longo do tempo (Tabela 4). A análise do BLASTp do segundo transportador mais hiper-expresso (jgi|97259) revelou 65% de identidade com um transportador de alta afinidade por glicose do fungo *Acremonium chrysogenum*. Assim, pode-se sugerir que o gene deste transportador pode codificar para um transportador de hexoses. Interessantemente, Zhang e colaboradores (2013) mostraram que a deleção e expressão constitutiva de *crt1* levou à drástica redução e ao aumento da expressão de celulases, respectivamente, e que a assimilação de celobiose e sofrorse foi a mesma no mutante  $\Delta crt1$  em relação ao selvagem. Por essas e outras observações, os autores sugerem que Crt1 possa transmitir um sinal de indução para a expressão das enzimas lignocelulolíticas independente do transporte de açúcares para o interior da célula (Zhang *et al.*, 2013). Assim, são necessários outros estudos para poder distinguir, com certeza, se Crt1 atua como sinalizador e/ou transportador de açúcares indutores. Independente disso, pode-se sugerir que Crt1 esteja atuando positivamente sobre a expressão das celulases, uma vez que o aumento da expressão dessas enzimas foi acompanhado do aumento de *crt1*. Curiosamente, *crt1* também foi o gene mais expresso após 24 horas de crescimento de *T. reesei* em palha de trigo (Ries *et al.*, 2013) e em celulose e sofrorse (dos Santos Castro *et al.*, 2014), o que demonstra a importância de sua expressão em vários substratos diferentes (complexos e simples). A hiper-expressão de *crt1* em colmo e bagaço também foi confirmada por qPCR (Figura 11) e novamente pode-se notar maior expressão desse gene nos pontos mais tardios (12 e 24 horas) de crescimento.



**Figura 11.** Expressão gênica do transportador *Crt1* (*jgi*|109243) nas linhagens RUT C30 com o repressor *cre1* truncado e íntegro (*full*). A expressão foi normalizada pela expressão dos dois genes normalizadores *act* (actina) e *sar1* (GTPase). As amostras foram coletadas após 30 min ('), 2, 4, 6, 12 e 24 horas em colmo (C) e bagaço (B). Como condição controle, foi usado o crescimento em frutose após 24 horas. As barras de erro indicam o desvio padrão para as réplicas biológicas. As réplicas foram analisadas pelo *t test* não pareado (\*  $p < 0,05$ ; \*\*  $p < 0,01$ ;  $p^* < 0,001$ )

Comparando as cepas de *T. reesei* RUT C30 com o *cre1* truncado e íntegro (*full*), pode-se observar uma diferença grande em bagaço 24h, com a expressão de *crt1* em *cre1* truncado sendo quatro vezes maior que em *cre1 full*, o que indica que *Crt1* é regulado negativamente por CRE1, pelo menos em bagaço. O mesmo tempo em colmo apresentou uma expressão de *crt1* menor na cepa com *cre1* truncado, e com nível de transcritos sendo maior em *cre1 full*. A maior expressão desse transportador em bagaço na linhagem RUT C30 com o gene *cre1* truncado apóia a hiper-expressão de várias celulases (celobiohidrolases, endoglucanases e  $\beta$ -glicosidases) nesse substrato (Tabela 3), e a indução de genes de transportadores na presença de colmo e bagaço (Tabela 4) indica que as frações de celulose e hemicelulose da parede vegetal estão sendo degradadas e liberando açúcares (Anexo I) que podem ser transportados para o interior do micélio para geração de energia e indução do sistema lignocelulolítico do fungo *T. reesei*.

### **5.1.3 Perfil transcricional de proteínas de função desconhecida**

Como citado anteriormente, foram classificadas nessa categoria proteínas que não tinham função conhecida ou que tinham domínios ainda não caracterizados (como por exemplo, os *DUF domain*) quando buscadas contra os bancos de dados citados anteriormente (vide item 4.4). Os genes hiper-expressos dessas proteínas representaram de 19 a 22% do número total de genes hiper-expressos no transcriptoma de *T. reesei* (Figura 8), sendo o colmo o substrato com maior número de genes hiper-expressos. No total, foram identificados mais de 200 e 150 genes hiper-expressos em colmo e bagaço, respectivamente, e, para simplificar, são mostradas as proteínas com peptídeo sinal predito e apenas as 15 mais expressas em colmo 24 horas sem o peptídeo sinal (Anexo IV). De maneira geral, a expressão dos genes que codificam a estas proteínas foi inferior à das enzimas lignocelulolíticas, proteínas acessórias e dos transportadores (Tabela 3 e 4), mas muitas delas apresentaram elevados valores de *fold change*. Diversos deles estavam presentes em todas as condições analisadas, em especial dois que codificam à proteínas com peptídeo sinal predito (jgi|131184 e 5086) e outras duas sem (jgi|4983 e 42889). Além de muitas proteínas terem o peptídeo sinal predito e serem candidatas a novas enzimas secretadas e com ação hidrolítica na parede vegetal, várias delas apresentaram um ou mais domínios transmembrânicos (TMHs), o que pode estar relacionado com a ancoragem dessas proteínas em membranas biológicas (Lodish *et al.*, 2000). Assim, existe uma grande oportunidade de se descobrir novos genes importantes para o sistema lignocelulolítico de *T. reesei* dentro dessa categoria de proteínas sem função conhecida.

### **5.1.4 Perfil transcricional de fatores de transcrição hipotéticos**

Os fatores de transcrição são proteínas que regulam os níveis de expressão de seus genes alvos por ligação a sequências específicas no DNA, localizadas nos promotores, aumentando ou reprimindo sua transcrição. Normalmente, eles são pouco expressos na célula, mas mudanças sutis nos níveis de transcritos podem alterar significativamente o metabolismo de um organismo (Caldana *et al.*, 2007). Dessa forma, a identificação e caracterização funcional dos fatores de transcrição são essenciais para a reconstrução e

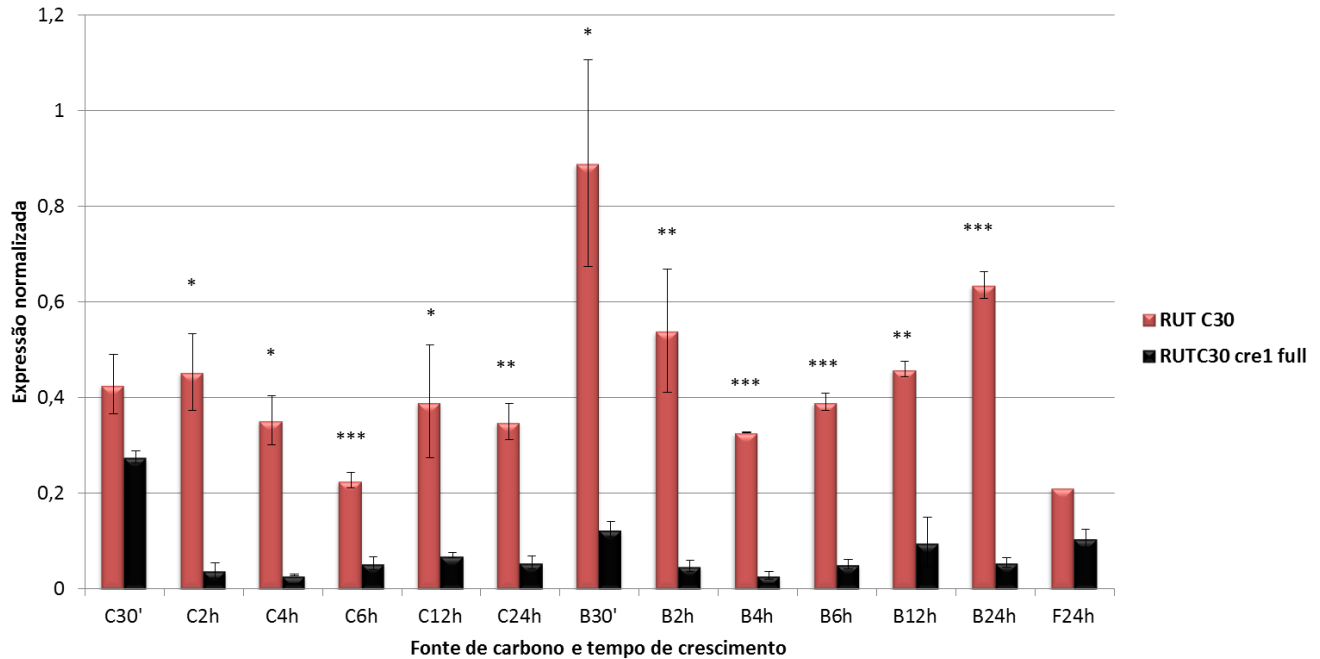
entendimento acerca do funcionamento das redes gênicas biológicas, já que eles regulam os principais processos celulares e de desenvolvimento.

Foram identificados 94 fatores de transcrição hipotéticos sendo hiper-expressos no transcriptoma de *T. reesei* (Anexo V). Juntos eles representam a menor porcentagem do número total de genes hiper-expressos (de 4 a 6%) (Figura 8). Desses 94 fatores de transcrição hipotéticos, 49 foram compartilhados entre os dois substratos, 33 foram encontrados exclusivamente em colmo e 12 somente em bagaço, e dos 49 compartilhados entre colmo e bagaço, 16 foram hiper-expressos em todas as condições. Em relação aos domínios dos fatores de transcrição hipotéticos, foram encontrados principalmente domínios *zinc finger* do tipo C2H2 (Anexo V). Esse é o tipo mais comum de fator de transcrição em eucariotos e também foi encontrado em abundância no transcriptoma de *T. reesei* crescido em celulose e soforose (dos Santos Castro *et al.*, 2014). Além do tipo C2H2, também foram encontrados domínios *zinc finger* do tipo AN1, C3H e RING. Fatores de transcrição hipotéticos com outros domínios também foram hiper-expressos, como os domínios *homeobox*, *bZIP* (*basic-leucine zipper*) e *NmrA-like* (Anexo X).

Apesar de já serem conhecidos vários reguladores positivos (XYR1, ACE2, ACE3, LAE1, VEL1, complexo HAP2/3/5) e dois negativos (CRE1 e ACE1) da expressão de CAZymes, só foi possível identificar os ativadores *xyr1* (jgi|98788) e *ace3* (jgi|98455) sendo hiper-expressos após 12 e 24 horas (Anexo V). Como já citado anteriormente, XYR1 é o principal ativador de genes de hidrolases e sua deleção leva à perda da indução de celulases, mesmo na presença dos substratos indutores celulose e soforose (dos Santos Castro *et al.*, 2014). ACE3 foi recentemente descoberto e a deleção de seu gene reduziu a atividade xilanolítica e a expressão de celulases e hemicelulases em *T. reesei* (Häkkinen *et al.*, 2014). Portanto, a hiper-expressão de *xyr1* junto com *ace3*, muito provavelmente, deve estar induzindo a expressão das principais celulases e hemicelulases em colmo e bagaço de cana. Além disso, o perfil transcricional crescente ao longo do tempo (Anexo V) reforça o modelo de indução discutido anteriormente, que sugere a liberação de açúcares solúveis nas primeiras horas da desconstrução da parede vegetal, transcrição de *xlnR* (homólogo de *xyr1*) e ativação das enzimas lignocelulolíticas.

Analisando o perfil transcricional de fatores de transcrição de *T. reesei* crescido em soforose, glicose e celulose, dos Santos Castro e colaboradores (2014) agruparam os fatores de acordo com a expressão deles em cada um dos substratos. Especificamente em celulose, três genes hiper-expressos foram compartilhados com o presente trabalho: um fator *zinc finger* do tipo Zn2Cys6 (jgi|128051), um fator bZIP (jgi|10981) e outro *zinc finger* do tipo C2H2 (jgi|95791) (Anexo V) (dos Santos Castro *et al.*, 2014). Os dois primeiros foram hiper-expressos unicamente em colmo, e o último teve seus transcritos detectados em todas as condições (Anexo V). Assim, por terem sido regulados por celulose, pode-se sugerir que esses fatores de transcrição tenham algum papel na regulação da resposta do fungo à parede celular vegetal.

Apesar de não ter sido encontrado o gene da proteína metiltransferase LAE1 (jgi|9778) no transcriptoma de *T. reesei* segundo os parâmetros definidos, seu gene foi investigado por qPCR nas linhagens RUT C30 com o repressor *cre1* truncado e íntegro (Figura 12), uma vez que foi mostrada a importância da sua expressão na regulação positiva das enzimas lignocelulolíticas em linhagens com o gene deletado ou sendo super-expresso (Seiboth *et al.*, 2012). Uma das razões para seu gene não ter sido encontrado nos dados de RNA-Seq é o fato dele ser pouco expresso nas condições de crescimento utilizadas neste trabalho.



**Figura 12.** Expressão gênica da metiltransferase LAE1 (jgi|9778) nas linhagens RUT C30 com o repressor *cre1* truncado e íntegro (*full*). A expressão foi normalizada pela expressão dos dois genes normalizadores *act* (actina) e *sar1* (GTPase). As amostras foram coletadas após 30 min ('), 2, 4, 6, 12 e 24 horas em colmo (C) e bagaço (B). Como condição controle, foi usado o crescimento em frutose após 24 horas. As barras de erro indicam o desvio padrão para as réplicas biológicas. As réplicas foram analisadas pelo *t test* não pareado (\*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$ )

Na Figura 12, é notável a diferença de expressão entre as linhagens de *T. reesei* em todas as condições analisadas. Na linhagem com o *cre1* truncado, LAE1 é bem mais expresso do que na linhagem reprimida catabolicamente (*cre1 full*). Na condição bagaço 30 minutos, por exemplo, a expressão da metiltransferase é maior que 0,8 na linhagem com *cre1* truncado, enquanto que com o gene íntegro, esse valor não chega a 0,2, ou seja, quatro vezes maior. Aparentemente, *lae1* também foi mais expresso em colmo e bagaço em relação à condição controle (frutose 24 horas). O resultado encontrado sugere que esse gene pode ser regulado pelo repressor catabólico CRE1 e que os maiores níveis de expressão em colmo e bagaço podem estar ativando a expressão do sistema celulolítico de *T. reesei*, refletindo na hiper-expressão de celulasas, hemicelulasas e proteínas acessórias (Tabela 3, Anexo II). Os fatores identificados aqui como hiper-expressos e não caracterizados poderão ser melhor

estudados, usando mutantes com seus genes deletados, do mesmo modo que proteínas de função desconhecida com algum potencial na sacarificação da biomassa.

Além de LAE1, a expressão dos ativadores transcricionais XYR1 e ACE2 também foi investigada por qPCR junto com a dos repressores CRE1 e ACE1 e de um fator de transcrição hipotético (jgi|26638) (Anexo III). Esses reguladores são fatores de transcrição e, portanto, apresentam normalmente uma baixa expressão na célula (Caldana *et al.*, 2007), o que justifica não terem sido encontrados nos dados de RNA-Seq segundo os parâmetros definidos (vide item 4.4). Como já esperado, todos apresentaram uma menor expressão em relação aos genes de enzimas hidrolíticas, com exceção de *cre1*, que curiosamente apresentou uma alta expressão. Notavelmente, foi observada uma expressão decrescente do fator de transcrição hipotético em bagaço para a cepa RUT C30 com *cre1* truncado (Anexo III), sendo maior nos pontos iniciais de crescimento nesse substrato. Em todos os tempos amostrados do bagaço, a expressão foi maior que em frutose, mostrando que esse fator hipotético é induzido por bagaço e é um bom candidato para estudos posteriores. Além disso, foi possível observar que a expressão de *xyr1* e *ace2* diminuiu na cepa com *cre1* íntegro, o que confirma o papel de CRE1 como repressor desses dois ativadores (Anexo III).

## **5.2 Análise do secretoma após sacarificação dos substratos**

A fim de analisar as proteínas secretadas pelo fungo para degradar os substratos e a atividade enzimática foram feitos ensaios enzimáticos e digestão das proteínas do sobrenadante em gel SDS-PAGE e posterior processamento dos peptídeos no espectrômetro de massas (LTQ Velos Orbitrap). Foi realizada ainda a análise dos açúcares liberados da desconstrução da parede por cromatografia de troca iônica de alta resolução (HPAEC-PAD). Todos os resultados encontrados nesses experimentos são apresentados e discutidos no manuscrito que segue no Anexo I. Esse manuscrito já foi submetido para a revista ***PLoS ONE*** e apresenta os resultados obtidos para o *T. reesei*, bem como para o fungo *Aspergillus niger*. Este fungo filamentoso também apresenta um grande potencial biotecnológico por ser um grande produtor e secretor de enzimas hidrolíticas, e tem se tornado um modelo útil para maior compreensão de como é feita a regulação do sistema celulolítico dos fungos (de Souza



*et al.*, 2011). Dessa forma, a comparação do secretoma dos dois fungos hipercelulolíticos é mais interessante e poderá contribuir para maiores *insights* a cerca da degradação de materiais lignocelulósicos.



## 6. CONCLUSÃO

Foi possível identificar no transcriptoma de *T. reesei* as principais celulases, hemicelulases e proteínas acessórias relacionadas direta ou indiretamente com a desconstrução da parede vegetal. De modo geral, as celulases e hemicelulases apresentaram uma expressão maior do que outras enzimas, e o nível dos seus transcritos foi crescente ao longo do tempo tanto em colmo quanto no bagaço. A grande maioria dos genes de CAZymes e proteínas acessórias hiper-expressos foram compartilhados pelos dois substratos, o que demonstra que a estratégia usada por *T. reesei* para degradar a parede do colmo e do bagaço é similar. Também foi observada a hiper-expressão de 83 transportadores hipotéticos de açúcar, 94 fatores de transcrição hipotéticos, 3 proteínas com atividade auxiliar predita mas não anotada no genoma e vários genes de função desconhecida. Esses genes são bons candidatos para maiores estudos que possam demonstrar alguma função relacionada com a degradação da biomassa. A análise da expressão gênica por qPCR validou os dados de RNA-Seq, e também permitiu supor que alguns genes estão sujeitos à repressão catabólica mediada por CRE1. Em paralelo, foi possível identificar várias CAZymes sendo secretadas no sobrenadante do fungo em todas as condições amostradas. Mesmo no ponto 6 horas, várias enzimas hidrolíticas estavam presentes no colmo e no bagaço. Ensaio enzimáticos também foram realizados com xilano, arabinano, CMC, xiloglucano e  $\beta$ -glicano, mostrando atividades celulo e hemicelulolíticas contra os substratos. A análise dos açúcares solúveis no sobrenadante após exposição e ação das enzimas hidrolíticas também indicou que tanto a celulose quanto hemicelulose estão sendo quebradas logo após 6 horas de crescimento nesses substratos. Assim, descrevemos pela primeira vez o arsenal de enzimas transcritas e secretadas por *T. reesei* RUT C30, desde pontos iniciais de crescimento, em bagaço explodido e colmo de cana-de-açúcar. Por fim, este trabalho também permitiu a identificação de vários genes, com função predita ou não, que podem abrir caminho para a descoberta de novos atuantes na resposta do fungo ao substrato lignocelulósico.



## 7. REFERÊNCIAS BIBLIOGRÁFICAS

- Adav, S.S., Chao, L.T., Sze, S.K. (2012) Quantitative secretomic analysis of *Trichoderma reesei* strains reveals enzymatic composition for lignocellulosic biomass degradation. *Mol Cell Proteomics*, 11:1-15.
- Adsul, M.G., Singhvi, M.S., Gaikawai, S.A., Gokhale, D.V. (2011) Development of biocatalysts for production of commodity chemicals from lignocellulosic biomass. *Bioresour Technol*, 102: 4304-4312.
- Agger, J., Vikso-Nielsen, A., Meyer, A.S. (2010) Enzymatic xylose release from pretreated corn bran arabinoxylan: differential effects of deacetylation and deferuloylation on insoluble and soluble substrate fractions. *J Agric Food Chem*, 58: 6141–6148.
- Akel, E., Metz, B., Seiboth, B., Kubicek, C. P. (2009) Molecular regulation of arabinan and L-arabinose metabolism in *Hypocrea jecorina* (*Trichoderma reesei*). *Eukaryot Cell*, 8:1837–1844.
- Amore, A., Giacobbe, S., Faraco, V. (2013) Regulation of cellulase and hemicellulase gene expression in fungi. *Curr Genomics*, 14(4):230-49.
- Amorim, H.V., Lopes, M.L., de Castro Oliveira, J.V., Buckeridge, M.S., Goldman, G.H. (2011) Scientific challenges of bioethanol production in Brazil. *Appl Microbiol Biotechnol*, 91(5): 1267-75.
- Andberg, M., Penttilä, M., Saloheimo, M. (2015) Swollenin from *Trichoderma reesei* exhibits hydrolytic activity against cellulosic substrates with features of both endoglucanases and cellobiohydrolases. *Bioresour Technol*, 181:105-113.
- Anders, S., Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol*, 11(10):R106.
- Andreoli, C., Souza, P.S. (2007) Cana-de-açúcar: a melhor alternativa para conversão de energia solar e fóssil em etanol. *Economia & Energia* 59:27-33.
- Antizar-Ladislao, B., Turrion-Gomez, J.L. (2008) Second-generation biofuels and local bioenergy systems. *Biofuel Bioprod Bior*, 2(5):455-469.
- Arantes, V., Saddler, J.N. (2010) Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol Biofuels*, 3:4.
- Aro, N., Saloheimo, A., Ilmén, M., Pentilla, M. (2001) ACEII, a novel transcriptional activator involved in regulation of cellulase and xylanase genes of *Trichoderma reesei*. *J Biol Chem*, 276: 24309-24314.
- Aro, N., Ilmén, M., Saloheimo, A., Pentilla, M. (2003) ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. *Appl Environ Microbiol*, 69: 56-65.

Axtell, M.J. (2014) Butter: High-precision genomic alignment of small RNA-seq data. *New Results*, doi: <http://dx.doi.org/10.1101/007427>.

Balat, M. (2011) Production of bioethanol from lignocellulosic materials via the biochemical pathway: a review. *Energ Convers Manage*, 52: 858-875.

Banerjee, G., Scott-Craig, J. S. and Walton, J. D. (2010a) Improving enzymes for biomass conversion: a basic research perspective. *Bioenergy Research* 3: 82-92.

Banerjee, S., Mudliar, S., Sen, R., Giri, B., Satpute, D., Chakrabarti, T., Pandey, R.A. (2010b) Commercializing lignocellulosic bioethanol: technology bottlenecks and possible remedies. *Biofuel Bioprod Bior*, 4: 77-93.

Basso, T.P., Basso, T.O., Gallo, C.R., Basso, L.C. (2013) Chapter 14 – Towards the Production of Second Generation Ethanol from Sugarcane Bagasse in Brazil. In *Biomass Now – Cultivation and Utilization*. Edited by Miodrag Darko Matovic.

Beeson, W.T., Phillips, C.M., Cate, J.H.D., Marletta, M.A. (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *J Am Chem Soc*, 134:890-892.

Biotium. Product Information. (2015) GelRed™ Nucleic Acid Gel Stain, 10,000x. Biotium – Glowing Products for Science™, CA, United States.

Bothast, R.J., Schlicher, M.A. (2005) Biotechnological processes for conversion of corn into ethanol. *Appl Microbiol Biotechnol*, 67: 19-25.

Bourne, Y., Henrissat, B. (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr Opin Struc Biol*, 11: 593-600.

Brotman, Y., Briff, E., Viterbo, A., Chet, I. (2008) Role of swollenin, an expansin-like protein from *Trichoderma*, in plant root colonization. *Plant Physiol*, 147:779–789

Caldana, C., Scheible, W.R., Mueller-Roeber, B., Ruzicic, S. (2007) A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. *Plant Methods*, 3:7.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): na expert resource for Glycogenomics. *Nucleic Acids Res*, 37: D233-238.

Cardona, C.A., Quintero, J.A., Paz, I.C. (2010) Production of bioethanol from sugarcane bagasse: status and perspectives. *Bioresource Technol*, 101::4754-4766.

Carle-Urioste, J.C., Escobar-Vera, J., El-Gogary, S., Henrique-Silva, F., Torigoi, E., Crivellaro, O., Herrera-Estrella, A., El-Dorry, H. (1997) Cellulase induction in *Trichoderma reesei* by cellulose requires its own basal expression. *J Biol Chem*, 272: 10169–10174.

Carpita, N.C., Gibeaut, D.M. (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J*, 3(1):1-30.

Chen, X.A., Ishida, N., Todaka, N., Nakamura, R., Maruyama, J., Takahashi, H., Kitamoto, K. (2010) Promotion of eficiente saccharification of crystalline clelulose by *Aspergillus fumigatus* Swo1. *Appl Environ Microbiol*, 76(8):2556-61.

Chen, X., Luo, Y., Yu, H., Sun, Y., Wu, H., Song, S., Hu, S., Dong, Z. (2014) Transcriptional profiling of biomass degradation-related genes during *Trichoderma reesei* growth on different carbon sources. *J Biotechnol*, 173: 59-64.

Clark, J.H., Luque, R., Matharu, A.S. (2012) Green Chemistry, Biofuels, and Biorefinery. *Annu Rev Chem Biomol Eng*, 3:183-207.

Conab (2014) Acompanhamento da safra brasileira: cana-de-açúcar, primeiro levantamento, abril/2014. Companhia Nacional de Abastecimento – Brasília. ([http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14\\_04\\_15\\_15\\_44\\_37\\_boletim\\_cana\\_portugues\\_-\\_1o\\_lev\\_-\\_14.pdf](http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14_04_15_15_44_37_boletim_cana_portugues_-_1o_lev_-_14.pdf))

Cortez, L., Magalhães, P. (1992) Principais subprodutos da agroindústria canavieira e sua valorização. *Revista Brasileira de Energia*, 2: 1-17.

Coughlan MP, Tuohy MG, Filho EXF, Puls J, Claeysens M, Vrsanská M, Hughes M (1993) Enzymological aspects of microbial hemicellulases with emphasis on fungal systems. In: Coughlan MP (ed) Hemicellulose and hemicellulases. Portland, London, pp 53–84

Criscuolo, A., Brisse, S. (2013) AlienTrimmer: a tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. *Genomics*, 102(5-6):500-6.

Cubero, B., Scazzocchio, C. (1994) Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans*. *EMBO J*, 13:407-415.

Cziferszky, A., Mach, R.L., Kubicek, C.P. (2002) Phosphorylation positively regulates DNA binding of the carbon catabolite repressor Cre1 of *Hypocrea jecorina* (*Trichoderma reesei*). *J Biol Chem*, 277:14688-14699.

Dashtban, M., Schraft, H., Qin, W. (2009) Fungal bioconversion of lignocellulosic residues; opportunities & perspectives. *Int J Biol Sci*, 5(6): 578-95.

Dashtban, M., Schraft, H., Syed, T.A., Qin, W. (2010) Fungal biodegradation and enzymatic modification of lignina. *Int J Biochem Mol Biol*, 1(1): 36-50.

Davies, G., Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure*, 3:853–859.

de Graaff, L.H., H.C. van den Broeck, H.C., van Ooijen, A.J.J., Visser, J. (1994) Regulation of the xylanase-encoding *xlnA* gene of *Aspergillus tubingensis*. *Mol Microbiol*, 12: 479–490.

de Souza, W.R., de Gouvea, P.F., Savoldi, M., Malavazi, I., de Souza Bernardes, L.A., Goldman, M.H., de Vries, R.P., de Castro Oliveira, J.V., Goldman, G.H. (2011) Transcriptome analysis of *Aspergillus niger* grown on sugarcane bagasse. *Biotechnol Biofuels*, 4:40.

De Vries, R.P., Poulsen, C.H., Madrid, S., Visser, J. (1998) *AguA*, the gene encoding an extracellular  $\alpha$ -glucuronidase from *Aspergillus tubingensis*, is specifically induced on xylose and not on glucuronic acid. *J Bacteriol*, 180: 243–249.

De Vries, R.P., Visser, J. (2001) *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. *Microbiol Mol Biol Rev*, 65(4):497–522.

Delmas, S., Pullan, S.T., Gaddipati, S., Kokolski, M., Malla, S., Blythe, M.J., Ibbett, R., Campbell, M., Liddell, S., Aboobaker, A., Tucker, G.A., Archer, D.B. (2012) Uncovering the genome-wide transcriptional responses of the filamentous fungus *Aspergillus niger* to lignocellulose using RNA sequencing. *PLoS Genet*, 8(8):e1002875.

Desmet, T., Cantaert, T., Gualfetti, P., Nerinckx, W., Gross, L., Mitchinson, C., Piens, K. (2007) An investigation of the substrate specificity of the xyloglucanase Cel74A from *Hypocrea jecorina*. *FEBS J*, 274(2):356-63.

Divne, C., Ståhlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J.K.C., Teeri, T.T., Jones, T.A. (1994) The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science*, 265: 524–528.

Dodd, D., Cann, I.K. (2009) Enzymatic deconstruction of xylan for biofuel production. *Glob Change Biol Bioenergy*, 1(1):2-17.

dos Santos Castro, L., Pedersoli, W.R., Antoniêto, A.C., Steindorff, A.S., Silva-Rocha, R., Martinez-Rossi, N.M., Rossi, A., Brown, N.A., Goldman, G.H., Faça, V.M., Persinoti, G.F., Silva, R.N. (2014) Comparative metabolism of cellulose, sophorose and glucose in *Trichoderma reesei* using high-throughput genomic and proteomic analyses. *Biotechnol Biofuels*, 7(1): 41.



DuPont S/A. (2013) Product sheet of Accellerase® 1500 – Cellulase Enzyme Complex for Lignocellulosic Biomass Hydrolysis. DuPont, USA.

Eijsink, V.G.H., Vaaje-Kolstad, G., Vårum, K.M., Horn, S.J. (2008) Towards new enzymes for biofuels: lessons from chitinase research. *Trends Biotechnol*, 26:228–235.

El-Gogary, S., Leite, A., Crivellaro, O., Eveleigh, D.E., El-Dorry, H. (1989) Mechanism by which cellulose triggers cellobiohydrolase I gene expression in *Trichoderma reesei*. *Proc Natl Acad Sci USA*, 86(16):6138-6141.

Eveleigh, D.E., Montenecourt, B.S. (1979) Increasing yields of extracellular enzymes. *Adv Appl Microbiol*, 25:57–74.

Eveleigh, D. E. (1982) Reducing the cost of cellulase production - selection of the hypercellulolytic *Trichoderma reesei* RUT-C30 mutant. New Brunswick:Rutgers University.

Fan, L.T., Lee, Y.H., Beardmore, D.H. (1980) Mechanism of the enzymatic hydrolysis of cellulose: effect of major structural features of cellulose on enzymatic hydrolysis. *Biotechnol Bioeng*, 23: 177-199.

Fonseca, G.G., Heinzle, E., Wittmann, C., Gombert, A.K. (2008) The yeast *Kluyveromyces marxianus* and its biotechnological potential. *Appl Microbiol Biotechnol*, 79(3): 339-354.

Foreman, P.K., Brown, D., Dankmeyer, L., Dean, R., Diener, S., Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J., Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C., Olivares, H.A., Teunissen, P.J., Yao, J., Ward, M. (2003) Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. *J Biol Chem*, 278:31988–31997.

Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunæs, A.C., Stenstrøm, Y., MacKenzie, A., Sørli, M., Horn, S.J., Eijsink, V.G. (2011) Cleavage of cellulose by a CBM33 protein. *Protein Sci*, 20: 1479–1483.

Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M.Jr., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A., Wang, Z.Y. (2011) Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc Natl Acad Sci USA*, 108(9):3803-8.

Fukuda, H., Kondo, A. and Tamalampudi, S. (2009) Bioenergy: Sustainable fuels from biomass by yeast and fungal whole-cell biocatalysts. *Biochemical Engineering Journal* 44: 2-12.

Furukawa, T., Shida, Y., Kitagami, N., Mori, K., Kato, M., Kobayashi, T., Okada, H., Ogasawara, W., Morikawa, Y. (2009) Identification of specific binding sites for XYR1, a transcriptional activator of cellulolytic and xylanolytic genes in *Trichoderma reesei*. *Fungal Genet. Biol*, 46:564–574.

Geysens, S, Pakula, T., Uusitalo, J., Dewerte, I., Penttilä, M., Contreras, R. (2005) Cloning and characterization of the glucosidase II alpha subunit gene of *Trichoderma reesei*: A frameshift mutation results in the aberrant glycosylation profile of the hypercellulolytic strain Rut-C30. *Appl Environ Microbiol*, 71:2910–2924.

Ghazalpour, A., Bennett, B., Petyuk, V.A., Orozco, L., Hagopian, R., Mungrue, I.N., Farber, C.R., Sinsheimer, J., Kang, H.M., Furlotte, N., Park, C.C., Wen, P., Brewer, H., Weitz, K., Camp, D.G., Pan, C., Yordanova, R., Neuhaus, I., Tilford, C., Siemers, N., Gargalovic, P., Eskin, E., Kirchgessner, T., Smith, D.J., Smith, R.D., Lusic, A.J. (2011) Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genet*, 7:e1001393.

Gielkens, M.M.C., Visser, J., de Graaff, L.H. (1997) Arabinoxylan degradation by fungi: characterisation of the arabinoxylan arabinofuranohydrolase encoding genes from *Aspergillus niger* and *Aspergillus tubingensis*. *Curr. Genet*, 31: 22–29.

Gírio, F.M., Fonseca, C., Carvalheiro, F., Duarte, L.C., Marques, S. & Bogel-Lukasik, R. (2010) Hemicellulose for fuel ethanol: A review. *Bioresource Technol*, 10:4775-4800.

Glaesener, S., Honecker, F., Veltman, I.M., Gillis, A.J., Rohlfing, T., Streichert, T., Otto, B., Brummendorf, T.H., Looijenga, L.H., Bokemeyer, C., Balabanov, S. (2008) Comparative proteome, transcriptome, and genome analysis of a gonadal and na extragonadal germ cell tumor cell line. *J Proteome Res*, 7(9):3890-9.

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

Goldemberg, J., Guardabassi, P. (2010) The potential for first-generation ethanol production from sugarcane. *Biofuel Bioprod Bior*, 4(1):17-24.

Gourlay, K., Hu, J., Arantes, V., Andberg, M., Saloheimo, M., Penttilä, M., Saddler, J. (2013) Swollenin aids in the amorphogenesis step during the enzymatic hydrolysis of pretreated biomass. *Bioresour Technol*, 142:498-503.

Gusakov, A.V., Sinitsyn, A.P. (1992) A theoretical analysis of cellulase product inhibition: Effect of cellulase binding constant, enzyme/substrate ratio, and  $\beta$ -glucosidase activity on the inhibition pattern. *Biotechnol Bioeng*, 40: 663–671.

Häkkinen, M., Arvas, M., Oja M., Aro, N., Penttila, M., Saloheimo, M., Pakula, T.M. (2012) Re-annotation of the CAZy genes of *Trichoderma reesei* and transcription in the presence of lignocellulosic substrates. *Microb Cell Fact*, 11:134.

Häkkinen, M., Valkonen, M.J., Westerholm-Parvinen, A., Aro, N., Arvas, M., Vitikainen, M., Penttilä, M., Saloheimo, M., Pakula, T.M. (2014) Screening of candidate regulators for cellulase and

hemicellulase production in *Trichoderma reesei* and identification of a factor essential for cellulase production. *Biotechnol Biofuels*, 7:14.

Hakulinen, N., Paavilainen, S., Korpela, T., Rouvinen, J. (2000) The crystal structure of  $\beta$ -glucosidase from *Bacillus circulans* sp. *alkalophilus*: Ability to form long polymeric assemblies. *J Struct Biol*, 129: 69–79.

Hall, M., Bansal, P., Lee, J.H., Realff, M.J., Bommarius, A.S. (2010) Cellulose crystallinity – a key predictor of the enzymatic hydrolysis rate. *FEBS J*, 277: 1571-1582.

Harris, P.V., Welner, D., McFarland, K.C., Re, E., Navarro Poulsen, J.C., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S., Lo Leggio, L. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry*, 49(15):3305-16.

Hassan, O., Ling, T.P., Maskat, M.Y., Illias, R.M., Badri, K., Jahim, J., Mahadi, N.M. (2013) Optimization of pretreatments of the hydrolysis of oil palm empty fruit bunch fiber (EFBF) using enzyme mixtures. *Biomass Bioenerg*, 56:137-146.

Hasunuma, T., Kondon, A. (2012) Development of yeast cell factories for consolidated bioprocessing of lignocellulose to bioethanol through cell surface engineering. *Biotechnol Adv*, 30: 1207-1218.

Herpoël-Gimbert I., Margeot A., Dolla A., Jan G., Mollé D., Lignon S., Mathis H., Sigoillot J.C., Monot F., Asther M. (2008) Comparative secretome analyses of two *Trichoderma reesei* RUT C30 and CL847 hypersecretory strains. *Biotechnol Biofuels*, 1(1):18.

Hiller, D., Jiang, H., Xu, W., Wong, W.H. (2009) Identifiability of isoform deconvolution from junction arrays and RNA-Seq. *Bioinformatics*, 25(23):3056-9.

Ho, D.P., Ngo, H.H., Guo, W. (2014) A mini review on renewable sources for biofuel. *Bioresour Technol*, 169C: 742-749.

Horn, S.J., Vaaje-Kolstad, G., Westereng, B., Eijsink, V.G.H. (2012) Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels*, 5:45.

IEA, International Energy Agency, 2013. World Energy Outlook 2013, International Energy Agency, IEA/OECD, Paris.

Ilmén, M., Thrane, C., Penttilä, M. (1996) The glucose repressor gene *cre1* of *Trichoderma*: Isolation and expression of a full-length and a truncated mutant form. *Mol Gen Genet*, 251:451–460.

Ilmén, M., Saloheimo, A., Onnela, M.L., Penttilä, M.E. (1997) Regulation of cellulase gene expression in the filamentous fungus *Trichoderma reesei*. *Appl Environ Microbiol*, 63:1298–306

Ilmén, M., Onnela, M.L., Klemsdal, S., Keränen, S., Penttilä, M. (1998) Functional analysis of the cellobiohydrolase I promoter of the filamentous fungus *Trichoderma reesei*. *Mol Gen Genet*, 257(3):386.

Ivanova, C., Bååth, J.A., Seiboth, B., Kubicek, C.P. (2013) Systems analysis of lactose metabolism in *Trichoderma reesei* identifies a lactose permease that is essential for cellulase induction. *PLoS One*, 8(5): e62631.

Jäger, G., Girfoglio, M., Dollo, F., Rinaldi, R., Bongard, H., Commandeur, U., Fischer, R., Spiess, A.C., Büchs, J. (2011) How recombinant swollenin from *Kluyveromyces lactis* affects cellulosic substrates and accelerates their hydrolysis. *Biotechnol Biofuels*, 4:33.

Karimi Aghchegh, R., Németh, Z., Atanasova, L., Fekete, E., Paholcsek, M., Sándor, E., Aquino, B., Druzhinina, I.S., Karaffa, L., Kubicek, C.P. (2014) The VELVET A orthologue VEL1 of *Trichoderma reesei* regulates fungal development and is essential for cellulase gene expression. *PLoS One*, 9(11):e112799. doi: 10.1371/journal.pone.0112799.

Kalliainen, A. (2014) Development of pretreatment technology and enzymatic hydrolysis for biorefineries. Doctoral thesis. VTT Technical Research Centre of Finland, Finland.

Karlsson, J., Saloheimo, M., Siika-Aho, M., Tenkanen, M., Penttilä, M., Tjerneld, F. (2001) Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. *Eur J Biochem*, 268(24):6498-507.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*, 14(4):R36.

Kim, E.S., Lee, H.J., Bang, W.G., Choi, I.G., Kim, K.H. (2009) Functional characterization of a bacterial expansin from *Bacillus subtilis* for enhanced enzymatic hydrolysis of cellulose. *Biotechnol Bioeng*, 102(5):1342-53.

Kim, Y., Nandakumar, M.P., Marten, M.R. (2007) Proteomics of filamentous fungi. *Trends Biotechnol*, 25(9):395-400.

Khuong le, D., Kondo, R., De Leon, R., Anh, T.K., Meguro, S., Shimizu, K., Kamei, I. (2014) Effect of chemical factors on integrated fungal fermentation of sugarcane bagasse for ethanol production by a white-rot fungus, *Phlebia* sp. MG-60. *Bioresour Technol*, 167: 33-40.

Kiran, E.U., Trzcinski, A.P., Ng, W.J., Liu, Y. (2014) Bioconversion of food waste to energy: a review. *Fuel*, 134: 389-399.

Klein-Marcuschamer, D., Oleskowicz-Popiel, P., Simmons, B.A., Blanch, H.W. (2012) The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnol Bioeng*, 109(4):1083-7.

Knob, A., Terrasan, C.R.F., Carmona, E.C. (2010)  $\beta$ -Xylosidases from filamentous fungi: an overview. *World J Microbiol Biotech*, 26:389–407.

Kolasa, M., Ahring, B.K., Lübeck, P.S., Lübeck, M. (2014) Co-cultivation of *Trichoderma reesei* RutC30 with three black *Aspergillus* strains facilitates efficient hydrolysis of pretreated wheat straw and shows promises for on-site enzyme production. *Bioresour Technol*, 169:143-148.

Koike, H., Aerts, A., LaButti, K., Grigoriev, I.V., Baker, S.E. (2013) Comparative genomics analysis of *Trichoderma reesei* strains. *Ind Biotech*, 9(6):352-367.

Kricka, W., Fitzpatrick, J., Bond, U. (2014) Metabolic engineering of yeasts by heterologous enzyme production for degradation of cellulose and hemicellulose from biomass: a perspective. *Front Microbiol*, 5: 174.

Kubicek, C.P., Baker, S., Gamauf, C., Kenerley, C.M., Druzhinina, I.S. (2008) Purifying selection and birth-and-death evolution in the class II hydrophobin gene families of the ascomycete *Trichoderma/Hypocrea*. *BMC Evol Biol*, 8:4.

Kubicek, C.P., Mikus, M., Schuster, A., Schmoll, M., Seiboth, B. (2009) Metabolic engineering strategies for the improvement of cellulase production by *Hypocrea jecorina*. *Biotechnol Biofuels*, 2:19.

Kubicek, C.P., Herrera-Estrella, A., Seidl-Seiboth, V., Martinez, D.A., Druzhinina, I.S., Thon, M., Zeilinger, S., Casas-Flores, S., Horwitz, B.A., Mukherjee, P.K., Mukherjee, M., Kredics, L., Alcaraz, L.D., Aerts, A., Antal, Z., Atanasova, L., Cervantes-Badillo, M.G., Challacombe, J., Chertkov, O., cCluskey, K., Culpier, F., Deshpande, N., von Doehren, H., Ebbole, D.J., Esquivel-Naranjo, E.U., Fekete, E., Flipphi, M., Glaser, F., Gomez-Rodriguez, Y., Gruber, S., Han, C., Henrissat, B., Hermosa, R., Hernandez-Onate, M., Karaffa, L., Kosti, I., Le Crom, S., Lindquist, E., Lucas, S., Lubeck, M., Lubeck, S., Margeot, A., Metz, B., Misra, M., Nevalainen, H., Omann, M., Packer, N., Perrone, G., Uresti-Rivera, E.E., Salamov, A., Schmoll, M., Seiboth, B., Shapiro, H., Sukno, S., Tamayo-Ramos, J.A., Tisch, D., Wiest, A., Wilkinson, H.H., Zhang, M., Coutinho, P.M., Kenerley, C.M., Monte, E., Baker, S.E., Grigoriev, I.V. (2011) Comparative genome sequence analysis underscores mycoparasitism as the ancestral life style of *Trichoderma*. *Genome Biol*, 12(4):R40

Kubicek, C.P. (2013) Chapter 1 – The Plant Biomass. In *Fungi and Lignocellulosic Biomass*. Wiley-Blackwell.

Kubicek, C.P. (2013) Systems biological approaches towards understanding cellulase production by *Trichoderma reesei*. *J Biotechnol*, 163: 133-142.

Kubicek, C.P., Starr, T.L., Glass, N.L. (2014) Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. *Annu Rev Phytopathol*, 52: 427-51.

Kumar, R., Singh, S., Singh, O.V. (2008) Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *J Ind Microbiol Biotechnol*, 35(5): 377-91.

Langmead, B., Salzberg, S.L. (2012) Fast gapped-read alignment with bowtie 2. *Nat Methods*, 9(4): 357-9.

Langston, J.A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E., Sweeney, M.D. () Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl Environ Microbiol*, 77(19):7007-15.

Le Crom, S., Schackwitz, W., Pennacchio, L., Magnuson, J.K., Culley, J.R., Martin, J., Druzhinina, I.S., Mathis, H., Monot, F., Seiboth, B., Cherry, B., Rey, M., Berka, R., Kubicek, C.P., Baker, S.E., Margeot, A. (2009) Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. *Proc Natl Acad Sci USA*, 106(38):16151-6.

Lee, R.A., Lavoie, J.M. (2013) From first- to third-generation biofuels: challenges of producing a commodity from a biomass of increasing complexity. *Animal Frontiers*, 3(2): 6–11.

Lee, S.B., Kim, I.H., Ryu, D.D.Y., Taguchi, H. (1983) Structural properties of cellulose and cellulase reaction mechanism. *Biotechnol Bioeng*, 25: 33-51.

Leite, R.C.C., Leal, M.R.L.V., Cortez, L.A.B., Griffin, W.M., Scandiffio, M.I.G. (2009) Can Brazil replace 5% of the 2025 gasoline world demand with ethanol? *Energy*, 34:655-661.

Levasseur, A., Drula, E., Lombard, V., Coutinho, P.M., Henrissat, B. (2013) Expansion the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*, 6:41.

Li, J., Zhou, P., Liu, H., Wu, K., Xiao, W., Gong, Y., Lin, J., Liu, Z. (2014) Ethanol production from xylan-removed sugarcane bagasse using low loading of commercial cellulase. *Bioresourc Technol*, 163: 390-4.

Li, X.-L., S. Spániková, R. P. de Vries, Biely, P. (2007) Identification of genes encoding microbial glucuronoyl esterases. *FEBS Lett*, 581:4029– 4035.

Li, X., Skory, C.D., Cotta, M.A., Puchart, V., Biely, P. (2008) Novel family of carbohydrate esterases, based on identification of the *Hypocrea jecorina* acetyl esterase gene. *Appl Environ Microbiol*, 74(24):7482–7489.

Lichius, A., Seidl-Seiboth, V., Seiboth, B., Kubicek, C.P. (2014) Nucleo-cytoplasmic shuttling dynamics of the transcriptional regulators XYR1 and CRE1 under conditions of cellulase and xylanase gene expression in *Trichoderma reesei*. *Mol Microbiol*, DOI: 10.1111/mmi.12824.

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

Lin, Y., Tanaka, S. (2006) Ethanol fermentation from biomass resources: current state and prospects. *Appl. Microbiol. Biotechnol.* 69: 627-642.

Linder, M.B., Szilvay, G.R., Nakari-Setälä, T., Penttilä, M. (2005) Hydrophobins: the protein-amphiphiles of filamentous fungi. *FEMS Microbiol Rev*, 29:877-896.

Linder, M.B. (2009) Hydrophobins: proteins that self assemble at interfaces. *Curr Opin Colloid Interface Sci*, 14: 356–363.

Liu, T., Wang, T., Li, X., Liu, X. (2008) Improved heterologous gene expression in *Trichoderma reesei* by cellobiohidrolase I gene (*cbh1*) promoter optimization. *Acta Bioch Bioph Sin*, 40(2): 158-65.

Liu, Y.K., Yang, C.A., Chen, W.C., Wei, Y.H. (2012) Producing bioethanol from cellulosic hydrolyzate via co-immobilized cultivation strategy. *J Biosci Bioeng*, 114 (2): 198-203.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J. (2000) Molecular Cell Biology. New York: W.H. Freeman. 4<sup>th</sup> edition. Disponível em: <http://www.ncbi.nlm.nih.gov/books/NBK21475/>

Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M., Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*, 42: D490-5.

Lynd, L.R., Weimer, P.J., van Zyl, W.H., Pretorius, I.S. (2002) Microbial Cellulose Utilization: Fundamentals and Biotechnology. *Microbiol Mol Biol Rev*, 66(3): 506-577.

Lynd L.R., Zhang Y. (2002) Quantitative determination of cellulase concentration as distinct from cell concentration in studies of microbial cellulose utilization: analytical framework and methodological approach. *Biotechnol Bioeng*, 77:467-475.

Macedo, I. Sugar Cane's Energy – Twelve studies on Brazilian sugar cane agribusiness and its sustainability. UNICA (São Paulo Sugarcane Agro Industry Union), 2005. Disponível em: <<http://sugarcane.org/resource-library/books/Sugar%20Canes%20Energy%20-%20Full%20book.pdf>>.

Macedo, I.C., Seabra, J.E.A., Silva, J.E. (2008) Green house gases emissions in the production and use of ethanol from sugarcane in Brazil: the 2005/2006 averages and a prediction for 2020. *Biomass Bioenerg*, 32:582–595.

Mach, R.L., Strauss, J., Zeilinger, S., Schindler, M., Kubicek, C.P. (1996) Carbon catabolite repression of xylanase I (*xyn1*) gene expression in *Trichoderma reesei*. *Mol Microbiol*, 21:1273-1281.

Mach-Aigner, A.R., Pucher, M.E., Steiger, M.G., Bauer, G.E., Preis, S.J., Mach, R.L. (2008) Transcriptional regulation of *xyr1*, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in *Hypocrea jecorina*. *Appl Environ Microbiol*, 74:6554-62.

Margeot, A., Hahn-Hagerdal, B., Edlund, M., Slade, R., Monot, F. (2009) New improvements for lignocellulosic ethanol. *Curr Opin Biotechnol*, 20(3):372-80.

Martens-Uzunova, E.S., Schaap, P.J. (2009) Assessment of the pectin degrading enzyme network of *Aspergillus niger* by functional genomics. *Fungal Genet Biol*, 46(Suppl 1):S170–S179.

Martinez D., Berka R.M., Henrissat B., Saloheimo M., Arvas M., Baker S.E., Chapman J., Chertkov O., Coutinho P.M., Cullen D., Danchin E.G., Grigoriev I.V., Harris P., Jackson M., Kubicek C.P., Han C.S., Ho I., Larrondo L.F., de Leon A.L., Magnuson J.K., Merino S., Misra M., Nelson B., Putnam N., Robbertse B., Salamov A.A., Schmoll M., Terry A., Thayer N., Westerholm-Parvinen A., Schoch C.L., Yao J., Barabote R., Nelson M.A., Detter C., Bruce D., Kuske C.R., Xie G., Richardson P., Rokhsar D.S., Lucas S.M., Rubin E.M., Dunn-Coleman N., Ward M., Brettin T.S. (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol*, 26:553-560.

Marx, I.J., van Wyk, N., Smit, S., Jacobson, D., Viljoen-Bloom, M., Volschenk, H. (2013) Comparative secretome analysis of *Trichoderma asperellum* S4F8 and *Trichoderma reesei* Rut C30 during solid-state fermentation on sugarcane bagasse. *Biotechnol Biofuels*, 6(1): 172.

Master, E.R., Zheng, Y., Storms, R., Tsang, A., Powlowski, J. (2008) A xyloglucan-specific family 12 glycosyl hydrolase from *Aspergillus niger*: Recombinant expression, purification and characterization. *Biochem J*, 411:161–170.

McMillan, J.D. (1993) Pretreatment of lignocellulosic biomass. In: Himmel ME, Baker O, Overend RP, editors. Enzymatic conversion of biomass for fuel production. Washington, D.C: American Chemical Society; 1993. p. 292e323. Mandels, M., Reese, E. T. (1957) Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. *J Bacteriol*, 73: 269–278.

McQueen-Mason S., Cosgrove D.J. (1994) Disruption of hydrogen bonding between plant cell wall polymers by proteins that induce wall extension. *Proc Natl Acad Sci USA*, 91(14):6574-8.

Mello-de-Sousa, T.M., Gorsche, R., Rassinger, A., Poças-Fonseca, M.J., Mach, R.L., Mach-Aigner, A.R. (2014) A truncated form of the Carbon catabolite repressor 1 increases cellulase production in *Trichoderma reesei*. *Biotechnol Biofuels*, 7: 129.



Menon, V., Rao, M. (2012) Trends in bioconversion of lignocellulose: Biofuels, platform chemicals & biorefinery concept. *Prog Energ Combust*, 38: 522-550.

Merino, S.T., Cherry, J. (2007) Progress and challenges in enzyme development for biomass production. *Adv Biochem Eng Biotechnol*, 108:95-120.

Michaud, P., Da Costa, A., Courtois, B., Courtois, J. (2003) Polysaccharide lyases: recent developments as biotechnological tools. *Crit Rev Biotechnol*, 23(4): 233-66.

Montenecourt, B. S. (1983) *Trichoderma reesei* cellulases. *Trends Biotechnol*. 1:156–161

Morgenstern, I., Powlowski, J., Tsang, A. (2014) Fungal cellulose degradation by oxidative enzymes: from dysfunctional GH61 family to powerful lytic polysaccharide monooxygenase family. *Brief Funct Genomics*, 13(6):471-81.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods*, 5:621-628.

Moser, F., Irwin, D., Chen, S., Wilson, D.B. (2008) Regulation and characterization of *Thermobifida fusca* carbohydrate-binding module proteins E7 and E8. *Biotechnol Bioeng*, 100:1066–1077.

Müller, S., Baldin, C., Groth, M., Guthke, R., Kniemeyer, O., Brakhage, A.A., Valiante, V. (2012) Comparison of transcriptome technologies in the pathogenic fungus *Aspergillus fumigatus* reveals novel insights into the genome and MpkA dependent gene expression. *BMC Genomics*, 13:519.

Mustacchi, R., Hohmann, S., Nielsen, J. (2006) Yeast systems biology to unravel the network of life. *Yeast*, 23: 227–238.

Nagalakshmi, U., Waern, K., Snyder, M. (2010) RNA-Seq: a method for comprehensive transcriptome analysis. *Curr Protoc Mol Biol*, chapter 4:unit 4.11.1-13.

Narendja, F.M., Davis, M.A., Hynes, M.J. (1999) AnCF, the CCAAT binding complex of *Aspergillus nidulans*, is essential for the formation of a DNase I-hypersensitive site in the 5' region of the amdS gene. *Mol Cell Biol*, 19: 6523–6531.

Nass, L.L., Pereira, P.A.A., Ellis, D. (2007) Biofuels in Brazil: An overview. *Crop Science*, 47: 2228-2237.

Nickel, W., Rabouille, C. (2009) Mechanisms of regulated unconventional protein secretion. *Nat Rev Mol Cell Biol*, 10(2):148-55.

Nielsen, J., Jewett, M.C. (2007) Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS Yeast Res*, 8(1):122-131.

Nigam, P.S., Singh, A. (2011) Production of liquid biofuels from renewable resources. *Prog Energ Combust*, 37: 52-68.

Nitsche, B.M., Jorgensen, T.R., Akeroyd, M., Meyer, V., Ram, A.F. (2012) The carbon starvation response of *Aspergillus niger* during submerged cultivation: insights from the transcriptome and secretome. *BMC Genomics*, 13:380.

Nogawa, M., Goto, M., Okada, H., Morikawa, Y. (2001) L-sorbose induces cellulase gene transcription in the cellulolytic fungus *Trichoderma reesei*. *Curr Genet*, 38(6):329-34.

Novozymes A/S. (2012) Product sheet of Cellic® CTec3. Novozymes A/S, Denmark.

Oliveira, F.M., Pinheiro, I.O., Souto-Maior, A.M., Martin, C., Gonçalves, A.R., Rocha, G.J. (2013) Industrial-scale steam explosion pretreatment of sugarcane straw for enzymatic hydrolysis of cellulose for production of second generation ethanol and value-added products. *Bioresour Technol*, 130: 168-73.

Olsson, L., Christensen, T., Hansen, K.P., Palmqvist, E.A. (2003) Influence of the carbon source on production of Ohtaki, S., Maeda, H., Takahashi, T., Yamagata, Y., Hasegawa, F., Gomi, K., Nakajima, T., Abe, K. (2006) Novel hydrophobic surface binding protein, HsbA, produced by *Aspergillus oryzae*. *Appl Environ Microb*, 72:2407-2413.

Otto, C., Stadler, P.F., Hoffmann, S. (2014) Lacking alignments? The next-generation sequencing mapper segemehl revisited. *Bioinformatics*, 30(13):1837-43. doi: 10.1093/bioinformatics/btu146.

Ozsolak, F., Milos, P.M. (2011) RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet*, 12:87-98.

Penttilä, Lehtovaara, P., Nevalainen, H., Bhikhabhai, R., Knowles, J. (1986) Homology between cellulose genes of *Trichoderma reesei*: complete nucleotide sequence of the endoglucanase I gene. *Gene*, 45(3):253-63.

Pakula, T., Saloheimo, M., Hakkinen, M., Westerholm-Parvinen, A., Vitikainen, M., inventors; Teknologian Tutkimuskeskus VTT, assignee. Production of proteins in filamentous fungi. Finland patent 20130244276. May 30, 2011a.

Pakula, T., Saloheimo, M., Hakkinen, M., Westerholm-Parvinen, A., Penttilä, M., Vitikainen, M., inventors; Teknologian Tutkimuskeskus VTT, assignee. Method for improved protein production in filamentous fungi. Finland patent 20130084604. May 30, 2011b.

Pakula, T., Saloheimo, M., Hakkinen, M., Westerholm-Parvinen, A., Penttilä, M., Vitikainen, M., inventors; Teknologian Tutkimuskeskus VTT, assignee. Method for protein production in filamentous fungi. Finland patent 20130078674. May 30, 2011c.

Pauly, M., Keegstra, K. (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J*, 54:559–568.

Peng, F., Peng, P., Xu, F., Sun, R.C. (2012) Fractional purification and bioconversion of hemicelluloses. *Biotechnol Adv*, 30(4): 879-903.

Percival Zhang, Y.H., Himmel, M.E., Mielenz, J.R. (2006) Outlook for cellulase improvement: screening and selection strategies. *Biotechnol Adv*, 24: 452-481.

Pokkuluri, P.R., Duke, N.E., Wood, S.J., Cotta, M.A., Li, X.L., Biely, P., Schiffer, M. (2011) Structure of the catalytic domain of glucuronoyl esterase Cip2 from *Hypocrea jecorina*. *Proteins*, 79(8):2588-92.

Peterson, R., Nevalainen, H. (2012) *Trichoderma reesei* RUT-C30 – thirty years of strain improvement. *Microbiology*, 158: 58-68.

Polizeli, M.L., Rizzatti, A.C., Monti, R., Terenzi, H.F., Jorge, J.A., Amorim, D.S. (2005) Xylanases from fungi: properties and industrial applications. *Appl Microbiol Biot*, 67:577-591.

Portnoy, T., Margeot, A., Seidl-Seiboth, V., Le Crom, S., Ben Chaabane, F., Linke, R., Seiboth, B., Kubicek, C.P. (2011) Differential regulation of the cellulase transcription factors XYR1, ACE2, and ACE1 in *Trichoderma reesei* strains producing high and low levels of cellulase. *Eukaryotic Cell*, 10:262–271.

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., Gu, Y. (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*, 13:341. doi: 10.1186/1471-2164-13-341.

Quinlan, R.J., Sweeney, M.D., Lo Leggio, L., Otten, H., Poulsen, J.-C.N., Johansen, K.S., Krogh, K.B.R.M., Jørgensen, C.I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C.P., Dupree, P., Xu, F., Davies, G.J., Walton, P.H. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proc Natl Acad Sci USA*, 108:15079-15084.

Ragauskas, A.J., Williams, C.K., Davidson, B.H., Britovsek, G., Cairney, J., Eckert, C.A., Frederick Jr., W.J., Hallett, J.P., Leak, D.J., Liotta, C.L., Mielenz, J.R., Murphy, R., Templer, R., Tschaplinski, T. (2006) The path forward for biofuels and biomaterials. *Science*, 311: 484-489.

Rahman, Z., Shida, Y., Furukawa, T., Suzuki, Y., Okada, H., Ogasawara, W., Morikawa, Y. (2009) Application of *Trichoderma reesei* cellulose and xylanase promoters through homologous recombinant for enhanced production of extracellular  $\beta$ -glucosidase I. *Biosci Biotechnol Biochem*, 73(5):1083-1089.

Ravanal, M.C. Callegari, E., Eyzaguirre, J. (2010) Novel bifunctional alpha-L-arabinofuranosidase/xylobiohydrolase (ABF3) from *Penicillium purpurogenum*. *Appl Environ Microbiol*, 76(15):5247-53.

Rauscher, R., Wurleiner, E., Wacenovsky, C., Aro, N., Stricker, A.R., Zeilinger, S., Kubicek, C.P., Penttila, M., Mach, R.L. (2006) Transcriptional regulation of *xyn1*, encoding xylanase I, in *Hypocrea jecorina*. *Eukaryotic Cell*, 5: 447-456.

Reddy, V.S., Shlykov, M.A., Castillo, R., Sun, E.I., Saier, M.H.Jr. (2012) The major facilitator superfamily (MFS) revisited. *FEBS J*, 279(11):2022-35.

REN21. 2014. Renewables 2014 Global Status Report (Paris: REN21 Secretariat).

Ries, L., Pullan, S.T., Delmas, S., Malla, S., Blythe, M.J., Archer, D.B. (2013) Genome-wide transcriptional response of *Trichoderma reesei* to lignocellulose using RNA sequencing and comparison with *Aspergillus niger*. *BMC Genomics*, 14:541.

Ries, L.N.A.R. (2013) Regulation of genes encoding enzymes involved in plant cell wall deconstruction in *Trichoderma reesei*. Doctoral thesis. University of Nottingham. United Kingdom.

Robinson, M.D., McCarthy, D.J., Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139-140.

Ronne, H. (1995) Glucose repression in fungi. *Trends Genet*, 11: 12-17.

Rosgaard, L., Pedersen, S., Langston, J., Akerhielm, D., Cherry, J.R., Meyer, A.S. (2007) Evaluation of minimal *Trichoderma reesei* cellulase mixtures on differently pretreated barley straw substrates. *Biotechnol Prog*, 23:1270–1276.

Rubin, E.M. (2008) Genomics of cellulosic biofuels. *Nature*, 454; 841-845.

Ruijter, G.J.G., Visser, J. (1997) Carbon repression in *Aspergilli*. *FEMS Microbiol Lett*, 151: 103–114.

Saha, B.C. (2000) Alpha-L-arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. *Biotechnol Adv*, 18(5):403-23.

Saha, B.D. (2003) Hemicellulose bioconversion. *J Ind Microbiol Biotechnol*, 30: 279e91.8;25.

Saloheimo, M., Nakari-Setälä, T., Tenkanen, M., Penttilä, M. (1997) cDNA cloning of a *Trichoderma reesei* cellulase and demonstration of endoglucanase activity by expression in yeast. *Eur J Biochem*, 249(2):584-91.

Saloheimo, A., Aro, N., Ilmen, M., Penttila, M. (2000) Isolation of the *ace1* gene encoding a Cys(2)-His(2) transcription factor involved in regulation of activity of the cellulase promoter *cbh1* of *Trichoderma reesei*. *J Biol Chem*, 275(8):5817–5825.

Saloheimo, M., Paloheimo, M., Hakola, S., Pere, J., Swanson, B., Nyysönen, E., Bhatia, A., Ward, M., Penttilä, M. (2002) Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. *Eur J Biochem*, 269(17):4202-11.

Saloheimo, A., Rauta, J., Stasyk, O.V., Sibirny, A.A., Penttilä, M., Ruohonen, L. (2007) Xylose transport studies with xylose-utilizing *Saccharomyces cerevisiae* strains expressing heterologous and homologous permeases. *Appl Microbiol Biotechnol*, 74:1041-1052.

Sanchez, C. (2009) Lignocellulosic residues: biodegradation and bioconversion by fungi. *Biotechnol Adv*, 27: 185-194

Sánchez, Ó.J., Cardona, C.A. (2008) Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresour Technol*, 99(13): 5270-5295.

Sarkar, N., Ghosh, S.K., Bannerjee, S., Aikat, K. (2012) Bioethanol production from agricultural wastes: an overview. *Renew Energ*, 37(1): 19-27.

Scheller, H.V., Ulvskov, P. (2010) Hemicelluloses. *Annu Rev Plant Biol*, 61: 263-89.

Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Fungal Barcoding Consortium, Fungal Barcoding Consortium Author List. (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc Natl Acad Sci USA*, 109:6241–6246.

Schuster, A., Schmoll, M. (2010) Biology and biotechnology of *Trichoderma*. *Appl Microbiol Biotechnol*, 87:787-99.

Seiboth, B., Karimi, R.A., Phatale, P.A., Linke, R., Hartl, L., Sauer, D.G., Smith, K.M., Baker, S.E., Freitag, M., Kubicek, C.P. (2012) The putative protein methyltransferase LAE1 controls cellulase gene expression in *Trichoderma reesei*. *Mol Microbiol*, 84(6):1150-64.

Seidl, V., Gamauf, C., Druzhinina, I.S., Seiboth, B., Hartl, L., Kubicek, C.P. (2008) The *Hypocrea jecorina* (*Trichoderma reesei*) hypercellulolytic mutant RUT C30 lacks a 85 kb (29 gene-encoding) region of the wild-type genome. *BMC Genomics*, 9:327.

Seidl, V., Seibel, C., Kubicek, C. P., Schmoll, M. (2009) Sexual development in the industrial workhorse *Trichoderma reesei*. *Proc Natl Acad Sci USA* 106, 13909–13914.

Sendelius, J., 2005. Steam pretreatment optimisation for sugarcane bagasse in bioethanol production, Master of Science Thesis. Lund University, Sweden.

Simmons, E. G. (1977). Classification of some cellulase-producing *Trichoderma* species. In *Second International Mycological Congress, Abstracts*, pp. 618. Edited by Simmons, E. G. University of South Florida, Tampa, FL.

Silva-Rocha, R., Castro, L.dos S., Antoniêto, A.C., Guazzaroni, M.E., Persinoti, G.F., Silva, R.N. (2014) Deciphering the cis-regulatory elements for XYR1 and CRE1 regulators in *Trichoderma reesei*. *PLoS One*, 9(6):e99366.

Shoemaker, S., Watt, K., Tsitovsky, G., Cox, R. (1983) Characterization and properties of cellulases purified from *Trichoderma reesei* strain L27. *Biotechnology (N. Y.)* 1: 687–690.

Souza A.P., Leite D.C.C., Pattathil S., Hahn M.G., Buckeridge M.S. (2012) Composition and Structure of Sugarcane Cell Wall Polysaccharides: Implications for Second-Generation Bioethanol Production. *Bioenerg Res*, 1-16.

Stals, I., Sandra, K., Geysens, S., Contreras, R., Beeumen, J.V., Claeysens, M. (2004) Factor influencing glycosylation of *Trichoderma reesei* cellulases. I: Postsecretorial changes of the O- and N-glycosylation pattern of Cel7A. *Glycobiology*, 14(8):713-24.

Steiger, M.G., Mach, R.L., Mach-Aigner, A.R. (2010) An accurate normalization strategy for RT-qPCR in *Hypocrea jecorina* (*Trichoderma reesei*). *J Biotechnol*, 145(1):30-7.

Stricker, A.R., Mach, R.L., de Graaff, L.H. (2008) Regulation of transcription of cellulases- and hemicellulases- encoding genes in *Aspergillus niger* and *Hypocrea jecorina* (*Trichoderma reesei*). *Appl Microbiol Biotechnol*, 78: 211-220.

Sun, J., Tian, C., Diamond, S., Glass, N.L. (2012) Deciphering transcriptional regulatory mechanisms associated with hemicellulose degradation in *Neurospora crassa*. *Eukaryotic cell*, 11:482–493.

Sygmund, C., Kracher, D., Scheiblbrandner, S., Zahma, K., Felice, A.K., Harreither, W., Kittl, R., Ludwig, R. (2012) Characterization of the two *Neurospora crassa* celobiose dehydrogenases and their connection to oxidative cellulose degradation. *Appl Environ Microbiol*, 78(17):6161-71.

Takahashi, N., Koshijima, T. (1988) Ester linkages between lignin and glucuronoxylan in a lignin-carbohydrate complex from beech (*Fagus crenata*) wood. *Wood Sci Technol*, 22:231–241.

Teeri, T. T., A. Koivula, M. Linder, G. Wohlfahrt, C. Divne, and T. A. Jones. (1998) *Trichoderma reesei* cellobiohydrolases: why so efficient on crystalline cellulose? *Biochem Soc Trans*, 26:173–178.

Teeri, T. T. (1997) Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends Biotechnol*, 15: 160–167.

The World Bank, 2014. World Development Indicators: Energy production and use. *Last updated: 22-Jul-2014*.

Thevenot, M., Dignac, M.F., Rumpel, C. (2010) Fate of lignins in soils: a review. *Soil Biol Biochem*, 42: 1200-1211.

- Tian, C., Beeson, W.T., Iavarone, A.T., Sun, J., Marletta, M.A., Cate, J.H., Glass, N.L. (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*. *Proc Natl Acad Sci USA*, 106:22157–22162.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*, 7(3):562-78.
- Triana, C.A.R. (2011) Energetics of Brazilian ethanol: comparison between assessment approaches. *Energ Policy*, 39(8):4605-4613.
- Tsukagoshi, N., Kobayashi, T., Kato, M. (2001) Regulation of the amylolytic and (hemi-)cellulolytic genes in aspergilli. *J Gen Appl Microbiol*, 47: 1–19.
- Turro, E., Su, S.Y., Gonçalves, Â., Coin, L.J., Richardson, S., Lewin, A. (2011) Haplotype and isoform specific expression estimation using multi-mapping RNA-seq reads. *Genome Biol*, 12(2):R13. doi: 10.1186/gb-2011-12-2-r13.
- Vaaje-Kolstad, G., Westereng, B., Horn, S.J., Liu, Z.L., Zhai, H., Sørli, M., Eijsink, V.G.H. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science*, 330:219–222.
- Vaaje-Kolstad, G., Bøhle, L.A., Gaseidnes, S., Dalhus, B., Bjøras, M., Mathiesen, G., Eijsink, V.G.H. (2012) Characterization of the chitinolytic machinery of *Enterococcus faecalis* V583 and high-resolution structure of its oxidative CBM33 enzyme. *J Mol Biol*, 416:239–254.
- van den Brink, J., de Vries, R.P. (2011) Fungal enzyme sets for plant polysaccharide degradation. *Appl Microbiol Biotechnol*, 91(6): 1477-92.
- van Peij, N.N., Brinkmann, J., Vrsanska, M., Visser, J., de Graaff, L.H. (1997)  $\beta$ -Xylosidase activity, encoded by *xlnD*, is essential for complete hydrolysis of xylan by *Aspergillus niger* but not for induction of the xylanolytic enzyme spectrum. *Eur J Biochem*, 245: 164–173.
- van Peij, N.N., Visser, J., de Graaff, L.H. (1998) Isolation and analysis of *xlnR*, encoding a transcriptional activator co-ordinating xylanolytic expression in *Aspergillus niger*. *Mol Microbiol*, 27(1):131–142.
- Verbeke, J., Coutinho, P., Mathis, H., Quenot, A., Record, E., Asther, M., Heiss-Blanquet, S. (2009) Transcriptional profiling of cellulase and expansin-related genes in a hypercellulolytic *Trichoderma reesei*. *Biotechnol Lett*, 31(9):1399-1405.
- Verma D., Jin S., Kanagaraj A., Singh N.D., Daniel J., Kolattukudy P.E., Miller M., Daniell H. (2013) Expression of fungal cutinase and swollenin in tobacco chloroplasts reveals novel enzyme functions and/or substrates. *PLoS One*, 8(2):e57187.

Wanderley, M.C., Martín, C., Rocha, G.J., Gouveia, E.R. (2013) Increase in ethanol production from sugarcane bagasse based on combined pretreatments and fed-batch enzymatic hydrolysis. *Bioresour Technol*, 128:448-53.

Wang, K., Singh, D., Zeng, Z., Coleman, S.J., Huang, Y., Savich, G.L., He, X., Mieczkowski, P., Grimm, S.A., Perou, C.M., MacLeod, J.N., Chiang, D.Y., Prins, J.F., Liu, J. (2010) MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res*, 38:1-14.

Wang, L., Wang, S., Li, W. (2012) RSeQC: Quality Control of RNA-seq experiments. *Bioinformatics*, 28(16): 2184-5.

Wang, X., Zhang, X., Liu, L., Xiang, M., Wang, W., Sun, X., Che, Y., Guo, L., Liu, G., Guo, L., Wang, C., Yin, W.B., Stadler, M., Zhang, X., Liu, X. (2015) Genomic and transcriptomic analysis of the endophytic fungus *Pestalotiopsis fici* reveals its lifestyle and high potential for synthesis of natural products. *BMC Genomics*, 16(1):28.

Wang, Z., Gerstein, M., Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, 10(1):57-63.

Westereng, B., Ishida, T., Vaaje-Kolstad, G., Wu, M., Eijsink, V.G.H., Igarashi, K., Samejima, M., Ståhlberg, J., Horn, S.J., Sandgren, M. (2011) The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* is a metal-dependent oxidative enzyme that cleaves cellulose. *PLoS One*, 6:e27807.

Wick, A.N., Drury, D.R., Nakada, H.I., Wolfe, J.B. (1957) Localization of the primary metabolic block produced by 2-deoxyglucose. *J Biol Chem*, 224:963–969.

Wong, K.K.Y., Saddler, J.N. (1993) Applications of hemicellulases in the food, feed, and pulp and paper industries. In: Coughlan MP, Hazlewood GP (eds) Hemicellulose and hemicellulases. Portland, London, pp 127–143

World Energy Resources - 2013 Survey Council. World Energy Council, 2013. Disponível em: [http://www.worldenergy.org/wp-content/uploads/2013/09/Complete\\_WER\\_2013\\_Survey.pdf](http://www.worldenergy.org/wp-content/uploads/2013/09/Complete_WER_2013_Survey.pdf)

Wu, C., Yang, F., Smith, K.M., Peterson, M., Dekhang, R., Zhang, Y., Zucker, J., Bredeweg, E.L., Mallapa, C., Zhou, X., Lyubetskaya, A., Townsend, J.P., Galagan, J.E., Freitag, M., Dunlap, J.C., Bell-Pedersen, D., Sachs, M.S. (2014) Genome-wide characterization of light-regulated genes in *Neurospora crassa*. *G3 (Bethesda)*, 4(9):1731-45.

Xie, S., Syrenne, R., Sun, S., Yuan, J.S. (2014) Exploration of Natural Biomass Utilization Systems (NBUS) for advanced biofuel – from systems biology to synthetic design. *Curr Opin Biotech*, 27: 195-203.



Xu, Z., Huang, F. (2014) Pretreatment methods for bioethanol production. *Appl Biochem Biotechnol*, 174: 43-62.

Yan, N. (2013) Structural advances for the major facilitator superfamily (MFS) transporters. *Trends Biochem Sci*, 38(3):151-9.

Yanase, S., Hasunuma, T., Yamada, R., Tanaka, T., Ogino, C., Fukuda, H., Kondo, A. (2010) Direct ethanol production from cellulosic materials at high temperature using the thermotolerant yeast *Kluyveromyces marxianus* displaying cellulolytic enzymes. *Appl Microbiol Biotechnol*, 88: 381-388.

Yip, V.L., Withers, S.G. (2006) Breakdown of oligosaccharides by the process of elimination. *Curr Opin Chem Biol*, 10: 147–155.

Yuan, J.S., Tiller, K.H., Al-Ahmad, H., Stewart, N.R., Stewart Jr, C.N. (2008) Plants to power: bioenergy to fuel to the future. *Trends Plant Sci*, 13(8): 421-9.

Zhang, N.N., Liu, X., Sun, J., Wu, Y., Li, Q.W. (2009) Nonclassical mechanisms of secretory protein in eukaryotic cells. *Yi Chuan*, 31:-29-35.

Zhang, W., Kou, Y., Xu, J., Cao, Y., Zhao, G., Shao, J., Wang, H., Wang, Z., Bao, X., Chen, G., Liu, W. (2013) Two major facilitator superfamily sugar transporters from *Trichoderma reesei* and their roles in induction of cellulose biosynthesis. *J Biol Chem*, 288(46):32861-72.

Zhang, Y.P., Lynd, L.R. (2004) Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. *Biotechnol Bioeng*, 88: 797–824.



# **ANEXOS**



**ANEXO I.** Manuscrito submetido para a revista *PLoS ONE*.

Comparative secretome analysis of *Trichoderma reesei* and  
*Aspergillus niger* during growth on sugarcane biomass

Gustavo Pagotto Borin<sup>1</sup>: gustavo.borin@bioetanol.org.br

Camila Cristina Sanchez<sup>1</sup>: camila.sanchez@bioetanol.org.br

Amanda Pereira de Souza<sup>2</sup>: amanda.psouza@usp.br

Eliane Silva de Santana<sup>1</sup>: eliane.santana@bioetanol.org.br

Aline Tieppo de Souza<sup>1</sup>: aline.souza@bioetanol.org.br

Adriana Franco Paes Leme<sup>3</sup>: adriana.paesleme@lnbio.cnpem.br

Fabio Marcio Squina<sup>1</sup>: fabio.squina@bioetanol.org.br

Marcos Buckeridge<sup>2</sup>: msbuckeridge@gmail.com

Gustavo Henrique Goldman<sup>1,4</sup>: ggoldman@usp.br

\*Juliana Velasco de Castro Oliveira<sup>1</sup>, corresponding author: juliana.velasco@bioetanol.org.br

<sup>1</sup> Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Caixa Postal 6170, 13083-970, Campinas, São Paulo, Brazil.

<sup>2</sup> Laboratório de Fisiologia Ecológica de Plantas (LAFIECO), Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, Rua do Matão 277, Cidade Universitária, Butantã, Caixa Postal 11416, CEP 05508-090, São Paulo, Brazil.

<sup>3</sup> Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Caixa Postal 6192, 13083-970, Campinas, São Paulo, Brazil.

<sup>4</sup> Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Av do Café S/N, CEP 14040-903, Ribeirão Preto, São Paulo, Brazil.

# Abstract

## Background

Our dependence on fossil fuel sources and concern about the environment has generated a worldwide interest in establishing new sources of fuel and energy. Thus, the use of ethanol as a fuel is advantageous because it is an inexhaustible energy source and has minimal environmental impact. Currently, Brazil is the world's second largest producer of ethanol, which is produced from sugarcane juice fermentation. However, several studies suggest that Brazil could double its production per hectare by using sugarcane bagasse and straw, known as second-generation (2G) bioethanol. Nevertheless, the use of this biomass presents a challenge because the plant cell wall structure, which is composed of complex sugars (cellulose and hemicelluloses), must be broken down into fermentable sugar, such as glucose and xylose. To achieve this goal, several types of hydrolytic enzymes are necessary, and these enzymes represent the majority of the cost associated with 2G bioethanol processing. Reducing the cost of the saccharification process can be achieved via a comprehensive understanding of the hydrolytic mechanisms and enzyme secretion of polysaccharide-hydrolyzing microorganisms. In many natural habitats, several microorganisms degrade lignocellulosic biomass through a set of enzymes that act synergistically. In this study, two fungal species, *Aspergillus niger* and *Trichoderma reesei*, were grown on sugarcane biomass with two levels of cell wall complexity, culm *in natura* and pretreated bagasse. The production of enzymes related to biomass degradation was monitored using secretome analyses after 6, 12 and 24 hours. Concurrently, we analyzed the sugars in the supernatant.

## Results

Analyzing the concentration of monosaccharides in the supernatant, we observed that both species are able to disassemble the polysaccharides of sugarcane cell walls since 6 hours post-inoculation. The sugars from the polysaccharides such as arabinoxylan and  $\beta$ -glucan (that compose the most external part of the cell wall in sugarcane) are likely the first to be released and assimilated by both species of fungi. At all time points tested, *A. niger* produced more enzymes (quantitatively and qualitatively) than *T. reesei*. However, the most important enzymes related to biomass degradation, including cellobiohydrolases, endoglucanases,  $\beta$ -glucosidases,  $\beta$ -xylosidases, endoxylanases, xyloglucanases, and  $\alpha$ -arabinofuranosidases, were identified in both secretomes. We also noticed that the both fungi produce more enzymes when grown in culm as a single carbon source.

## Conclusion

Our work provides the most detailed quantitative secretome analysis of *A. niger* and *T. reesei* grown on sugarcane biomass. Our data indicate that a combination of enzymes from both fungi is an interesting option to increase saccharification efficiency. In other words, these two fungal species might be combined for their usage in industrial processes.

## Introduction

The increasing demand for sustainable energy has promoted considerable efforts to replace fossil fuels with biofuels. As the second world's largest producer and exporter of ethanol from sugarcane, approximately half of Brazil's fuel supply is produced from renewable energy sources [1]. Currently, Brazilian production relies on the fermentation of

sucrose, known as first-generation (1G) bioethanol. If second-generation (2G) bioethanol was commercialized, Brazil could increase bioethanol production by approximately 40% [2]. To reach this level in industrial processes, the obstacle of cell wall recalcitrance must be overcome. Cell wall recalcitrance is a phenomenon directly related to the enormous complexity of the plant cell wall [3]. In the case of sugarcane, de Souza *et al.* [4] proposed a model for the architecture of polymers within the cell walls of the leaf and culm (the stem of the sugarcane) that included the structural complexity of hemicelluloses, such as arabinoxylan, xyloglucan, and mixed-linkage- $\beta$ -glucans, as well as pectins, such as homogalacturonan and arabinogalactans. They found that sugarcane tissues are composed of ca. 30% of cellulose and 60% hemicelluloses, with pectins and lignin accounting for the rest of the biomass [4].

The biomass of sugarcane displaying these features is transformed into bagasse, a major residue from the Brazilian agroindustry (280 kg per 1 ton of sugarcane crushed) [5]. Bagasse is obtained from a process that crushes and washes biomass (to obtain sucrose), changing the composition in relation to culm. Bagasse is composed of cellulose (40-50 %), hemicellulose (25-35 %), and lignin (15-20 %) [6,7], highlighting the fact that a portion of hemicelluloses and pectins are washed out during sucrose extraction. Thus, the compositions of sugarcane culms *in natura* and bagasse are considerably different, with the former displaying higher complexity and proportionally higher levels of soluble polymers that belong to the classes of hemicelluloses and pectins.

Although considerable progress has been made in the saccharification of recalcitrant plant biomass, the cost of 2G bioethanol will not become economically competitive unless the full conversion of lignocellulose biomass can be reached. Complete hydrolysis of cellulose yields glucose, whereas hemicellulose hydrolysis can produce monomers of xylose, arabinose, mannose, glucose and galactose. In order to break down the cell wall and release these



monomers, pre-treatment with enzymatic cocktails are necessary prior to hydrolysis, and this step constitutes the majority of the cost in 2G bioethanol processing [8]. A better understanding of the hydrolytic mechanisms and enzyme secretion of polysaccharide-hydrolyzing microorganisms is needed to overcome the cost associated with enzyme pretreatments. In many natural habitats, plant biomass is degraded by a variety of lignocellulolytic microorganisms that work together to break down the recalcitrant structure of lignocellulosic materials. Although bacteria and yeast (the latter more rarely) produce hydrolytic enzymes [9,10], most enzymes used in commercial cocktails are derived from fungi, such as *Aspergillus niger* and *Trichoderma reesei*, due to their efficiency in producing and secreting a broad range of cellulases and hemicellulases.

*A. niger* is industrially used to produce many pectinases [11,12] and hemicellulases [13,14]. A sequencing effort reported that *A. niger* contains 14,056 genes [15], and it has one of the most remarkable sets of genes encoding hydrolytic enzymes among sequenced fungal genomes. According to the Carbohydrate-Active Enzymes (CAZY) database (<http://www.cazy.org/>), *A. niger* has more than 250 glycoside hydrolases (GHs). Another hyper producer of cellulolytic enzymes is *T. reesei* RUT-C30. This strain was obtained from the wild-type strain, QM6a, after three rounds of random mutagenesis, with the aim of increasing cellulase production.

Due to their great potential for producing hydrolytic enzymes, both of these fungi have been the focus of several studies on GH discovery and there is a marked effort to understand the regulation of the expression of genes that encoding them. To date, only one master carbon repression regulator has been described (CreA/Cre1). The *A. niger* transcription factor, XlnR (and the *T. reesei* orthologues Xyr1), is a major positive transcriptional regulator of xylanases and cellulases encoding genes for this species. In *A. niger*, the expression of most

cellulases and hemicellulases is co-regulated by the same inducer (xylose), but for *T. reesei*, at least four different inducers have been described (xylose, xylobiose, sophorose and lactose) [16,17]. Several differences in the regulation of GH production between these two fungi have been already described [16,18,19], but comparative studies could provide a more comprehensive overview of how these important industrial species sense and produce hydrolytic enzymes.

Although many secretome studies have been performed using *A. niger* and *T. reesei* [20-24], very few were based on sugarcane culm and/or bagasse [25,26]. In the present work, these two fungal species were grown on sugarcane biomass with two levels of cell wall complexity: culm *in natura* and pretreated bagasse. The production of enzymes related to biomass degradation was monitored using secretome analyses after 6, 12 and 24 hours. Concomitantly, we analyzed the sugars released in the supernatant. Our experiments demonstrate that both species degrade biomass after 6 hours post-inoculation, but comparative secretome analysis of *A. niger* and *T. reesei* revealed that it can occur through different mechanisms. This study provides a better understanding of the saccharification process, and it can be used as a basis for the production of optimized enzymatic cocktails.

## **Materials and methods**

### **Fungi strains and media**

The species used in this work were the filamentous fungi *Trichoderma reesei* RUT-C30 and *Aspergillus niger* N402. Both strains were maintained on potato dextrose agar (PDA) at 29°C and 30°C, respectively. The basic culture medium (BCM) (pH 5.5) was composed of 0.05% yeast extract (w/v), 50 mL/L salt solution (6 g/L NaNO<sub>3</sub>, 1.5 g/L KH<sub>2</sub>PO<sub>4</sub>, 0.5 g/L KCl

and 0.5 g/L MgSO<sub>4</sub>), 200 µL/L trace elements (10 g/L ethylenediaminetetraacetic acid, 4.4 g/L ZnSO<sub>4</sub>·7 H<sub>2</sub>O, 1.0 g/L MnCl<sub>2</sub>·4H<sub>2</sub>O, 0.32 g/L CoCl<sub>2</sub>·6H<sub>2</sub>O, 0.315 g/L CuSO<sub>4</sub>·5H<sub>2</sub>O, 0.22 g/L (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>·4H<sub>2</sub>O), 1.47 g/L CaCl<sub>2</sub>·2H<sub>2</sub>O and 1 g/L FeSO<sub>4</sub>·7H<sub>2</sub>O) and a predetermined concentration of carbon source, according to our experimental conditions (see below).

The exploded bagasse was prepared as described by Souza *et al.* [25]. Briefly, sugarcane bagasse *in natura* was treated with 14 kg/cm<sup>2</sup> water steam, washed exhaustively with distilled water until reducing sugars were not detected by dinitrosalicylic acid (DNS) [27] and dried at 40°C for several days. The culm was ground into particles with a 2-mm diameter, washed six times with 25 mL of 80 % (v/v) ethanol at 80 °C for 20 min, washed with distilled water to remove ethanol, and dried. After drying, both sugarcane exploded bagasse (SEB) and sugarcane culm (SC) were sifted on a 600-µm industrial sieve.

## Substrate-based induction conditions

*T. reesei* and *A. niger* spores were harvested from fresh potato dextrose agar (PDA) plates by adding 1 mL of sterile distilled water. The spore suspensions were inoculated in triplicate to a final concentration of  $1 \times 10^6$  spores per 30 mL of BCM (pH5.5) in a 250-mL Erlenmeyer flask containing 1% fructose (w/v) as the sole carbon source. *T. reesei* and *A. niger* spores were grown at 29°C and 30°C, respectively, for 24 hours (*A. niger*) on a rotary shaker with agitation at 200 rpm. *T. reesei* was also grown on a rotary shaker with agitation of 200 rpm, but it was grown for 48 hours to achieve an initial mycelial mass similar to that of *A. niger*. After, mycelia were removed by filtration through Whatman grade 1 filters (GE Healthcare), and they were washed with sterile water and transferred to fresh BCM media (without yeast extract) containing 0.5% of SEB or SC (w/v) as the sole carbon source for 6, 12

or 24 hours. *T. reesei* cultures were grown in a controlled environmental growth chamber under constant illumination with white light.

The mycelia and biomass used as carbon sources were harvested by filtration, washed thoroughly with sterile water and quickly frozen in liquid nitrogen for further fungal growth analyses. The supernatant was stored at -20°C for enzymatic, soluble supernatant sugar and mass spectrometry analyses.

## **Fungal growth**

Nitrogen content, an indirect measure of fungal growth, was measured based on the Pregl-Dumas' classical method [28]. The mycelial mass of *A. niger* and *T. reesei* grown on BCM with bagasse or culm for 6, 12 or 24 hours was rinsed with distilled water to remove traces of medium, and it was dried at 80°C for 4 h. The sample was macerated, and 2 mg (weighed with a digital electronic balance) was burned at approximately 975 °C in the presence of pure oxygen. The process released nitrogen, carbon dioxide and water, which were passed through special columns that absorbed the carbon dioxide and water. A column carrying a thermal conductivity detector separated the nitrogen from any other residue, and the resulting nitrogen content was measured. The instrument (PerkinElmer, model 2400, series II) was previously calibrated by analyzing a pure standard of known nitrogen, and the amount of nitrogen in each sample was given as a percentage in relation to the initial mass.

## **Monosaccharide Analyses in Culture Supernatant**

Monosaccharide analysis was performed on the supernatant. Each sample (1.8 mL) was completely dried using a Refrigerated CentriVap Concentrator (LABCONCO),

resuspended in 500  $\mu$ L of sterile deionized water and filtered through a 0.45- $\mu$ m pore size, 13-mm diameter (Durapore, Millex). The samples were subsequently analyzed by HPAEC-PAD on a CarboPac PA-1 column (DX-500 system, Dionex®). The elution of carbohydrates occurred in a gradient mixture of water and 200 mM sodium hydroxide at a flow rate of 0.8 mL/min for 50 min. Sugars were identified and quantified by comparing the retention times and ratios of sample peak area to internal standard peak area in relation to ratios determined for external standards using a Chromeleon 6.8 Chromatography Data System software.

## **Supernatant preparation and SDS-PAGE analysis**

To analyze the secretome profiles of *T. reesei* and *A. niger*, triplicate supernatants (~90 mL) containing enzymes from each time-point were pooled and clarified by filtration through a 0.22- $\mu$ m filter (Hydrophilic Millex, Millipore). The clarified supernatant was concentrated using a 3-kDa membrane (Vivaspin 20, GE HealthCare) to a final volume of 200  $\mu$ L, and 20  $\mu$ L was separated by 10% SDS-PAGE (110 V, 90 min). Three biological replicates were performed for the secretome experiments. The proteins were visualized by staining with 0.1 % Coomassie Brilliant Blue R250 (w/v), followed by destaining with 45% methanol and 10% acetic acid solution (v/v). All bands from triplicate SDS-PAGE gels were manually excised, reduced, alkylated and digested in gel with trypsin-modified sequencing-grade reagents (Promega), according to a previously described method [29].

## Mass spectrometry and protein identification

Peptides were resuspended in 0.1% formic acid (v/v), and an aliquot (4.5  $\mu$ L) was analyzed on an ETD-enabled LTQ Velos Orbitrap Mass Spectrometer (Thermo Fisher Scientific) coupled to a nanoflow liquid chromatography column (LC-MS/MS) via an EASY-nLC System (Proxeon Biosystem) through a Proxeon nanoelectrospray ion source. Peptides were separated by a 2-90% acetonitrile gradient in 0.1% formic acid using an analytical column PicoFrit Column (20 cm x ID75  $\mu$ m, 5- $\mu$ m particle size, New Objective), with a flow of 300 nL/min over 27 min. The nanoelectrospray voltage was set to 2.2 kV, and the source temperature was 275°C. All instrument methods for the LTQ Orbitrap Velos were set up in the data-dependent analysis (DDA) mode. Full scan MS spectra (m/z 300-1,600) were acquired in the Orbitrap analyzer after accumulation to a target value of  $1e^6$ . The resolution in the Orbitrap was set to  $r = 60,000$ . The 20 most intense peptide ions with charge states  $\geq 2$  were sequentially isolated to a target value of 5,000 and fragmented in the linear ion trap by low-energy CID (normalized collision energy of 35%). The signal threshold for triggering an MS/MS event was set to 1,000 counts. Dynamic exclusion was enabled with an exclusion size list of 500 and exclusion duration of 60 s. The activation Q-value was 0.25 and the activation time was 10 ms.

Data were acquired using the Xcalibur software package, and the raw data files were converted to a peak list format (mgf), without summing the scans, using Mascot Distiller v.2.3.2.0 (Matrix Science Ltd.). The database search was performed against the *Trichoderma* (13,808 proteins) and *Aspergillus niger* (36,414 proteins) from the NCBI database using the

Mascot v2.3.02 engine (Matrix Science Ltd.), with carbamidomethylation as a fixed modification, oxidation of methionine as a variable modification and one trypsin missed cleavage. The precursor mass tolerance was set to 10 ppm, and the fragment mass tolerance was set to 0.1 Da. For protein identification, the resulting search data were analyzed in Scaffold 3.5.1 (Proteome Software). The defined parameters were: minimum protein probability of 80%, minimum peptide probability of 90% and unique different minimum peptide of 1. We accepted proteins with up to 10% FDR protein and 5% FDR peptide. To access the signal peptide presence, we used the signal peptide prediction program, SignalP version 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>). The CAZy database (<http://www.cazy.org/>) was used to classify the identified proteins according to their families.

## **Enzymatic assays**

Enzymatic activity was determined from the amount of reducing sugar liberated from different polysaccharide substrates by the 3,5-dinitrosalicylic acid (DNS) method [27] using glucose as standard. First, 30 mL of supernatant from the samples induced with SEB or SC for 24 hours was concentrated using a centrifugal concentrator (Vivaspin 20, 10 kDA, GE HealthCare) to a final volume of 5 mL. The supernatant activity was assayed using xylan from beechwood,  $\beta$ -glucan, debranched arabinan from sugar beet, xyloglucan from tamarind and carboxymethylcellulose (CMC) (purchased from Sigma–Aldrich and Megazyme) as substrates at a 0.5% final concentration. Briefly, activity was measured using 50  $\mu$ L of the substrate solution, and the supernatant was diluted in 100 mM sodium acetate buffer (pH 5.5) to achieve a final volume of 100  $\mu$ L. The reaction was incubated at 40°C for 5 min for xylan and  $\beta$ -glucan and for 10, 70 and 180 min, for xyloglucan, CMC and debranched arabinan,

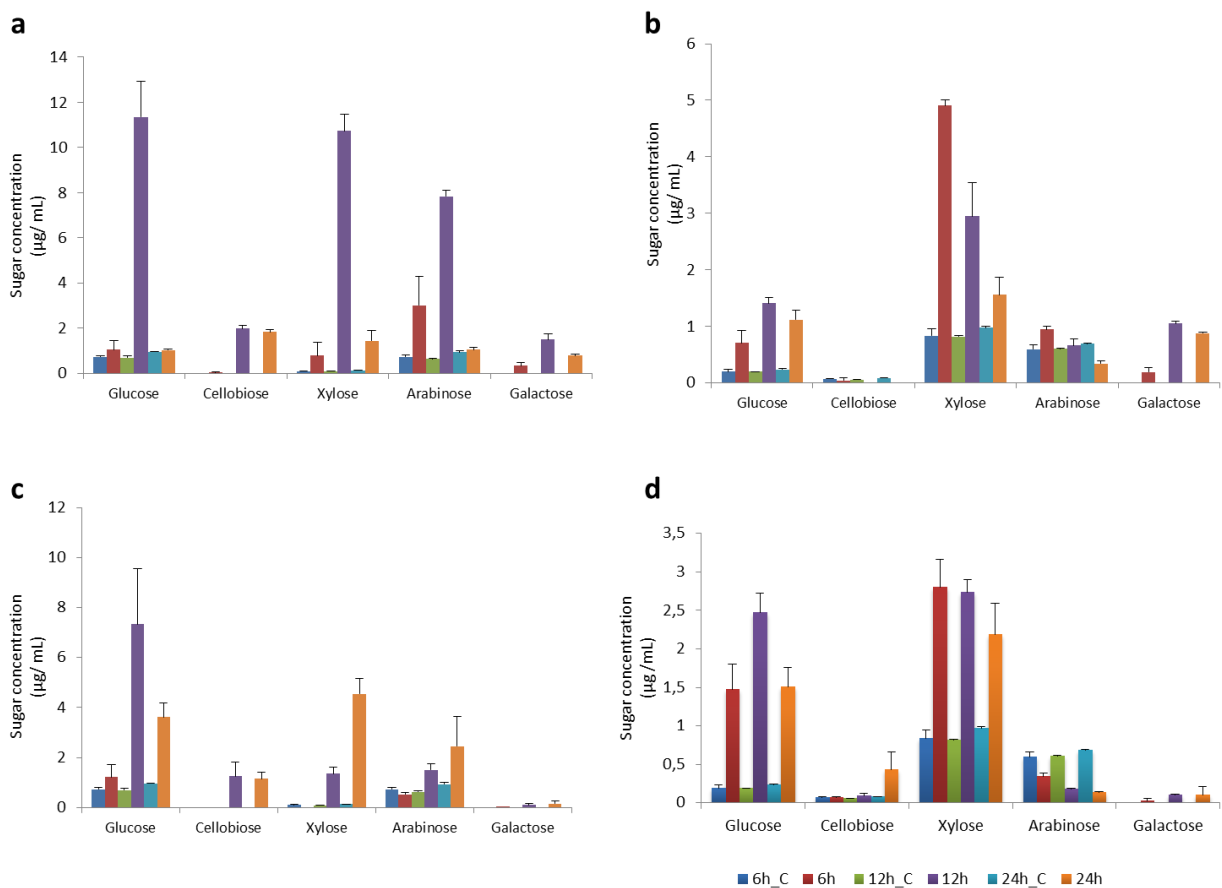
respectively. The reaction was stopped by adding 100  $\mu$ L of DNS. One unit (U) of enzymatic activity was defined as the amount of enzyme required to release 1  $\mu$ mol of reducing sugar per minute.

## Results and discussion

### Saccharification of sugarcane biomass

To identify when the degradation of sugarcane exploded bagasse (SEB) and culm (SC) were initiated, we analyzed the sugars that were released into the *A. niger* and *T. reesei* culture supernatants using high-performance liquid chromatography (HPAEC-PAD), over a time course of 6, 12 and 24 hours post-inoculation. The supernatant from each time point was also collected from control samples that were not inoculated with fungi. Fig. 1 shows the changes in the proportions of glucose, arabinose, xylose, galactose and cellobiose over the course of the experiment. We observed bagasse degradation after 6 hours post-inoculation since the total concentrations of glucose and xylose were higher than in the control samples (Fig. 1a, 1c). In culm, this feature was more noticeable after 12 hours (Fig. 1b, d).



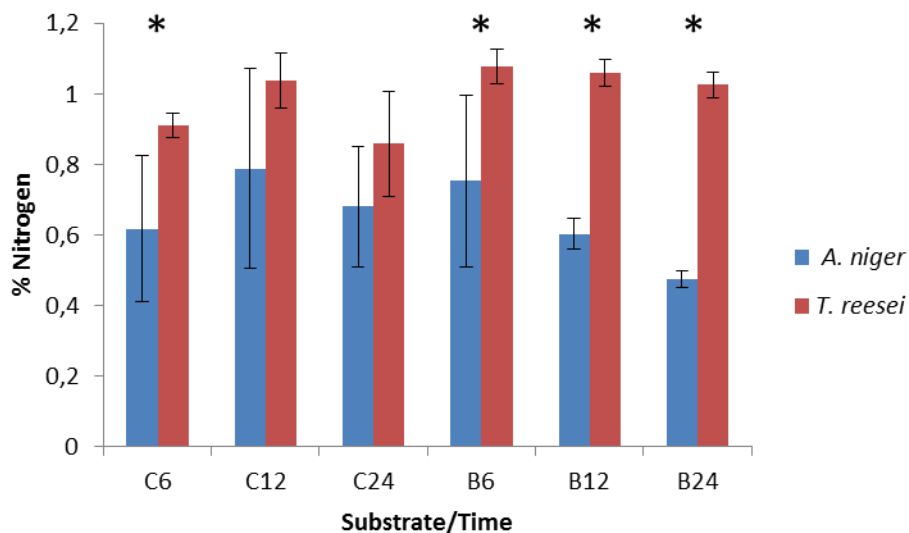


**Fig. 1. Degradation of sugarcane biomass and the sugars released by *A. niger* and *T. reesei*.** Free sugars in the supernatant after the transfer of mycelia to culm/bagasse media at 6, 12 and 24 hours (h). Each bar represents the mean and the standard deviation of values from three independent experiments. The samples marked C are the control samples that were not inoculated with fungi. a) *A. niger* grown on 0.5% culm; b) *A. niger* grown on 0.5% bagasse; c) *T. reesei* grown on 0.5% culm; and d) *T. reesei* grown on 0.5% bagasse.

The changes in the proportion of sugars (Fig. 1) led us to conclude that both cellulose and hemicellulose were being degraded. We observed only galactose in the samples with fungi that used bagasse and culm as carbon source, suggesting that the degradation of branched

hemicelluloses (mainly composed of xylose and galactose) started within 6 hours. It is also likely that the fungi were consuming sugars derived from arabinoxylan, since we detected changes in arabinose concentration. Together with pectin and  $\beta$ -glucan, arabinoxylans form the most water-soluble and accessible part of the cell wall in sugarcane [4]. Therefore, these sugars are readily released by the fungal species as soon as they come into contact with the substrate due the actions of multiple enzymes, such as  $\alpha$ -arabinofuranosidases (GH3, GH43, GH51, GH54 and GH62), which remove the residues of arabinose, and  $\beta$ -1,4-endoxylanase (GH10, GH11, GH30) and  $\beta$ -1,4-xylosidase (GH3, GH43), which hydrolyze xylose from the xylan backbone [30].

Despite the higher recalcitrance of the culm, we also noticed that the samples from fungi grown on this substrate presented the highest amount of sugars in the supernatant (Fig. 1b, 1d). These results suggest that both *T. reesei* and *A. niger* secreted more enzymes using culm as a carbon source compared to bagasse, since there were no differences in growth between the two substrates (indirectly measured by nitrogen content, Fig. 2). Because both fungi consume sugars, these values do not represent the real rate of sugar release; however, they suggest that both fungi released different amounts of monosaccharides and/or have different rates of sugar uptake.



**Fig. 2. Nitrogen content.** Total nitrogen content of fungi growing on sugarcane culm (C) and bagasse (B) for 6, 12 and 24 hours (h). Each bar represents the mean and the standard deviation of values from three independent experiments. The asterisk (\*) indicates samples with a significantly different % of nitrogen between fungal species ( $p < 0.05$ ) by unpaired t-test, which suggest that *T. reesei* has a higher growth at these substrates/time points than *A. niger*. There is no difference in nitrogen content among the substrates/time points examined in both species.

## CAZymes profiling of *T. reesei* and *A. niger* secretomes

Because the sugars released are dependent on secreted hydrolases, we investigated the *T. reesei* and *A. niger* secretomes. These hydrolysis systems appear depend on the microorganism and carbon source used [23,31,32]. To date, few studies have utilized lignocellulosic biomass as a major carbon source, especially with sugarcane biomass [25,26]. Our data showed that *A. niger* secreted a large number of CAZymes at all time points/carbon sources. In total, we identified 45 different CAZymes in *T. reesei* (Table 1) and 89 in *A. niger*

(Table 2) (Glycosyl Transferases, GTs, were not included). Table S1 lists all these enzymes and their respective molecular weights, numbers of identified peptides, peptide sequences, mass-to-charge ratios ( $m/z$ ), numbers of unique peptides, and presence/absence of a signal peptide.

**Table 1** – CAZymes families detected in the secretome of *Trichoderma reesei*.

CAZy Family	Predicted Protein	JGI Protein ID	Substrate/Time point	Possible Polysaccharide Substrate/Classification
GH17	Candidate glucan 1,3- $\beta$ -glucosidase	24326	B24h, C6h, C24h	1,3-1,4- $\beta$ -Glucan
GH17	Candidate glucan endo-1,3- $\beta$ -glucosidase	110434	C12h, C24h	1,3-1,4- $\beta$ -Glucan
GH55	Exo-1,3-b-glucanase	25104	B12h, C6h, C12h, C24h	1,3-1,4- $\beta$ -Glucan
GH55	$\beta$ -1,3-glucanase	93142	B6h, B12h, B24h, C6h, C12h, C24h	1,3-1,4- $\beta$ -Glucan
GH1	$\beta$ -glucosidase CEL1A, bgl2	127115	B12h, B24h	Cellulose
GH3	$\beta$ -glucosidase, CEL3B	25095	B6h, B12h, B24h, C24h	Cellulose
GH3	$\beta$ -glucosidase, bgl3i	109567	B24h, C6h, C24h	Cellulose
GH3	$\beta$ -glucosidase, bgl1, cel3a	136547	B24h, C24h	Cellulose
GH5	Endoglucanase CEL5A, egl2	72489	C24h	Cellulose
GH6	Cellobiohydrolase CEL6A, cbh2	122470	B24h, C12h, C24h	Cellulose
GH7	Endoglucanase CEL7B, egl1	5304	C24h	Cellulose
GH7	Cellobiohydrolase CEL7A, cbh1	125125	B6h, B12h, B24h, C6h, C12h, C24h	Cellulose
AA9 (GH61)	Copper-dependent monooxygenase, CEL61a	139633	B24h, C24h	Cellulose
GH5	Endo- $\beta$ -1,4-mannosidase, man5a	122377	C12h, C24h	Mannan
GH79	Candidate $\beta$ -glucuronidase	69609	C12h	Pectin
GH43	Candidate $\beta$ -xylosidase/arabinosidase	133200	C6h, C12h, C24h	Xylan/Arabinoxylan
GH3	B-xylosidase, bxl1	140746	B6h, B12h, B24h, C24h	Xylan/Arabinoxylan
GH11	Xylanase, xyn2	124931	B12h, B24h, C6h	Xylan/Arabinoxylan
GH30	Endo- $\beta$ -1,4-xylanase, xyn4	90847	B6h, B12h, B24h, C6h	Xylan/Arabinoxylan
CE5	Acetyl xylan esterase, axe2	94461	C6h, C12h, C24h	Xylan/Arabinoxylan
CE5	Acetyl xylan esterase, axe1	139631	B24h	Xylan/Arabinoxylan
GH54	$\alpha$ -L-arabinofuranosidase, abf3	72252	C24h	Xylan/Arabinoxylan
GH74	Xyloglucanase, CEL74A	111943	B24h, C12h, C24h	Xyloglucan
GH16	Cell wall glucanase	96805	B6h, B12h, B24h, C6h, C12h, C24h	Carbohydrate transport and metabolism
GH27	$\alpha$ -D-galactosidase	75420	B12h	Carbohydrate transport and metabolism

<b>CAZy Family</b>	<b>Predicted Protein</b>	<b>JGI Protein ID</b>	<b>Substrate/Time point</b>	<b>Possible Polysaccharide Substrate/Classification</b>
GH3	Candidate $\beta$ -N-acetylglucosaminidase	12475	B6h	Carbohydrate transport and metabolism
GH30	Candidate endo- $\beta$ -1,4-xylanase	93498	B24h	Carbohydrate transport and metabolism
GH31	Candidate $\alpha$ -glucosidase	104546	B6h, B24h, C12h, C24h	Carbohydrate transport and metabolism
GH37	Candidate $\alpha,\alpha$ -trehalase	102372	C6h, C12h, C24h	Carbohydrate transport and metabolism
GH65	Candidate $\alpha,\alpha$ -trehalase	139320	B24h, C12h, C24h	Carbohydrate transport and metabolism
GH72	Candidate $\beta$ -1,3-glucanosyltransferase	98936	C6h, C12h, C24h	Carbohydrate transport and metabolism
GT20	Bifunctional trehalose-6-phosphate synthase	72420	B6h, B12h	Carbohydrate transport and metabolism
GT20	Bifunctional trehalose-6-phosphate synthase	67350	B12h	Carbohydrate transport and metabolism
GT31	Distantly related to $\beta$ -glycosyltransferases	101599	B12h	Carbohydrate transport and metabolism
GT35	Glycosyl transferase	23636	B6h, B12h, B24h, C6h	Carbohydrate transport and metabolism
GH16	Candidate glucanosyltransferase	66752	B6h, B12h, B24h, C6h, C12h, C24h	Cell wall biosynthesis and morphogenesis
GH72	Candidate $\beta$ -1,3-glucanosyltransferase	9074	B6h, B12h, C6h, C12h, C24h	Cell wall biosynthesis and morphogenesis
GH72	Candidate $\beta$ -1,3-glucanosyltransglycosylase	124639	B6h, B12h, B24h, C6h, C12h, C24h	Cell wall biosynthesis and morphogenesis
GT2	Glycosyl transferase	114628	B6h, B12h, B24h, C6h, C12h, C24h	Cell wall biosynthesis and morphogenesis
GT48	Candidate $\beta$ -1,3-glucan synthase	139875	C6h, C12h	Cell wall biosynthesis and morphogenesis
GH18	Chitinase (chi18-5, chi46)	124526	C12h, C24h	Chitin
GH18	Chitinase (chi18-8)	130024	B12h	Chitin
GH20	N-acetyl- $\beta$ -D-glucosaminidase (nag1)	99285	C12h, C24h	Chitin
GH13	Candidate glycogen debranching enzyme	73564	B6h, B12h, B24h	Starch
GH15	Glucoamylase (gla)	70185	C6h, C12h	Starch

**Table 2** - CAZymes families detected in the secretome of *Aspergillus niger*.

CAZy Family	Predicted protein	JGI Protein ID	Substrate/Time point	Possible Polysaccharide Substrate/Classification
GH17	Glucan endo-1,3- $\beta$ -glucosidase( <i>eglc</i> )	158521	B6h, B12h, B24h, C6h, C12h, C24h	1,3-1,4- $\beta$ -Glucan
GH55	Exo- $\beta$ -1,3-glucanase ( <i>bxga</i> , <i>exgo</i> )	156270	B12h, B24h, C6h, C12h, C24h	1,3-1,4- $\beta$ -Glucan
GH55	$\beta$ -1,3-exoglucanase	157838	B12h	1,3-1,4- $\beta$ -Glucan
GH3	$\beta$ -D-glucoside glucohydrolase M	163273	B6h, B12h, B24h, C12h	Cellulose
GH3	$\beta$ -glucosidase 2	163842	B12h, B24h	Cellulose
GH3	$\beta$ -glucosidase	165962	C24h	Cellulose
GH3	$\beta$ -glucosidase	168801	B6h, B12h, B24h, C6h, C12h	Cellulose
GH5	Endo- $\beta$ -1,4-glucanase A ( <i>egla</i> )	156195	B6h, B12h, B24h, C12h, C24h	Cellulose
GH5	Endo- $\beta$ -1,4-glucanase B ( <i>eng1</i> )	161114	B6h, B12h, B24h, C6h, C12h, C24h	Cellulose
GH5	Endoglucanase B ( <i>eglb</i> )	167967	B6h, B12h, C6h, C12h	Cellulose
GH5	Glucan 1,3- $\beta$ -glucosidase A ( <i>exga</i> )	168853	B6h, B12h	Cellulose
GH6	Exocellobiohydrolase	161440	B12h, B24h, C6h, C12h	Cellulose
GH6	$\beta$ -glucancellobiohydrolase C	164557	B6h, B12h, B24h, C6h	Cellulose
GH7	Cellobiohydrolase B ( <i>cbhb</i> )	156194	B6h, B12h, B24h, C6h, C12h, C24h	Cellulose
GH7	Cellobiohydrolase A ( <i>cbha</i> )	161153	B6h, B12h, B24h, C6h, C12h, C24h	Cellulose
GH30	Glucan endo-1,6- $\beta$ -glucosidase	158067	B6h, B12h, B24h, C6h, C12h, C24h	Cellulose
AA9 (GH61)	Endoglucanase IV	161785	B6h, B12h, C6h, C12h	Cellulose
AA9 (GH61)	Endoglucanase 4	166052	B24h	Cellulose
AA9 (GH61)	Putative endoglucanase IV	166976	B6h, B12h, B24h	Cellulose
CE8	Pectinesterase	157769	C6h, C12h, C24h	Pectin
CE8	Pectin methylesterase A ( <i>pmea</i> )	158617	B6h, C6h, C12h	Pectin
CE8	Pectin methylesterase A ( <i>pmea</i> )	159650	B6h, B12h, C6h, C12h, C24h	Pectin
CE12	Putative rhamnogalacturonan acetyl esterase	159617	C6h, C12h, C24h	Pectin
CE12	Rhamnogalacturonan acetyl esterase	162676	C6h, C12h	Pectin
CE16	Putative pectin acetylerase	156782	C6h	Pectin

CAZy Family	Predicted protein	JGI Protein ID	Substrate/Time point	Possible Polysaccharide Substrate/Classification
GH5	Endo- $\beta$ -1,6-galactanase	158118	B6h, B12h, B24h, C6h, C12h	Pectin
GH28	Endopolygalacturonase-1	156180	C12h	Pectin
GH28	Endopolygalacturonase B ( <i>pgab</i> )	157015	B6h, B12h, B24h, C6h	Pectin
GH28	Exopolygalacturonase ( <i>pgxb</i> )	158660	C6h, C12h, C24h	Pectin
GH28	Endo-xylogalacturonan hydrolase A ( <i>xgha</i> )	159651	B12h, B24h, C6h, C12h, C24h	Pectin
GH28	Endo-polygalacturonase D ( <i>pgad</i> )	162788	C6h, C12h	Pectin
GH28	Exo-xylogalacturonan hydrolase ( <i>pgxa</i> )	163648	C12h	Pectin
GH28	Rhamnogalacturonase ( <i>rhga</i> )	164433	C12h	Pectin
GH28	Exopolygalacturonase X ( <i>pgax</i> )	165048	C6h	Pectin
GH28	Rhamnogalacturonase B ( <i>rhgb</i> )	166203	C6h, C12h	Pectin
GH28	Rhamnogalacturonan $\alpha$ -galacturonohydrolase	168924	C12h	Pectin
GH53	Arabinogalactan endo-1,4- $\beta$ -galactosidas	169030	B24h, C6h	Pectin
PL1	Pectin lyase A ( <i>pela</i> )	166220	B12h, C6h, C12h	Pectin
CE1	Feruloyl esterase A ( <i>faea</i> )	162483	B6h, B12h, B24h, C6h, C12h, C24h	Phenylpropanoids
CE1	Feruloyl esterase C ( <i>faec</i> )	164585	B6h, B12h	Phenylpropanoids
CE1	Feruloyl esterase ( <i>faeb</i> )	165335	B6h, B12h, B24h, C6h, C12h, C24h	Phenylpropanoids
CE1	Acetyl xylan esterase ( <i>axea</i> , <i>acea</i> )	164821	B24h, C6h, C12h	Xylan/Arabinoxylan
CE16	Acetylerase	161113	B6h, B12h, C6h, C12h	Xylan/Arabinoxylan
GH3	Exo-1,4- $\beta$ -xylosidase ( <i>x1nd</i> )	156034	B6h, B12h, B24h, C6h, C12h	Xylan/Arabinoxylan
GH3	Bifunctional xylosidase-arabinosidase	168244	B24h	Xylan/Arabinoxylan
GH10	Endo-1,4- $\beta$ -xylanase C	158107	B6h, B12h, B24h, C12h	Xylan/Arabinoxylan
GH11	Xylanase 2 / B ( <i>xynb</i> , <i>xlnb</i> )	155137	B6h, B12h, B24h, C6h, C12h, C24h	Xylan/Arabinoxylan
GH11	Endo-1,4- $\beta$ -xylanase B	166974	C6h	Xylan/Arabinoxylan
GH43	Arabinan endo-1,5- $\alpha$ -L-arabinosidase C	157571	B6h, B12h, B24h, C6h, C12h, C24h	Xylan/Arabinoxylan
GH43	Xylan $\beta$ -xylosidase	161454	B12h, B24h	Xylan/Arabinoxylan
GH43	Glycosyl hydrolase family 43 protein	162327	B12h, B24h, C6h, C12h, C24h	Xylan/Arabinoxylan
GH43	Arabinan endo- $\alpha$ -1,5-L-arabinosidase A	162583	B6h, C6h, C24h	Xylan/Arabinoxylan
GH43	Endo-arabinase	166877	B12h	Xylan/Arabinoxylan



<b>CAZy Family</b>	<b>Predicted protein</b>	<b>JGI Protein ID</b>	<b>Substrate/Time point</b>	<b>Possible Polysaccharide Substrate/Classification</b>
GH51	$\alpha$ -L-arabinofuranosidase A (abfa, exoa)	155097	C6h, C12h, C24h	Xylan/Arabinoxylan
GH51	$\alpha$ -L-arabinofuranosidase E	162554	B6h, B24h, C12h	Xylan/Arabinoxylan
GH54	$\alpha$ -L-arabinofuranosidase B	166753	B6h, B12h, B24h, C6h, C12h, C24h	Xylan/Arabinoxylan
GH62	$\alpha$ -L-arabinofuranosidase (axha)	158109	B6h, B12h, B24h, C6h, C12h, C24h	Xylan/Arabinoxylan
GH67	$\alpha$ -glucuronidase A	166362	B6h, B12h, B24h	Xylan/Arabinoxylan
GH35	$\beta$ -galactosidase A (laca)	156240	B6h, B12h	Xyloglucan
GH12	Xyloglucan- endo- $\beta$ -1,4-glucanase	155384	B6h, C6h, C12h	Xyloglucan
GH12	Endo- $\beta$ -1,4-glucanase	158544	B6h, B12h, B24h, C12h, C24h	Xyloglucan
GH12	Endo- $\beta$ -1,4-glucanase (egla)	166061	B6h, B12h, B24h, C6h, C12h, C24h	Xyloglucan
GH74	Xyloglucanase	155242	B12h, B24h	Xyloglucan
GH32	Invertase (suca, suc1)	162354	B24h, C6h	Carbohydrate transport and metabolism
GH32	Exo-inulinase E (inue, inu1)	165128	C6h, C12h, C24h	Carbohydrate transport and metabolism
GH65	$\alpha$ -trehalase	155210	B6h, B12h, B24h	Carbohydrate transport and metabolism
GH95	Glycosyl hydrolase	167353	B24h	Carbohydrate transport and metabolism
GH27	$\alpha$ -galactosidase II (aglb)/ melibiase	157631	B12h, C12h	Carbohydrate transport and metabolism
GH27	$\alpha$ -galactosidase A (ag1A;agla)	159990	B12h, B24h, C12h	Carbohydrate transport and metabolism
GH27	$\alpha$ -galactosidase D (agld)	165965	B6h, C12	Carbohydrate transport and metabolism
GH16	$\beta$ -glucanase	155502	B24h, C24h	Cell wall biosynthesis and morphogenesis
GH16	Glycosidase crf1	156136	B6h, B12h, B24h, C6h, C12h, C24h	Cell wall biosynthesis and morphogenesis
GH16	GPI-anchored glucanosyltransferase	160973	B6h, C6h	Cell wall biosynthesis and morphogenesis
GH16	$\beta$ -glucanase	163407	B6h, B12h, B24h, C6h, C12h, C24h	Cell wall biosynthesis and morphogenesis
GH17	1,3- $\beta$ -glucanosyltransferase (bgt1)	161620	B6h, B12h, B24h	Cell wall biosynthesis and morphogenesis
GH72	1,3- $\beta$ -glucanosyltransferase	156831	B12h, C12h	Cell wall biosynthesis and morphogenesis
GH72	1,3- $\beta$ -glucanosyltransferase	161995	C6h	Cell wall biosynthesis and morphogenesis
GH72	1,3- $\beta$ -glucanosyltransferase	162537	B6h, B24h, C6h, C12h, C24h	Cell wall biosynthesis and morphogenesis
GH72	1,3- $\beta$ -glucanosyltransferase gell	163189	B6h, B12h	Cell wall biosynthesis and morphogenesis

<b>CAZy Family</b>	<b>Predicted protein</b>	<b>JGI Protein ID</b>	<b>Substrate/Time point</b>	<b>Possible Polysaccharide Substrate/Classification</b>
GH81	Glucan endo-1,3- $\beta$ -D-glucosidase	155359	B12h, B24h	Cell wall biosynthesis and morphogenesis
GH13	Cell-wall 4- $\alpha$ -glucanotransferase (agta)	162772	B6h, B12h	Starch
GH13	Acid $\alpha$ -amylase (aama)	163584	B6h, B12h, B24h, C6h, C12h, C24h	Starch
GH15	Glucoamylase (glaa)	158641	B6h, B12h, B24h, C6h, C12h, C24h	Starch
GH18	Class V endochitinase (chib)	157223	B12h, B24h	Chitin
GH18	Exo-chitinase (cfci)	157878	B12h	Chitin
GH20	N-Acetyl- $\beta$ -glucosaminidase (nag1)	162684	B24h	Chitin
GH5	$\beta$ -mannanase (mana)	159852	B6h, B12h, B24h	Mannan
GH47	$\alpha$ -1,2-mannosidase	156279	B6h, B12h, B24h	Mannan
GH92	Glycosyl hydrolase	166207	B12h, B24h	Mannan

Compared to previous studies, we identified a higher number of CAZymes in *A. niger*. Adav *et al.* [20] identified 30 enzymes produced while culturing *A. niger* in a bioflo fermenter containing minimal medium and glucose. Oliveira *et al.* [33] identified 40 hydrolytic enzymes in the secretome of *A. niger* grown on a bioreactor containing minimal medium added of D-xylose (an inducer of cellulases and hemicellulases) or D-maltose as the sole carbon sources. However, in those reports, the inducers were simple carbon sources, and not complex carbon sources such as sugarcane bagasse and culm. An exception is the work from Souza *et al.* [25] that used SEB as carbon source, but they identified only 17 proteins using a much less refined proteomics method.

Interestingly, only one enzyme was reported in common to all these studies, the acid  $\alpha$ -amylase (AamA). This enzyme belongs to the GH13 family and it cleaves internal  $\alpha$ -(1,4)-glycosidic bonds in starch and glycogen [34]. However, we identified the most important enzymes related to biomass degradation. To hydrolyze the cellulose chains into monomers, the main chain must be cleaved internally, and this event is performed by endoglucanases (GH5, GH12). Likewise, the release cellobiose occurs via the action of exoglucanases (GH6, GH7), and it is subsequently converted into glucose by  $\beta$ -glucosidases (GH3) [35]. The hemicellulose fraction is formed by arabinoxylan,  $\beta$ -glucan and xyloglucan, and due to this variety of substrates, the enzymatic mixture required to break it down is more diverse, including endoxylanases (GH10, GH11),  $\beta$ -xylosidase (GH3), arabinofuranosidases (GH43, GH51, GH54), galactosidases (GH35) and others [22]. In the *A. niger* secretome, all of these enzymes were present at the earliest time point examined in this study (6 h).

We observed some differences for *T. reesei* grown on other carbon sources since the number of *T. reesei* CAZymes was higher than that reported by Herpoël-Gimbert *et al.* [22]; they identified 22 hydrolytic enzymes using lactose and xylose as carbon sources. However,

Adav *et al.* [24] identified over 90 CAZymes using a quantitative proteomic approach, the iTRAQ system. Moreover, a recent study using untreated sugarcane bagasse investigated the secretome of two *Trichoderma* strains using solid-state fermentation [26]. In this study, they identified 39 GHs, and other proteins that play an important role in biomass degradation (for example, swollenin). These differences could be due either to the carbon source, culture conditions or experimental design. However, in contrast to *A. niger*, the most important proteins found in this study were also present in the afore mentioned reports, such as cellobiohydrolases (Cbh1/Cbh2), endoglucanases (Eg1, Eg2 and Eg4) and  $\beta$ -glucosidases (Bgl1/Bgl2). We also identified hemicellulases involved in the cleavage of the main chain of xylan and xyloglucan (endoxylanase and xyloglucanase) and the side chain of hemicelluloses, including  $\beta$ -xylosidase,  $\alpha$ -arabinofuranosidase and  $\alpha$ -galactosidase, and acetylxyylan esterase.

## **Differences in the enzyme secretion in distinct sugarcane biomass**

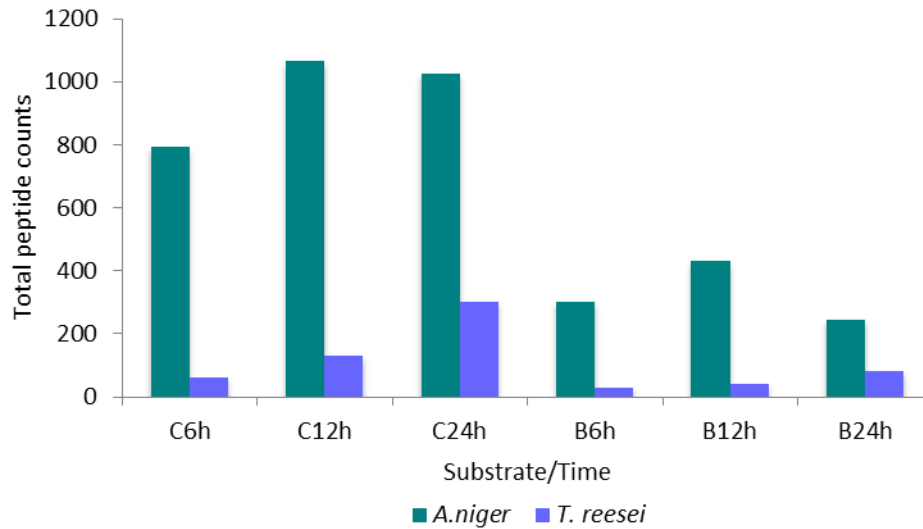
An additional goal of this study was to better understand the differences in enzyme patterns in distinct biomass types. Thus, in addition to bagasse, culm *in natura* was used as a sole carbon source. Fig. 3 shows a heat-map representing the number of enzymes identified for each class of enzymes (CAZymes classification was used) in each family of glycosyl hydrolase directly related to biomass degradation (see Table 1 and 2) in *A. niger* and *T. reesei* after 6, 12 and 24 hours growing on untreated sugarcane culm and pretreated bagasse.

CAZy Family/Putative Enzyme		<i>A. niger</i>						<i>T. reesei</i>						Occurrence					
		Culm			Bagasse			Culm			Bagasse			<i>A.niger</i>	<i>T.reesei</i>	UNIQUNESS			
		6h	12h	24h	6h	12h	24h	6h	12h	24h	6h	12h	24h						
<b>PHENYLPROPANOIDS</b>																			
CE1	Feruloyl esterase C (28.5KDa and 32.6KDa)	3	3	1	3	3	3	0	0	0	0	0	0	0	0	0	Y	N	AN
<b>SWOLLENIN</b>																			
SWOI	Swollenin (22KDa)	0	0	0	0	0	0	1	1	1	0	1	1	N	Y	TR			
<b>PECTINS</b>																			
CE8	Pectin methylesterase A/Pectinesterase (35.6KDa)	3	3	2	2	1	0	0	0	0	0	0	0	0	0	0	Y	N	AN
CE12	Rhamnogalacturonan acetyl esterase (27.7KDa)	2	2	1	0	0	0	0	0	0	0	0	0	0	0	0	Y	N	AN
CE16	Putative pectin acetyl esterase (35.1KDa - 39.1KDa)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	Y	N	AN
GH28	Endopolygalacturonase/Exoxylogalacturonan hydrolase/Rhamnogalacturonase (37 to 47.8KDa)	6	8	2	1	2	2	0	0	0	0	0	0	0	0	0	Y	N	AN
GH53	Arabinogalactan endo-1,4-beta-galactosidase A (38.7KDa)	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	Y	N	AN
PL1	Pectin lyase A (39.8KDa)	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	Y	N	AN
<b>1,3-1,4-β-GLUCAN</b>																			
GH17	Glucan endo-1,3-β-glucosidase (46.3KDa)	1	1	1	1	1	1	1	1	2	0	0	1	1	1	1	Y	Y	BOTH
GH55	Beta-1,3-exoglucanase (98.7KDa - 81KDa)	1	1	1	0	2	1	2	2	2	1	2	1	1	1	1	Y	Y	BOTH
<b>XYLAN / ARABINOXYLAN</b>																			
CE1	Acetyl xylan esterase	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	Y	N	AN
CE16	Acetyl esterase (35.1kDa - 39.1KDa)	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	Y	N	AN
GH10	Endo-1,4-β-xylanase (35.4kDa - 38.1KDa)	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	Y	N	AN
GH67	α-glucuronidase A (93.9KDa - 93.4KDa)	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	Y	N	AN
GH51	α-L-arabinofuranosidase A or E (67.9KDa)	1	2	1	1	0	1	0	0	0	0	0	0	0	0	0	Y	N	AN
GH62	α-L-arabinofuranosidase (axha) (35.8KDa - 34.8KDa)	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	Y	N	AN
CE5	Acetyl xylan esterase (26.7Kda - 30.5KDa)	0	0	0	0	0	0	1	1	1	0	0	1	0	0	1	N	Y	TR
GH30	Endo-β-1,4-xylanase (51.8KDa - 51.8)	0	0	0	0	0	0	1	0	0	1	1	2	0	0	0	N	Y	TR
GH11	Xylanase, xyn2 (23.5KDa - 21.5KDa)	2	1	1	1	1	1	1	0	0	0	1	1	0	0	0	Y	Y	BOTH
GH43	Xylan β-xylosidase/Arabinan endo-1,5-alpha-L-arabinosidase C (35.4KDa)	3	2	3	2	4	3	1	1	1	0	0	0	0	0	0	Y	Y	BOTH
GH54	α-L-arabinofuranosidase B (52.5KDa - 51.1KDa)	1	1	1	1	1	1	0	0	1	0	0	0	0	0	0	Y	Y	BOTH
GH3	β-xylosidase/Bifunctional xylosidase-arabinosidase/ (87.2KDa - 93.9KDa)	1	1	0	1	1	2	0	0	1	1	1	1	1	1	1	Y	Y	BOTH
<b>XYLOGLUCAN</b>																			
GH12	Xyloglucan-specific endo-β-1,4-glucanase (25.8KDa - 25.1KDa)	2	3	2	3	2	2	0	0	0	0	0	0	0	0	0	Y	N	AN
GH35	β-galactosidase A (109.2KDa - 111.3KDa)	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	Y	N	AN
GH74	Xyloglucanase (90.4KDa - 87.1KDa)	0	0	0	0	1	1	0	1	1	0	0	1	0	0	0	Y	Y	BOTH
<b>CELLULOSE</b>																			
GH1	β-glucosidase CEL1A, bg12 (48.3KDa - 52.2KDa)	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	N	Y	TR
GH3	β-glucosidase (87.2KDa - 93.9KDa)	1	2	1	2	3	3	1	0	3	1	1	3	0	0	0	Y	Y	BOTH
GH5	Endo-beta-1,4-glucanase A or B (36.5 KDa - 46.8KDa)	3	4	2	6	6	4	0	1	2	0	0	0	0	0	0	Y	Y	BOTH
GH6	Cellobiohydrolase cbh2 (42.1KDa - 49.6KDa)	2	1	0	1	2	2	0	1	1	0	0	1	0	0	0	Y	Y	BOTH
GH7	Cellobiohydrolase B (56.1KDa - 48.2KDa)	2	2	2	2	2	2	1	1	2	1	1	1	1	1	1	Y	Y	BOTH
AA9 (GH61)	Cu dependent monooxygenase	1	1	0	2	2	2	0	0	1	0	0	1	0	0	0	Y	Y	BOTH

**Fig. 3. Graphical representation of secreted CAZYmes.** Heat-map of the number of enzymes of each CAZY family secreted by *A. niger* and *T. reesei* after 6, 12 and 24 hours (h) growth on sugarcane culm and bagasse. This map includes only enzymes/proteins related to biomass degradation (Tables 1 and 2).

The pattern of enzyme production observed for *A. niger* was similar for culm and bagasse. However, analyzing the enzymes secreted exclusively when *A.niger* was grown on culm, we observed a range of pectinases (Fig. 3, Table 2), which have a variety of biotechnological and industrial applications, and several of them have been cloned, characterized, or have had their production optimized [12,36-38]. Although six pectinases were identified as being secreted using both carbon sources, ten pectinases were secreted exclusively when *A.niger* was grown on culm (Fig. 3, Table 2). It is likely that the fraction containing more soluble pectins was lost during the sugarcane crushing (probably because of the hot water treatment), rendering bagasse with proportionally less pectin compared to culm. Therefore, the induction of these enzymes is higher when the fungi are grown using culm as a carbon source. Curiously, although *T. reesei* secreted several enzymes when growing on culm and bagasse, none were related to pectin, even though its genome encodes several pectinases.

Despite the variety of enzymes produced by both fungi, a comparative analysis of the total number of peptides, which may be indicative of protein/enzyme abundance [39,40] (Fig. 4), suggested that *A. niger* and *T. reesei* secreted more enzymes when grown on culm. As mentioned previously, there were no differences in fungi growth between the two substrates (Fig. 2). Thus, the higher abundance of peptides produced by fungi growing on culm correlates with the greater amount of sugars in the supernatant (Fig. 1b, 1d). This result could be related to differences in the recalcitrance of substrates, which reflects the different levels of complexity of the cell walls, as culm did not receive any pretreatment. However, how fungi sense these differences remains unclear.



**Fig. 4. A comparison of the abundance of secreted enzymes by *A. niger* and *T. reesei*.** A semi-quantitative analysis of the amount of enzyme secreted by both fungi after 6, 12 and 24 hours growth on sugarcane culm (C) and bagasse (B).

### Comparison between *T. reesei* and *A. niger* secretomes

The secretome profiles of these two fungi were compared, and our data indicates they differed considerably (Fig. 3, Tables 1 and 2). As mentioned before, the assembly, modification and breakdown of complex carbohydrate and glycoconjugates are carried out by the CAZymes [41]. However glycosyl transferases and the carbohydrate esterases from the family CE10 genes are not involved in degradation of plant cell walls. In the *A. niger* supernatant, we detected 65 enzymes directly related to biomass degradation, approximately 22% of the total of carbohydrate-active enzymes encoded by its genome (292 proteins, excluding GTs), whereas *T. reesei* secreted 24 enzymes (approximately 10% of its 228

CAZymes, excluding GTs). GHs found exclusively in the *A. niger* secretome were related to hemicellulose degradation, such as GH10 (endo-1,4- $\beta$ -xylanase), GH51 and GH62 (both  $\alpha$ -L-arabinofuranosidases), and pectin, including endopolygalacturonase (GH28) and endo-1,4- $\beta$ -galactosidase (GH53) (Fig. 3). Important esterases were also present in *A. niger*, but not *T. reesei*, such as the acetyl xylan esterase (CE1) and acetylerase (CE16), that catalyzes the hydrolysis of acetyl groups from hemicellulases, and pectin methyl esterase (CE8), acetyl esterase (CE12), that have activity against pectin.

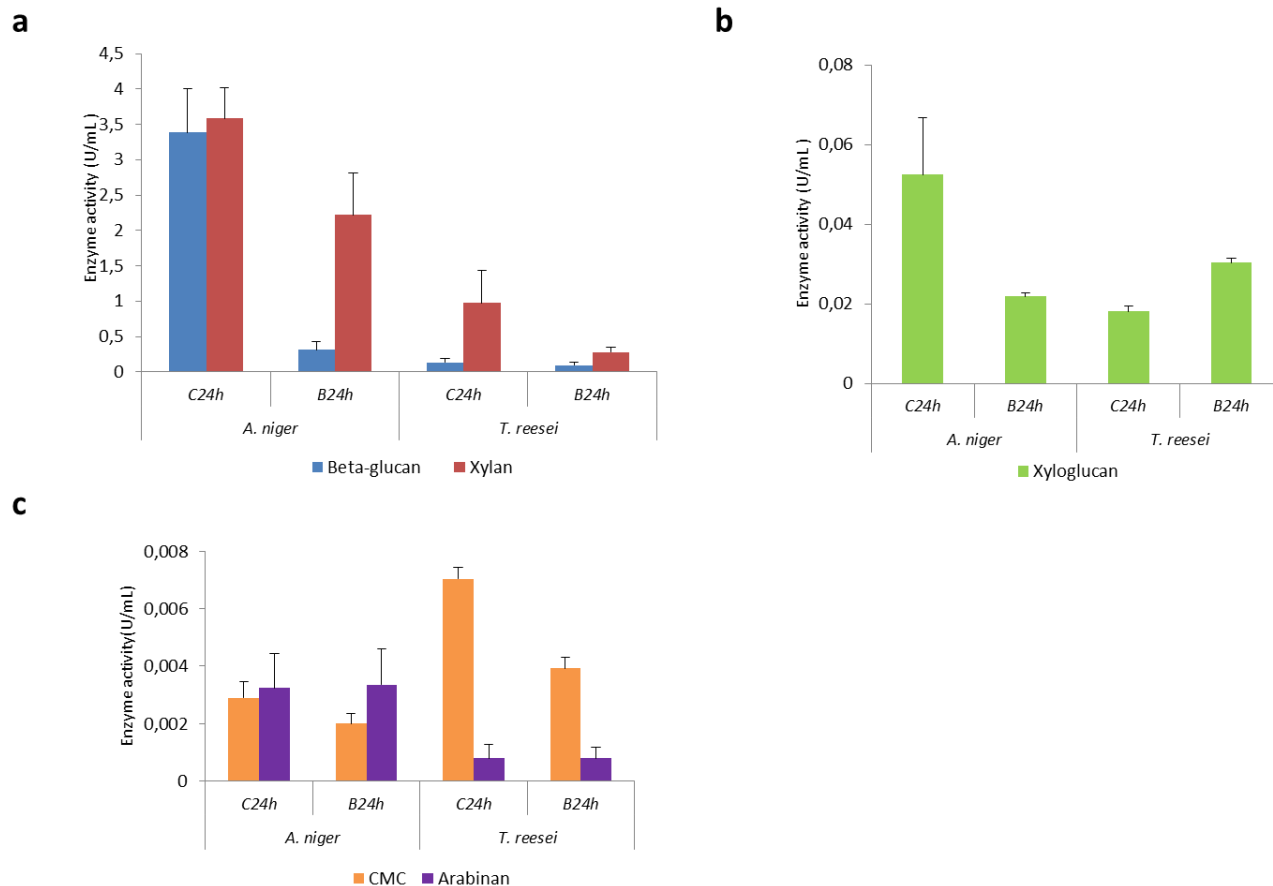
*T. reesei* produces the most commercially used cellulases [42], and among the GHs related to biomass degradation found exclusively on its secretome, we identified a  $\beta$ -glucosidase from the GH1 family. At earlier time points, the mode by which *T. reesei* attacks arabinoxylans seems to differ from that of *A. niger*, with an exclusive acetyl xylan esterase (CE5) and an endo-xylanase (GH30). No evidence of the presence of laccases were found for any of the species studied, suggesting that the hydrolytic attack of these fungi toward sugarcane biomass does not involve lignin degradation, at least on the time points tested.

We also performed a comparative analysis of the total number of peptides. In agreement with previous studies [23,24], the cellobiohydrolase Cel7A displayed the highest number of peptides in the *T. reesei* secretome (Table S1). The cellobiohydrolase Cel6A, was the third most abundant enzyme; together with CE17A, these enzymes represent 80% of the peptides from cellulases (Table S1), as reported by other authors [22,24]. An enzyme from the GH16, a putative glucanosyl transferase (ID 66752) related to cell wall biosynthesis and morphogenesis, was the second most abundant protein. Unlike *T. reesei*, the relative analysis of abundance revealed that the peptides from  $\alpha$ -L-arabinofuranosidase B (GH54) and xylanase 2 (GH11) were the most abundant peptides in *A. niger*. Although Adav *et al.* [20] described



the same arabinofuranosidase as the one of most abundant proteins after D-xylose induction, xylanase 2 was barely detected. This xylanase was induced by culm and bagasse in *A. niger*, but it was less prevalent in *T. reesei* (Fig. 3, Table S1).

Interestingly, when we compared the growth of both fungi, *T. reesei* had a larger mycelia mass compared to *A. niger* (indirectly measured by nitrogen content, Fig. 2). Considering the abundance of secreted enzymes in both fungi, *A. niger* produced not only a wider range of enzymes but also secreted higher quantities compared to *T. reesei* (Fig. 4). To verify whether the abundance of peptides reflected the abundance of enzymes, we performed enzymatic assays using beechwood,  $\beta$ -glucan, debranched arabinan from sugar beet, xyloglucan from tamarind and CMC as substrates. Given the low sensitivity of the assay and the small amount of protein, we measured the enzymatic activities at 24 hours, the time point with the highest enzyme production (Tables 1 and 2). Supernatants derived from both cultures were capable of hydrolyzing hemicelluloses, such as  $\beta$ -glucan, xylan (Fig. 5a), xyloglucan (Fig. 5b), arabinan (Fig. 5c), and cellulose: CMC (Fig. 5c), which is in agreement with the identified enzymes. Moreover, enzymes derived from the supernatant fraction of *A. niger* were able to hydrolyze the majority of substrates more effectively than *T. reesei*, corroborating the peptide proportion data presented in showed in Fig. 4. Furthermore, enzymatic activity toward CMC was higher in *T. reesei*, likely due to its remarkable cellulolytic ability.



**Figure 5. Estimation of enzyme activities.** Enzymatic activities (U/mL) against different substrates of *A. niger* and *T. reesei* after 24 hours (h) growth on sugarcane culm (C) and bagasse (B). a) Beta-glucan and xylan b) Xyloglucan, and c) CMC and Arabinan. Each bar represents the mean and the standard deviation of values from three independent experiments.

Therefore, we evaluated secreted enzymes involved in cell wall degradation by both fungi over the experimental time course and tried to correlate the monosaccharides released in the supernatant and enzymes related to the degradation of their polysaccharides. As we mentioned before, our data suggests that arabinoxylan is been consumed within the first 6 hours after attack by both fungal species. Hypothetically, at least for *A. niger*, the attack of

feruloyl esterases and the acetyl esterases might allow for arabinoxylan degradation. At early time points, the general performance of the enzyme production system of *A. niger* corroborated the model of sugarcane cell wall hydrolysis proposed by De Souza *et al.* [4] that reported that sugarcane cell walls are composed of domains that can be extracted with progressively higher concentrations of alkali. Pectins (ca. 10% of the wall) are more soluble, followed by 1,3 and 1,4- $\beta$ -glucans (ca.10% of the wall) and highly interactive arabinoxylans (ca. 40% of the wall) and a xyloglucan/xylan-cellulose domain (ca. 40% of the wall) that is less soluble. Thus pectinases and esterases should be the first enzymes to act on the walls, opening the way for the hemicellulases and cellulases. Observing the appearance of enzymes related to biomass degradation (Fig. 3), we found that the carbohydrate esterases (feruloyl and acetyl xylan esterases (CE1), pectin methylesterase (CE8), rhamnogalacturonan acetyl esterase (CE12), pectin and acetyl esterase (CE16)) were some of the first enzymes secreted by *A. niger*. These enzymes have molecular weights below the threshold that is considered the pore size of the cell wall, 35-40 Angstroms [43]. Thus, these enzymes are likely to penetrate the cell wall matrix before other glycosyl hydrolases that will attach polymer decorations and main chains. Thus, by deferuloylating, demethylating and deacetylating pectins and xylans, the hydrolytic system would hydrate the wall environment (due to the lack of hydrophobic branches), while also making the main chains of these polymers more accessible to endo-enzymes.

From the proteomics point of view, *T. reesei* displays a different strategy to disassemble sugarcane cell walls compared with *A. niger*. The former species produced relatively low quantities and variety of enzymes, and it did not produce pectinases at all. The fact that no feruloyl esterase is present in *T. reesei* extracts suggests that this species employs

a mechanism that uses proportionally fewer debranching enzymes in the early stages of biomass degradation and that it is able to attack cellulose microfibrils without a prior attack to the phenylpropanoids of the cell wall. On the other hand, swollenin was detected in *T.reesei* at all time points in culm and after 12 hours in bagasse. This swollenin isolated from *T.reesei* (the protein was named SWO1) behaves like a plant expansin. When purified, SWO1 disrupted the cell wall structure without the production of free glucose [44]. Swollenin and the acetyl-xylan-esterase (CE5 – 30.5 KDa), which were present only in the *T. reesei* genome and secretome (Fig. 3, Table 1), likely act by disrupting the cell wall architecture and loosening polymer-polymer interaction in a way that polysaccharides become more accessible to glycosyl hydrolases, such as 1,3-1,4- $\beta$ -glucanases, arabinoxylanases and cellulases.

Despite the distinct modes of attack to the biomass, both fungi are able to break down the biomass cell wall since the earliest time points, but we can suggest that *A. niger* invests more in cell wall hydrolysis in terms of number of enzymes and enzyme activities related to cell wall hydrolysis than *T. reesei*, which secreted fewer enzymes. These two different behaviors are probably associated with the biology of each species. Another important aspect that seems to influence the observed behavior of both fungi is that they are limited, to a certain extent, by their respective genomes, i.e., the enzyme arsenals to address biomass [15,45]. Notwithstanding, it must taken into consideration that *T. reesei* was submitted to several rounds of random mutagenesis to obtain the hypercellulolytic strain RUT-C30.

## Conclusions

Here, we provide the first comparative secretome analysis of the most important lignocellulolytic fungal species, *A. niger* and *T. reesei*, growing on sugarcane biomass. This

secretome study indicates that biomass degradation begins within the first 6 hours of fungal growth. A proteomic approach was used to analyze the secretome profiles of *A. niger* and *T. reesei*, and our data indicated that the two fungal species have different modes of attacking the same biomass, at least within 24 hours of the saccharification process. Thus, we advanced our understanding of the synergic mode of attack of swollenin, esterases, and glycosyl hydrolases in the context of enzyme cocktails and architecture of the plant cell wall. Using a semi-quantitative method based on peptide counts, we estimated the relative differences in the amount of extracellular enzyme production. We noticed that the induction of hydrolytic enzymes is higher when both fungi were growing using culm as a carbon source, probably due to the higher recalcitrance of this substrate.

At the time points measured, *A. niger* produced more enzymes (quantitatively and qualitatively) than *T. reesei*, but both species were able to disassemble the carbohydrate portion of sugarcane cell walls. Considering that *T. reesei* and *A. niger* have different mechanisms for degrading biomass, these data suggest that a combination of enzyme from the two species might be an interesting option to increase saccharification efficiency. In other words, the two fungal species might be combined for use in industrial processes.

## **Acknowledgements**

The authors would like to thank Dr. Marina Camara Mattos Martins and Ms. Lucia Wolff for HPAEC-PAD analysis, as well as Dr. Maria Teresa B. Pimenta and Ms. Karen Marabezi for CHNS analysis.

## References

1. EPE (2013) EPE: Balanço Energético Nacional.
2. 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.
3. Buckeridge MS, De Souza AP, Arundale RA, Anderson-Teixeira KJ, DeLucia E (2012) Ethanol from sugarcane in Brazil: a 'midway' strategy for increasing ethanol production while maximizing environmental benefits. *GCB Bioenergy* 4: 119-126.
4. 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. *Bioenerg Res*: 1-16.
5. Rodrigues RCLB, Felipe MGA, Silva JBA, Vitolo M (2003) Response surface methodology for xylitol production from sugarcane bagasse hemicellulosic hydrolyzate using controlled vacuum evaporation process variables. *Process Biochemistry* 38: 1231-1237.
6. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, et al. (2006) The path forward for biofuels and biomaterials. *Science* 311: 484-489.
7. Lin Y, Tanaka S (2006) Ethanol fermentation from biomass resources: current state and prospects. *Appl Microbiol Biotechnol* 69: 627-642.
8. Patel-Predd P (2006) Overcoming the hurdles to producing ethanol from cellulose. *Environ Sci Technol* 40: 4052-4053.
9. Maki M, Leung KT, Qin W (2009) The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. *Int J Biol Sci* 5: 500-516.
10. Borges TA, Souza AT, Squina FM, Riaño-Pachón DM, Santos RAC, et al. (2013) Biochemical characterization of an endoxylanase from *Pseudozyma brasiliensis* sp. nov. strain GHG001 isolated from the intestinal tract of Chrysomelidae larvae associated to sugarcane roots. *Process Biochemistry*.
11. Bussink HJ, Buxton FP, Fraaye BA, de Graaff LH, Visser J (1992) The polygalacturonases of *Aspergillus niger* are encoded by a family of diverged genes. *Eur J Biochem* 208: 83-90.
12. Patil SR, Dayanand A (2006) Optimization of process for the production of fungal pectinases from deseeded sunflower head in submerged and solid-state conditions. *Bioresour Technol* 97: 2340-2344.
13. Polizeli ML, Rizzatti AC, Monti R, Terenzi HF, Jorge JA, et al. (2005) Xylanases from fungi: properties and industrial applications. *Appl Microbiol Biotechnol* 67: 577-591.
14. Costa-Ferreira M, Dias A, Maximo C, Morgado MJ, Sena-Martins G, et al. (1994) Xylanolytic enzyme production by an *Aspergillus niger* isolate. *Appl Biochem Biotechnol* 44: 231-242.
15. Pel HJ, de Winde JH, Archer DB, Dyer PS, Hofmann G, et al. (2007) Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat Biotechnol* 25: 221-231.
16. Stricker AR, Mach RL, de Graaff LH (2008) Regulation of transcription of cellulases- and hemicellulases-encoding genes in *Aspergillus niger* and *Hypocrea jecorina* (*Trichoderma reesei*). *Appl Microbiol Biotechnol* 78: 211-220.

17. Hakkinen M, Valkonen MJ, Westerholm-Parvinen A, Aro N, Arvas M, et al. (2014) Screening of candidate regulators for cellulase and hemicellulase production in *Trichoderma reesei* and identification of a factor essential for cellulase production. *Biotechnol Biofuels* 7: 14.
18. Glass NL, Schmoll M, Cate JH, Coradetti S (2013) Plant cell wall deconstruction by ascomycete fungi. *Annu Rev Microbiol* 67: 477-498.
19. Tani S, Kawaguchi T, Kobayashi T (2014) Complex regulation of hydrolytic enzyme genes for cellulosic biomass degradation in filamentous fungi. *Appl Microbiol Biotechnol*.
20. Adav SS, Li AA, Manavalan A, Punt P, Sze SK (2010) Quantitative iTRAQ secretome analysis of *Aspergillus niger* reveals novel hydrolytic enzymes. *J Proteome Res* 9: 3932-3940.
21. Oliveira JM, van Passel MW, Schaap PJ, de Graaff LH (2011) Proteomic analysis of the secretory response of *Aspergillus niger* to D-maltose and D-xylose. *PLoS One* 6: 1-9.
22. Herpoel-Gimbert I, Margeot A, Dolla A, Jan G, Molle D, et al. (2008) Comparative secretome analyses of two *Trichoderma reesei* RUT-C30 and CL847 hypersecretory strains. *Biotechnol Biofuels* 1: 18.
23. Jun H, Kieselbach T, Jonsson LJ (2011) Enzyme production by filamentous fungi: analysis of the secretome of *Trichoderma reesei* grown on unconventional carbon source. *Microb Cell Fact* 10: 68.
24. Adav SS, Chao LT, Sze SK (2012) Quantitative secretomic analysis of *Trichoderma reesei* strains reveals enzymatic composition for lignocellulosic biomass degradation. *Mol Cell Proteomics* 11: 1-15.
25. de Souza WR, de Gouvea PF, Savoldi M, Malavazi I, de Souza Bernardes LA, et al. (2011) Transcriptome analysis of *Aspergillus niger* grown on sugarcane bagasse. *Biotechnol Biofuels* 4: 40.
26. Marx IJ, van Wyk N, Smit S, Jacobson D, Viljoen-Bloom M, et al. (2013) Comparative secretome analysis of *Trichoderma asperellum* S4F8 and *Trichoderma reesei* Rut C30 during solid-state fermentation on sugarcane bagasse. *Biotechnol Biofuels* 6: 172.
27. Miller GL (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Chem* 31: 426-428.
28. Patterson RK (1973) Automated Pregl-Dumas technique for determining total carbon, hydrogen, and nitrogen in atmospheric aerosols. *Anal Chem* 45: 605-609.
29. Paes Leme AF, Sherman NE, Smalley DM, Sizukusa LO, Oliveira AK, et al. (2012) Hemorrhagic activity of HF3, a snake venom metalloproteinase: insights from the proteomic analysis of mouse skin and blood plasma. *J Proteome Res* 11: 279-291.
30. van den Brink J, de Vries RP (2011) Fungal enzyme sets for plant polysaccharide degradation. *Appl Microbiol Biotechnol* 91: 1477-1492.
31. Lu X, Sun J, Nimtz M, Wissing J, Zeng AP, et al. (2010) The intra- and extracellular proteome of *Aspergillus niger* growing on defined medium with xylose or maltose as carbon substrate. *Microb Cell Fact* 9: 1-13.
32. Braaksma M, Martens-Uzunova ES, Punt PJ, Schaap PJ (2010) An inventory of the *Aspergillus niger* secretome by combining in silico predictions with shotgun proteomics data. *BMC Genomics* 11: 584.

33. Bussamara R, Fuentefria AM, de Oliveira ES, Broetto L, Simcikova M, et al. (2010) Isolation of a lipase-secreting yeast for enzyme production in a pilot-plant scale batch fermentation. *Bioresour Technol* 101: 268-275.
34. van der Kaaij RM, Yuan XL, Franken A, Ram AF, Punt PJ, et al. (2007) Two novel, putatively cell wall-associated and glycosylphosphatidylinositol-anchored alpha-glucanotransferase enzymes of *Aspergillus niger*. *Eukaryot Cell* 6: 1178-1188.
35. Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev* 66: 506-577, table of contents.
36. Harmsen JA, Kusters-van Someren MA, Visser J (1990) Cloning and expression of a second *Aspergillus niger* pectin lyase gene (peIA): indications of a pectin lyase gene family in *A. niger*. *Curr Genet* 18: 161-166.
37. Acuna-Arguelles ME, Gutierrez-Rojas M, Viniegra-Gonzalez G, Favela-Torres E (1995) Production and properties of three pectinolytic activities produced by *Aspergillus niger* in submerged and solid-state fermentation. *Appl Microbiol Biotechnol* 43: 808-814.
38. Solis-Pereyra S, Favela-Torres E, Gutierrez-Rojas M, Roussos S, Saucedo-Castaneda G, et al. (1996) Production of pectinases by *Aspergillus niger* in solid state fermentation at high initial glucose concentrations. *World J Microbiol Biotechnol* 12: 257-260.
39. Corbin RW, Paliy O, Yang F, Shabanowitz J, Platt M, et al. (2003) Toward a protein profile of *Escherichia coli*: comparison to its transcription profile. *Proc Natl Acad Sci U S A* 100: 9232-9237.
40. Gao J, Opiteck GJ, Friedrichs MS, Dongre AR, Hefta SA (2003) Changes in the protein expression of yeast as a function of carbon source. *J Proteome Res* 2: 643-649.
41. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels* 6: 41.
42. Mohanram S, D. A, J. C, A. A, L. N (2013) Novel perspectives for evolving enzyme cocktails for lignocellulose hydrolysis in biorefineries. *Sustainable chemical process* 1: 1-15.
43. Carpita N, Sabulase D, Montezinos D, Delmer DP (1979) Determination of the pore size of cell walls of living plant cells. *Science* 205: 1144-1147.
44. Saloheimo M, Paloheimo M, Hakola S, Pere J, Swanson B, et al. (2002) Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. *Eur J Biochem* 269: 4202-4211.
45. Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, et al. (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol* 26: 553-560.

## Supporting Information

**Por motivo de espaço, não é apresentado o material suplementar das proteínas**

**CAZymes identificadas no secretoma de *A. niger* e *T. reesei*.**



**ANEXO II.** 50 genes mais hiper-expressos no transcriptoma de *T. reesei* crescido em colmo e bagaço de cana, após 6, 12 e 24 horas. Os transcritos foram identificados pela técnica *high-throughput* de RNA-Seq.

**Tabela 1.** Genes mais hiper-expressos na condição colmo 6 horas. CAZymes, proteínas acessórias, transportadores hipotéticos de açúcar, fatores de transcrição hipotéticos e proteínas de função desconhecida são mostrados em negrito, nas cores **laranja**, **azul**, **verde**, **vermelho** e **amarelo**, respectivamente.

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	C6h	F24h		
<b>33630</b>	<b>26,8</b>	<b>0,1</b>	<b>7,7</b>	<b>Putative xylose permease</b>
134448	2,7	0,0	6,9	<i>GH31 Candidate <math>\alpha</math>-xylosidase/<math>\alpha</math>-glucosidase</i>
133383	2,0	0,0	6,9	<i>GH28 Candidate endo-polygalacturonase</i>
118070	9,2	0,1	6,2	<i>GH62 Candidate <math>\alpha</math>-L-arabinofuranosidase, abf2</i>
<b>77945</b>	<b>160,6</b>	<b>4,6</b>	<b>5,1</b>	<b>Putative MFS transporter</b>
11580	2,9	0,1	5,1	<i>GH30 Candidate endo-<math>\beta</math>-1,6-galactanase</i>
70716	1,2	0,0	5,0	<i>CE10 Alpha/beta-hydrolase</i>
91127	43,2	1,3	5,0	<b>Putative MFS transporter</b>
11734	64,8	2,4	4,8	<b>General substrate transporter</b>
109243	826,0	30,8	4,7	<b>General substrate transporter</b>
42889	8,4	0,3	4,7	<i>Hypothetical protein</i>
<b>26932</b>	<b>28,0</b>	<b>1,1</b>	<b>4,7</b>	<b>Putative MFS monosaccharide transporter</b>
26904	574,2	23,0	4,6	<i>Putative peptide transporter</i>
101532	15,2	0,7	4,5	<i>Quinate permease</i>
<b>114133</b>	<b>2,7</b>	<b>0,1</b>	<b>4,3</b>	<b>AA3 GMC oxidoreductase</b>
122377	2,2	0,1	4,3	<i>GH5 <math>\beta</math>-mannanase, man1</i>
42618	6,7	0,3	4,3	<i>Hypothetical protein</i>
<b>136712</b>	<b>31,3</b>	<b>1,7</b>	<b>4,2</b>	<b>D-xylose proton symporter</b>
61517	2,5	0,1	4,2	<i>Hypothetical protein</i>
134352	12,0	0,7	4,1	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>
90302	3,3	0,2	4,0	<i>GH67 <math>\alpha</math>-glucuronidase, glr1</i>
5304	18,7	1,2	4,0	<i>GH7 Endo-<math>\beta</math>-1,4-glucanase I, egl1</i>
25940	7,3	0,5	4,0	<i>GH45 Endo-<math>\beta</math>-1,4-glucanase, egl5/cel45a</i>
125268	67,7	4,5	3,9	<i>GH3 Candidate <math>\beta</math>-glucosidase, cel3c</i>
125125	129,2	8,9	3,9	<i>GH7 Exocellobiohydrolase I, cbh1/cel7a</i>
<b>25063</b>	<b>5,8</b>	<b>0,4</b>	<b>3,8</b>	<b>Putative MFS transporter</b>
68955	2,0	0,2	3,7	<i>Hypothetical protein</i>
26903	3,6	0,3	3,6	<i>Hypothetical protein</i>
<b>90719</b>	<b>9,8</b>	<b>0,8</b>	<b>3,6</b>	<b>Transmembrane transporter</b>
<b>91594</b>	<b>6,1</b>	<b>0,5</b>	<b>3,6</b>	<b>Putative maltose permease</b>
26551	10,9	1,0	3,4	<i>Putative oxidoreductase CipA-like protein</i>
<b>115868</b>	<b>48,6</b>	<b>4,5</b>	<b>3,4</b>	<b>AA3 Alcohol dehydrogenase</b>
<b>139633</b>	<b>16,1</b>	<b>1,5</b>	<b>3,4</b>	<b>AA9 Lytic polysaccharide monoxygenase, cel61a/egl4</b>
125575	1,0	0,1	3,4	<i>CE15 Glucuronoyl esterase, cip2</i>
136770	2,5	0,2	3,4	<i>CE16 Acetyl esterase, aes1</i>
131496	5,4	0,5	3,4	<i>Hypothetical protein</i>

(Continuação)

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	C6h	F24h		
31758	3,0	0,3	3,3	<i>Hypothetical protein</i>
121449	17,8	1,8	3,3	<i>CBM1 Cellulose induced protein, cip1</i>
38765	19,0	2,0	3,3	<i>General substrate transporter</i>
130622	56,0	5,8	3,3	<i>Hypothetical protein</i>
127115	239,0	26,0	3,2	<i>GH1 β-glucosidase, bgl2/cel1a</i>
136698	4,1	0,5	3,2	<i>Hypothetical protein</i>
122470	55,7	6,2	3,2	<i>GH6 Cellobiohydrolase II, cbh2/cel6a</i>
77521	6,7	0,8	3,1	<i>GH43 Candidate β-xylosidase/α-L-arabinofuranosidase</i>
12995	11,7	1,4	3,1	<i>GH125 DUF1237 domain protein</i>
88058	3,4	0,4	3,0	<i>MYND zinc finger</i>
128051	2,1	0,3	3,0	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>
39277	3,6	0,5	3,0	<i>GH27 α-galactosidase, agl3</i>
84089	7,5	1,0	2,9	<i>MSF superfamily transporter</i>
75420	12,5	1,6	2,9	<i>GH27 Alpha-galactosidase 3</i>

**Tabela 2.** Genes mais hiper-expressos na condição colmo 12 horas.

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	C12h	F24h		
33630	49,0	0,1	8,6	<i>Putative xylose permease</i>
122518	67,7	0,2	8,2	<i>AA9 Lytic polysaccharide monooxygenase, cel61b</i>
133383	3,8	0,0	7,8	<i>GH28 Candidate endo-polygalacturonase</i>
134448	5,0	0,0	7,8	<i>GH31 Candidate α-xylosidase/α-glucosidase</i>
118070	24,1	0,1	7,6	<i>GH62 Candidate α-L-arabinofuranosidase, abf2</i>
124438	77,9	0,5	7,4	<i>GH12 Endo-β-1,4-glucanase III, egl3/cel12a</i>
125575	13,1	0,1	7,1	<i>CE15 Glucuronoyl esterase, cip2</i>
131826	25,5	0,2	6,8	<i>Iron transporter</i>
122377	10,2	0,1	6,5	<i>GH5 β-mannanase, man1</i>
136770	19,4	0,2	6,3	<i>CE16 Acetyl esterase, aes1</i>
23616	9,6	0,1	6,3	<i>GH10 Endo-β-1,4-xylanase, xyn3</i>
140920	28,0	0,4	6,2	<i>Lactate/pyruvate transporter</i>
25940	31,8	0,5	6,1	<i>GH45 Endo-β-1,4-glucanase, egl5/cel45a</i>
109243	1931,6	30,8	6,0	<i>General substrate transporter</i>
125125	530,1	8,9	5,9	<i>GH7 Exocellobiohydrolase I, cbh1/cel7a</i>
70716	2,1	0,0	5,8	<i>CE10 Alpha/beta-hydrolase</i>
11580	4,8	0,1	5,8	<i>GH30 Candidate endo-β-1,6-galactanase</i>
121449	94,7	1,8	5,8	<i>CBM1 Cellulose induced protein, cip1</i>
26932	57,0	1,1	5,7	<i>Putative MFS monosaccharide transporter</i>
111088	7,6	0,2	5,6	<i>Zn-dependent exopeptidase</i>
139633	74,1	1,5	5,6	<i>AA9 Lytic polysaccharide monooxygenase, cel61a/egl4</i>
75795	16,9	0,4	5,6	<i>Hypothetical protein</i>
122470	257,9	6,2	5,4	<i>GH6 Cellobiohydrolase II, cbh2/cel6a</i>
72489	110,8	2,8	5,3	<i>GH5 Endo-β-1,4-glucanase II, egl2/cel5a</i>
73367	1,7	0,0	5,3	<i>MFS general substrate transporter</i>
94554	388,8	10,0	5,3	<i>Oligopeptide transporter</i>

(Continuação)

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	C12h	F24h		
90302	7,8	0,2	5,3	<i>GH67 α-glucuronidase, glr1</i>
5304	43,2	1,2	5,2	<i>GH7 Endo-β-1,4-glucanase I, egl1</i>
126605	6,4	0,2	5,2	<i>4-coumarate-CoA ligase</i>
139631	39,7	1,2	5,1	<i>CE5 Acetylxylan esterase, axe1</i>
26904	690,4	23,0	4,9	<i>Putative peptide transporter</i>
91594	14,5	0,5	4,8	<i>Putative maltose permease</i>
38144	1,4	0,1	4,8	<i>Ferric reductase-like transmembrane component</i>
5086	8,3	0,3	4,8	<i>Hypothetical protein</i>
90719	22,2	0,8	4,8	<i>Transmembrane transporter</i>
140746	25,4	1,1	4,6	<i>GH3 β-xylosidase, bxl1</i>
93498	10,1	0,5	4,5	<i>GH30 Candidate endo-β-1,4-xylanase</i>
53480	3,0	0,1	4,4	<i>Ribonuclease H-like</i>
92851	2,4	0,1	4,4	<i>HHE domain protein</i>
139450	191,9	9,1	4,4	<i>Dicarboxylic amino acid permease</i>
42889	6,9	0,3	4,4	<i>Hypothetical protein</i>
134325	36,3	1,8	4,4	<i>Hypothetical protein</i>
127115	528,9	26,0	4,3	<i>GH1 β-glucosidase, bgl2/cel1a</i>
102517	17,3	0,9	4,2	<i>GH54 α-L-arabinofuranosidase I, abf1</i>
131881	7,2	0,4	4,2	<i>Quinate dehydrogenase</i>
38545	4,5	0,3	4,1	<i>UbiE/COQ5 methyltransferase</i>
114133	2,3	0,1	4,1	<i>AA3 GMC oxidoreductase</i>
130622	99,2	5,8	4,1	<i>Hypothetical protein</i>
124931	62,5	3,7	4,1	<i>GH11 Endo-β-1,4-xylanase, xyn2</i>
105247	55,4	3,6	4,0	<i>Argininosuccinate synthase</i>

**Tabela 3.** Genes mais hiper-expressos na condição colmo 24 horas.

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	C24h	F24h		
134945	191,5	0,0	12,0	<i>GH11 Candidate endo-β-1,4-xylanase, xyn5</i>
134448	86,7	0,0	11,9	<i>GH31 Candidate α-xylosidase/α-glucosidase</i>
122377	370,1	0,1	11,7	<i>GH5 β-mannanase, man1</i>
133383	47,3	0,0	11,5	<i>GH28 Candidate endo-polygalacturonase</i>
125575	140,0	0,1	10,5	<i>CE15 Glucuronoyl esterase, cip2</i>
122518	334,6	0,2	10,5	<i>AA9 Lytic polysaccharide monooxygenase, cel61b</i>
118070	157,0	0,1	10,3	<i>GH62 Candidate α-L-arabinofuranosidase, abf2</i>
38144	53,4	0,1	10,0	<i>Ferric reductase-like transmembrane component</i>
33630	133,3	0,1	10,0	<i>Putative xylose permease</i>
136770	193,2	0,2	9,7	<i>CE16 Acetyl esterase, aes1</i>
121449	1386,4	1,8	9,6	<i>CBM1 Cellulose induced protein, cip1</i>
139633	1061,3	1,5	9,5	<i>AA9 Lytic polysaccharide monooxygenase, cel61a/egl4</i>
139631	825,2	1,2	9,5	<i>CE5 Acetylxylan esterase, axe1</i>
91667	10,3	0,0	9,4	<i>GH28 Candidate exo-polygalacturonase</i>

(Continuação)

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	C24h	F24h		
125125	5822,2	8,9	9,3	<i>GH7 Exocellobiohydrolase I, cbh1/cel7a</i>
25940	289,3	0,5	9,3	<i>GH45 Endo-β-1,4-glucanase, egl5/cel45a</i>
5304	716,4	1,2	9,3	<i>GH7 Endo-β-1,4-glucanase I, egl1</i>
91594	283,2	0,5	9,1	<i>Putative maltose permease</i>
72489	1307,4	2,8	8,9	<i>GH5 Endo-β-1,4-glucanase II, egl2/cel5a</i>
122470	2535,3	6,2	8,7	<i>GH6 Cellobiohydrolase II, cbh2/cel6a</i>
124931	1388,0	3,7	8,5	<i>GH11 Endo-β-1,4-xylanase, xyn2</i>
93498	139,8	0,5	8,2	<i>GH30 Candidate endo-β-1,4-xylanase</i>
73367	12,1	0,0	8,1	<i>MFS general substrate transporter</i>
6699	5,5	0,0	8,1	<i>GH105 Candidate rhamnogalacturonyl hydrolase</i>
38418	166,9	0,6	8,0	<i>GH11 Endo-β-1,4-xylanase, xyn1</i>
140746	260,3	1,1	7,9	<i>GH3 β-xylosidase, bxl1</i>
90302	48,3	0,2	7,9	<i>GH67 α-glucuronidase, glr1</i>
127981	196,3	0,8	7,9	<i>GH109 NAD-binding Rossmann fold oxidoreductase</i>
88887	77,8	0,3	7,9	<i>CE5 Candidate acetyl xylan esterase</i>
45640	4,3	0,0	7,7	<i>Hypothetical protein</i>
124438	91,2	0,5	7,6	<i>GH12 Endo-β-1,4-glucanase III, egl3/cel12a</i>
109243	5219,3	30,8	7,4	<i>General substrate transporter</i>
11580	13,0	0,1	7,3	<i>GH30 Candidate endo-β-1,6-galactanase</i>
75795	46,1	0,4	7,0	<i>Hypothetical protein</i>
90847	214,4	1,9	6,8	<i>GH30 Endo-β-1,4-xylanase, xyn4</i>
23616	10,4	0,1	6,4	<i>GH10 Endo-β-1,4-xylanase, xyn3</i>
131826	17,2	0,2	6,3	<i>Iron transporter</i>
104220	566,2	7,5	6,2	<i>CBM1 Swollenin, swo1</i>
71952	5,6	0,1	6,2	<i>Hypothetical protein</i>
127980	44,4	0,6	6,2	<i>MFS lactose permease</i>
136547	33,5	0,5	6,1	<i>GH3 β-glucosidase, bgl1/cel3a</i>
39271	2,1	0,0	6,0	<i>NmrA-like family</i>
97259	28,0	0,4	6,0	<i>General substrate transporter</i>
138519	494,0	8,0	6,0	<i>General substrate transporter</i>
111943	362,3	6,6	5,8	<i>GH74 Xyloglucanase, cel74a</i>
73291	1,2	0,0	5,7	<i>Cysteine synthase</i>
102517	47,0	0,9	5,7	<i>GH54 α-L-arabinofuranosidase I, abf1</i>
127115	1302,7	26,0	5,6	<i>GH1 β-glucosidase, bgl2/cel1a</i>
67663	12,5	0,3	5,5	<i>Acetamidase/Formamidase</i>
70716	1,8	0,0	5,5	<i>CE10 Alpha/beta-hydrolase</i>

**Tabela 4.** Genes mais hiper-expressos na condição bagaço 6 horas.

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	B6h	F24h		
33630	38,8	0,1	8,2	<i>Putative xylose permease</i>
134448	6,0	0,0	8,0	<i>GH31 Candidate <math>\alpha</math>-xylosidase/<math>\alpha</math>-glucosidase</i>
140746	268,8	1,1	8,0	<i>GH3 <math>\beta</math>-xylosidase, bxl1</i>
90302	28,0	0,2	7,1	<i>GH67 <math>\alpha</math>-glucuronidase, glr1</i>
37499	1,1	0,0	6,8	<i><math>\beta</math>-1,6-N-acetylglucosaminyltransferase, contains WSC domain</i>
124931	343,3	3,7	6,5	<i>GH11 Endo-<math>\beta</math>-1,4-xylanase, xyn2</i>
131826	20,7	0,2	6,5	<i>Iron transporter</i>
136770	21,3	0,2	6,5	<i>CE16 Acetyl esterase, aes1</i>
138519	655,5	8,0	6,4	<i>General substrate transporter</i>
118070	8,3	0,1	6,0	<i>GH62 Candidate <math>\alpha</math>-L-arabinofuranosidase, abf2</i>
26932	67,8	1,1	5,9	<i>Putative MFS monosaccharide transporter</i>
134945	2,5	0,0	5,7	<i>GH11 Candidate endo-<math>\beta</math>-1,4-xylanase, xyn5</i>
93518	2,1	0,0	5,7	<i>Putative short-chain dehydrogenases/reductase</i>
139450	446,7	9,1	5,6	<i>Dicarboxylic amino acid permease</i>
109243	1496,5	30,8	5,6	<i>General substrate transporter</i>
11580	3,6	0,1	5,4	<i>GH30 Candidate endo-<math>\beta</math>-1,6-galactanase</i>
90847	77,6	1,9	5,3	<i>GH30 Endo-<math>\beta</math>-1,4-xylanase, xyn4</i>
26391	121,9	3,1	5,3	<i>Fatty acid desaturase, type 1</i>
140920	15,2	0,4	5,3	<i>Lactate/pyruvate transporter</i>
90719	28,4	0,8	5,1	<i>Transmembrane transporter</i>
127852	1,6	0,0	5,0	<i>Amidohydrolase</i>
88887	10,6	0,3	5,0	<i>CE5 Candidate acetyl xylan esterase</i>
127981	26,3	0,8	5,0	<i>GH109 NAD-binding Rossmann fold oxidoreductase</i>
92240	83,1	2,7	5,0	<i>MIP family channel protein</i>
125575	2,8	0,1	4,9	<i>CE15 Glucuronoyl esterase, cip2</i>
111088	4,3	0,2	4,8	<i>Zn-dependent exopeptidase</i>
94554	277,0	10,0	4,8	<i>Oligopeptide transporter</i>
75795	9,4	0,4	4,7	<i>Hypothetical protein</i>
97585	63,3	2,4	4,7	<i>Cell wall protein</i>
88640	1,7	0,1	4,6	<i>Hypothetical protein</i>
56646	5,6	0,2	4,6	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>
94809	712,3	29,1	4,6	<i>NAD(P)H-dependent D-xylose reductase xyl1</i>
139631	28,2	1,2	4,6	<i>CE5 Acetylxylan esterase, axe1</i>
124397	388,0	16,3	4,6	<i>GT32 Glycosyltransferase</i>
126605	4,0	0,2	4,5	<i>4-coumarate-CoA ligase</i>
26904	492,7	23,0	4,4	<i>Putative peptide transporter</i>
115868	92,9	4,5	4,4	<i>AA3 Alcohol dehydrogenase</i>

(Continuação)

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	B6h	F24h		
77521	15,9	0,8	4,4	<i>GH43 Candidate β-xylosidase/α-L-arabinofuranosidase</i>
123870	86,8	4,3	4,3	<i>L-arabinitol 4-dehydrogenase</i>
91594	9,9	0,5	4,3	<i>Putative maltose permease</i>
130622	111,5	5,8	4,3	<i>Hypothetical protein</i>
33351	1,7	0,1	4,2	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>
91584	1,1	0,1	4,2	<i>Haloacid dehalogenase-like hydrolase</i>
71952	1,2	0,1	4,0	<i>Hypothetical protein</i>
121449	28,1	1,8	4,0	<i>CBM1 Cellulose induced protein, cip1</i>
139633	23,8	1,5	4,0	<i>AA9 Lytic polysaccharide monooxygenase, cel61a/egl4</i>
73914	90,9	5,8	4,0	<i>ABC transporter</i>
98410	2,6	0,2	3,9	<i>Phosphate permease</i>
127980	9,2	0,6	3,9	<i>MFS lactose permease</i>
127115	395,8	26,0	3,9	<i>GH1 β-glucosidase, bgl2/cel1a</i>

**Tabela 5.** Genes mais hiper-expressos na condição bagaço 12 horas.

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	B12h	F24h		
140746	231,0	1,1	7,8	<i>GH3 β-xylosidase, bxl1</i>
124931	789,0	3,7	7,7	<i>GH11 Endo-β-1,4-xylanase, xyn2</i>
136770	48,9	0,2	7,7	<i>CE16 Acetyl esterase, aes1</i>
90302	37,4	0,2	7,5	<i>GH67 α-glucuronidase, glr1</i>
134448	4,2	0,0	7,5	<i>GH31 Candidate α-xylosidase/α-glucosidase</i>
122518	23,8	0,2	6,7	<i>AA9 Lytic polysaccharide monooxygenase, cel61b</i>
109243	3105,5	30,8	6,7	<i>General substrate transporter</i>
139631	116,1	1,2	6,6	<i>CE5 Acetylxylan esterase, axe1</i>
125575	8,6	0,1	6,5	<i>CE15 Glucuronoyl esterase, cip2</i>
91667	1,4	0,0	6,5	<i>GH28 Candidate exo-polygalacturonase</i>
121449	147,1	1,8	6,4	<i>CBM1 Cellulose induced protein, cip1</i>
138519	670,5	8,0	6,4	<i>General substrate transporter</i>
33630	10,7	0,1	6,4	<i>Putative xylose permease</i>
139633	114,7	1,5	6,3	<i>AA9 Lytic polysaccharide monooxygenase, cel61a/egl4</i>
134945	3,7	0,0	6,3	<i>GH11 Candidate endo-β-1,4-xylanase, xyn5</i>
88887	24,5	0,3	6,2	<i>CE5 Candidate acetyl xylan esterase</i>
75795	25,0	0,4	6,1	<i>Hypothetical protein</i>
38144	3,5	0,1	6,1	<i>Ferric reductase-like transmembrane component</i>



(Continuação)

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	B12h	F24h		
125125	571,9	8,9	6,0	<i>GH7 Exocellobiohydrolase I, cbh1/cel7a</i>
73367	2,6	0,0	5,9	<i>MFS general substrate transporter</i>
72489	165,5	2,8	5,9	<i>GH5 Endo-β-1,4-glucanase II, egl2/cel5a</i>
131826	13,0	0,2	5,9	<i>Iron transporter</i>
127981	48,3	0,8	5,8	<i>GH109 NAD-binding Rossmann fold oxidoreductase</i>
6699	1,1	0,0	5,8	<i>GH105 Candidate rhamnogalacturonyl hydrolase</i>
93336	1,0	0,0	5,6	<i>Heterokaryon incompatibility protein</i>
25940	21,8	0,5	5,6	<i>GH45 Endo-β-1,4-glucanase, egl5/cel45a</i>
124438	21,6	0,5	5,6	<i>GH12 Endo-β-1,4-glucanase III, egl3/cel12a</i>
90719	36,9	0,8	5,5	<i>Transmembrane transporter</i>
93518	1,8	0,0	5,5	<i>Putative short-chain dehydrogenases/reductase</i>
122470	269,6	6,2	5,4	<i>GH6 Cellobiohydrolase II, cbh2/cel6a</i>
77945	194,5	4,6	5,4	<i>Putative MFS transporter</i>
111088	5,9	0,2	5,3	<i>Zn-dependent exopeptidase</i>
140920	14,7	0,4	5,2	<i>Lactate/pyruvate transporter</i>
5304	43,1	1,2	5,2	<i>GH7 Endo-β-1,4-glucanase I, egl1</i>
90847	65,0	1,9	5,1	<i>GH30 Endo-β-1,4-xylanase, xyn4</i>
118070	4,3	0,1	5,1	<i>GH62 Candidate α-L-arabinofuranosidase, abf2</i>
97259	14,3	0,4	5,1	<i>General substrate transporter</i>
122377	3,5	0,1	5,0	<i>GH5 β-mannanase, man1</i>
56646	7,3	0,2	5,0	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>
77521	24,0	0,8	5,0	<i>GH43 Candidate β-xylosidase/α-L-arabinofuranosidase</i>
23616	3,9	0,1	4,9	<i>GH10 Endo-β-1,4-xylanase, xyn3</i>
139450	268,7	9,1	4,9	<i>Dicarboxylic amino acid permease</i>
127852	1,3	0,0	4,8	<i>Amidohydrolase</i>
127115	704,2	26,0	4,8	<i>GH1 β-glucosidase, bgl2/cel1a</i>
94554	234,5	10,0	4,6	<i>Oligopeptide transporter</i>
124397	352,9	16,3	4,4	<i>GT32 Glycosyltransferase</i>
130622	123,1	5,8	4,4	<i>Hypothetical protein</i>
93498	9,7	0,5	4,4	<i>GH30 Candidate endo-β-1,4-xylanase</i>
94809	605,1	29,1	4,4	<i>NAD(P)H-dependent D-xylose reductase xyl1</i>
122639	19,2	0,9	4,3	<i>GH3 Candidate β-glucosidase, cel3d</i>

**Tabela 6.** Genes mais hiper-expressos na condição bagaço 24 horas.

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	B24h	F24h		
121449	1038,7	1,8	9,2	<i>CBM1 Cellulose induced protein, cip1</i>
139631	634,6	1,2	9,1	<i>CE5 Acetylxylan esterase, axe1</i>
125575	50,6	0,1	9,1	<i>CE15 Glucuronoyl esterase, cip2</i>
122518	120,2	0,2	9,0	<i>AA9 Lytic polysaccharide monoxygenase, cel61b</i>
124438	224,5	0,5	8,9	<i>GH12 Endo-β-1,4-glucanase III, egl3/cel12a</i>
88887	158,5	0,3	8,9	<i>CE5 Candidate acetyl xylan esterase</i>
124931	1731,5	3,7	8,9	<i>GH11 Endo-β-1,4-xylanase, xyn2</i>
91667	7,2	0,0	8,8	<i>GH28 Candidate exo-polygalacturonase</i>
122377	46,1	0,1	8,7	<i>GH5 β-mannanase, man1</i>
125125	3727,5	8,9	8,7	<i>GH7 Exocellobiohydrolase I, cbh1/cel7a</i>
139633	560,7	1,5	8,6	<i>AA9 Lytic polysaccharide monoxygenase, cel61a/egl4</i>
72489	1023,5	2,8	8,5	<i>GH5 Endo-β-1,4-glucanase II, egl2/cel5a</i>
23616	44,9	0,1	8,5	<i>GH10 Endo-β-1,4-xylanase, xyn3</i>
25940	163,3	0,5	8,5	<i>GH45 Endo-β-1,4-glucanase, egl5/cel45a</i>
5304	379,2	1,2	8,3	<i>GH7 Endo-β-1,4-glucanase I, egl1</i>
134448	7,0	0,0	8,3	<i>GH31 Candidate α-xylosidase/α-glucosidase</i>
140746	312,5	1,1	8,2	<i>GH3 β-xylosidase, bxl1</i>
122470	1811,5	6,2	8,2	<i>GH6 Cellobiohydrolase II, cbh2/cel6a</i>
38144	13,1	0,1	8,0	<i>Ferric reductase-like transmembrane component</i>
91065	4,9	0,0	7,7	<i>Putative sodium/phosphate symporter</i>
127981	173,1	0,8	7,7	<i>GH109 NAD-binding Rossmann fold oxidoreductase</i>
136770	49,2	0,2	7,7	<i>CE16 Acetyl esterase, aes1</i>
90302	38,6	0,2	7,6	<i>GH67 α-glucuronidase, glr1</i>
109243	5565,2	30,8	7,5	<i>General substrate transporter</i>
6699	3,5	0,0	7,4	<i>GH105 Candidate rhamnogalacturonyl hydrolase</i>
133383	2,7	0,0	7,3	<i>GH28 Candidate endo-polygalacturonase</i>
92233	234,8	1,5	7,3	<i>Chondroitin AC/alginate lyase</i>
75795	54,1	0,4	7,3	<i>Hypothetical protein</i>
93498	65,7	0,5	7,2	<i>GH30 Candidate endo-β-1,4-xylanase</i>
111088	21,4	0,2	7,1	<i>Zn-dependent exopeptidase</i>
136547	68,6	0,5	7,1	<i>GH3 β-glucosidase, bgl1/cel3a</i>
118070	16,2	0,1	7,0	<i>GH62 Candidate α-L-arabinofuranosidase, abf2</i>
97259	51,7	0,4	6,9	<i>General substrate transporter</i>
134945	5,3	0,0	6,8	<i>GH11 Candidate endo-β-1,4-xylanase, xyn5</i>
104220	790,6	7,5	6,7	<i>CBM1 Swollenin, swo1</i>
131826	21,1	0,2	6,6	<i>Iron transporter</i>
26391	272,6	3,1	6,5	<i>Fatty acid desaturase, type 1</i>
73367	3,7	0,0	6,4	<i>MFS general substrate transporter</i>

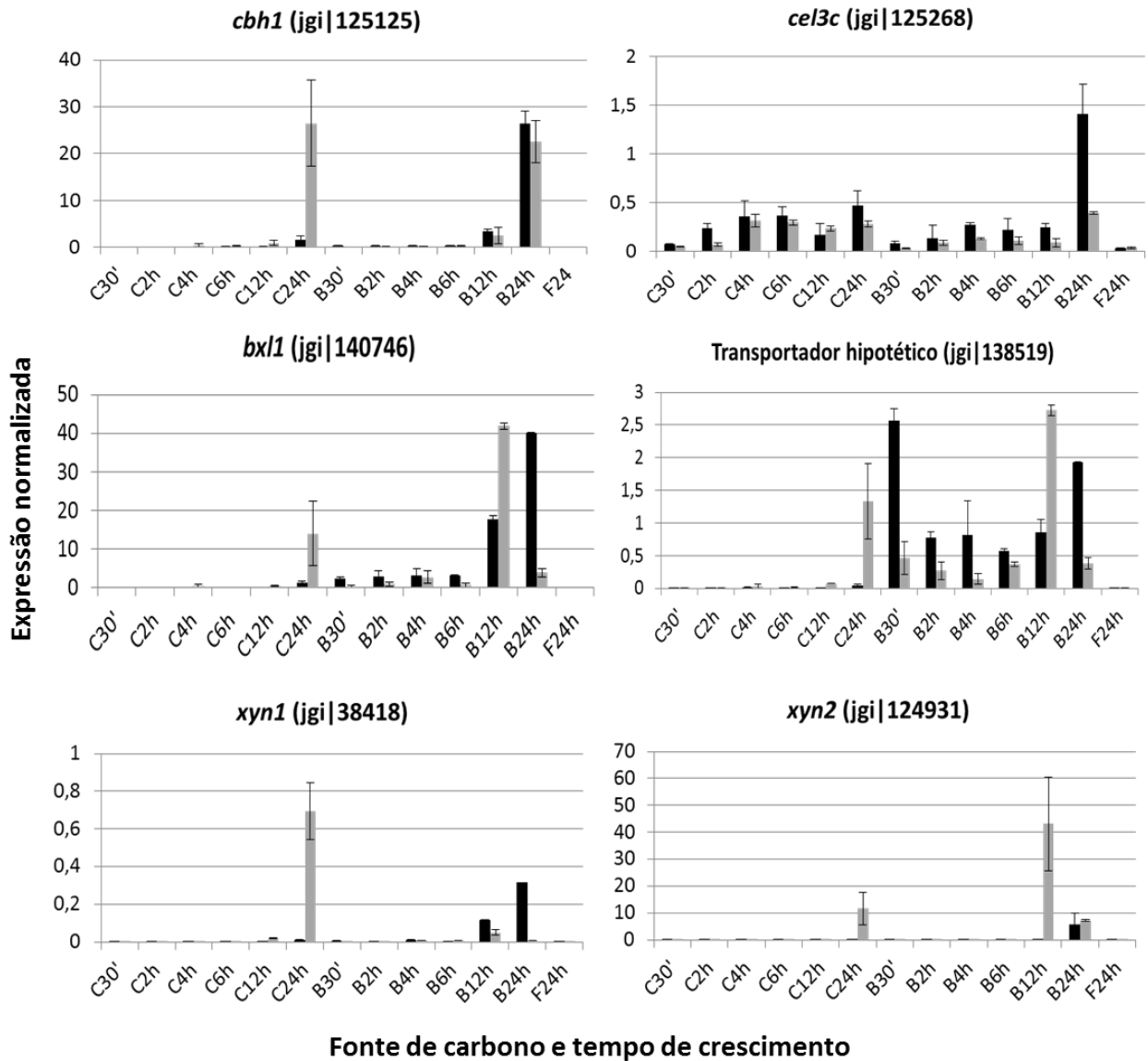


(Continuação)

ID JGI	RPKM		Log <sub>2</sub> (fold change)	Protein
	B24h	F24h		
138519	549,6	8,0	6,1	General substrate transporter
88640	4,7	0,1	6,1	Hypothetical protein
90847	126,2	1,9	6,0	GH30 Endo-β-1,4-xylanase, xyn4
71952	4,7	0,1	5,9	Hypothetical protein
68144	5,8	0,1	5,9	AA1 Ferroxidase-like protein
127115	1569,9	26,0	5,9	GH1 β-glucosidase, bgl2/cel1a
90719	48,5	0,8	5,9	Transmembrane transporter
127980	34,1	0,6	5,8	MFS lactose permease
93518	2,3	0,0	5,8	Putative short-chain dehydrogenases/reductase
122639	49,9	0,9	5,7	GH3 Candidate β-glucosidase, cel3d
140920	19,8	0,4	5,7	Lactate/pyruvate transporter
91594	23,7	0,5	5,5	Putative maltose permease

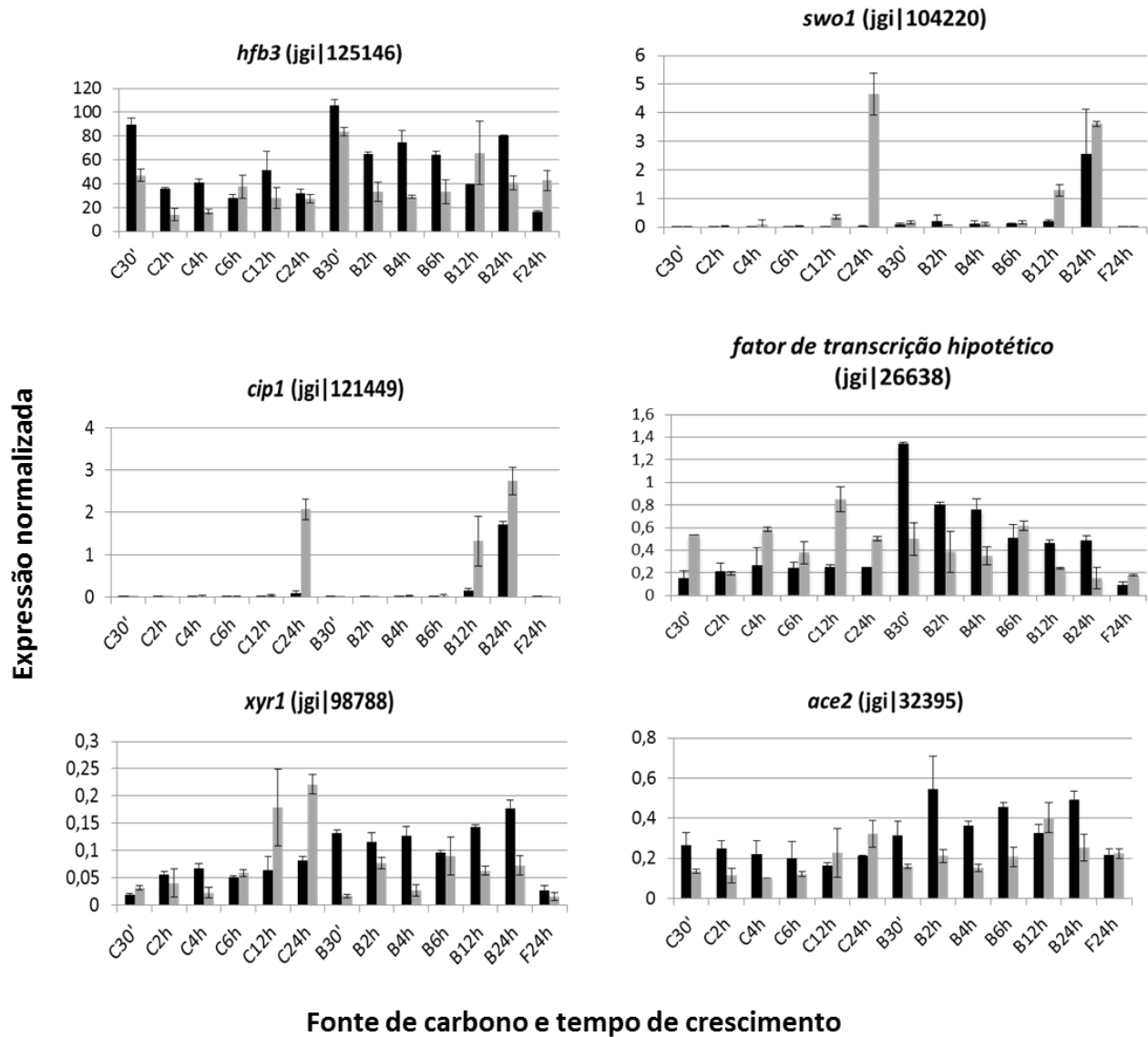


**ANEXO III.** Expressão de genes validados por qPCR para confirmação dos dados de RNA-Seq.



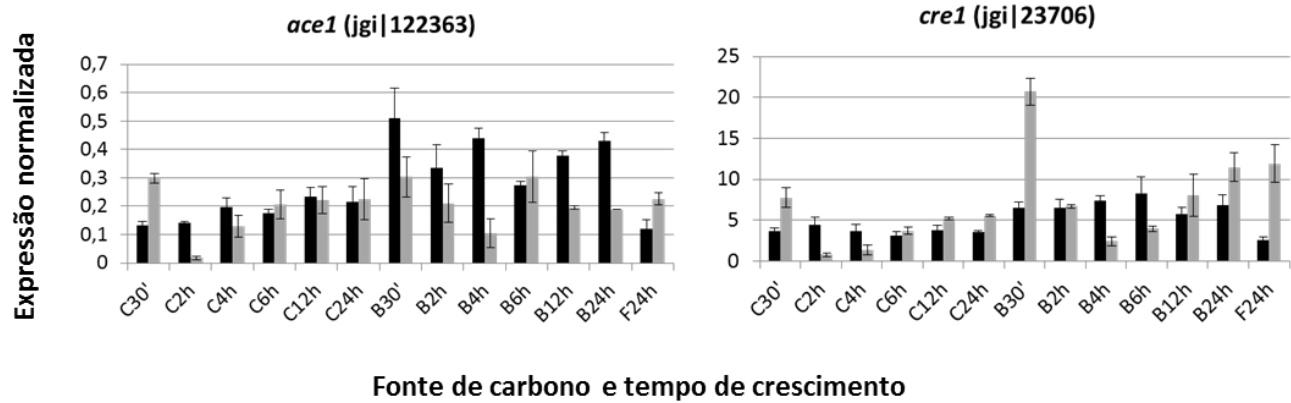
**Figura 1.** Expressão gênica validada por qPCR de celulases (*cbh1* e *cel3c*), hemicelulases (*bx1*, *xyn1* e *xyn2*) e um transportador hipotético de açúcar (jgi|138519). A expressão foi medida para a cepa de *T. reesei* com *cre1* truncado (coluna preta) e *cre1* full (coluna cinza), e normalizada pelos genes *act* e *sar1*. A expressão foi validada após os tempos 30 minutos ('), 2, 4, 6, 12 e 24 horas (h) de crescimento de *T. reesei* em colmo (C) e bagaço (B). Como controle, ou condição não-induzida, foi utilizada a condição frutose (F) 24 horas. A barra de erros indica o desvio padrão.

**ANEXO III. Expressão de genes validados por qPCR para confirmação dos dados de RNA-Seq**



**Figura 2.** Expressão gênica validada por qPCR de proteínas acessórias (*hfb3*, *swo1* e *cip1*), um fator de transcrição hipotético (jgi|26638) e ativadores transcrionais (*xyr1* e *ace2*) de enzimas. A expressão foi medida para a cepa de *T. reesei* com *cre1* truncado (coluna preta) e *cre1* full (coluna cinza), e normalizada pelos genes *act* e *sar1*. A expressão foi validada após os tempos 30 minutos ('), 2, 4, 6, 12 e 24 horas (h) de crescimento de *T. reesei* em colmo (C) e bagaço (B). Como controle, ou condição não-induzida, foi utilizada a condição frutose (F) 24 horas. A barra de erros indica o desvio padrão.

ANEXO III. Expressão de genes validados por qPCR para confirmação dos dados de RNA-Seq.



**Figura 3.** Expressão gênica validada por qPCR de repressores transcricionais (*ace1* e *cre1*). A expressão foi medida para a cepa de *T. reesei* com *cre1* truncado (coluna preta) e *cre1* full (coluna cinza), e normalizada pelos genes *act* e *sar1*. A expressão foi validada após os tempos 30 minutos ('), 2, 4, 6, 12 e 24 horas (h) de crescimento de *T. reesei* em colmo (C) e bagaço (B). Como controle, ou condição não-induzida, foi utilizada a condição frutose (F) 24 horas. A barra de erros indica o desvio padrão.



**ANEXO IV.** Genes hiper-expressos de proteínas de função desconhecida encontrados no transcriptoma de *T. reesei* crescido em biomassa de cana-de-açúcar.

**Tabela 1.** Genes de proteínas com função desconhecida sendo hiper-expressos no transcriptoma de *T. reesei* crescido em colmo (C) e bagaço (B), após 6, 12 e 24 horas (h). ). A identificação (ID) de cada gene é a disponível no banco de dados do JGI do *T. reesei* RUT C30 v1.0. A expressão é apresentada como o  $\log_2$  do *fold change*. A presença de peptídeo sinal (SignalP), segmentos transmembrânicos (TMH) e de domínios encontrados no banco de dados Pfam (Pfam) também é mostrado.

ID JGI	Proteína	$\log_2(\text{fold change})$						SignalP	TMH	Pfam
		C6h	C12h	C24h	B6h	B12h	B24h			
4983	proteína hipotética	1,1	2,0	5,5	1,1	2,7	4,2	-	-	-
42889	proteína hipotética	4,7	4,4	4,8	3,7	3,6	3,6	-	-	-
131184	proteína hipotética	2,2	2,1	4,6	2,1	2,1	4,1	Sim	-	-
88912	proteína hipotética	2,7	3,4	4,2	3,0	3,2	2,2	-	-	-
136698	proteína hipotética	3,2	3,5	4,1	3,2	3,0	4,3	-	-	DUF4459
5086	proteína hipotética	2,3	4,8	3,7	2,4	3,1	4,6	Sim	-	-
130622	proteína hipotética	3,3	4,1	3,4	4,3	4,4	4,5	-	2	-
129378	proteína hipotética	2,2	3,7	3,0	3,6	3,9	5,0	-	-	-
92595	proteína hipotética	2,4	3,1	3,0	3,7	3,6	3,4	-	-	-
64222	proteína hipotética	2,1	2,5	3,0	2,0	2,1	1,8	-	-	-
81213	proteína hipotética	3,0	3,1	2,5	2,5	1,4	2,1	Sim	-	-
140332	proteína hipotética	2,8	3,2	2,3	3,3	3,1	2,7	Sim	1	-
101778	proteína hipotética	2,1	2,2	2,1	1,8	1,6	2,4	Sim	1	-
24323	proteína hipotética	1,5	1,5	1,9	1,4	1,4	2,1	Sim	-	-
9756	proteína hipotética	1,3	1,6	1,3	1,7	1,6	1,5	Sim	3	-
8730	proteína hipotética	1,9	1,5	1,2	2,2	1,5	2,6	Sim	3	-
11207	proteína hipotética	3,7	2,7	1,9	3,1	2,1	-	Sim	-	-
34487	proteína hipotética	-	1,5	1,7	1,0	1,4	1,2	Sim	-	-
134325	proteína hipotética	1,4	4,4	-	2,9	3,1	1,7	Sim	-	-
128163	proteína hipotética	1,6	1,5	-	1,7	1,4	1,1	Sim	-	-

(Continuação)

ID JGI	Proteína	Log <sub>2</sub> (fold change)						SignalP	TMH	Pfam
		C6h	C12h	C24h	B6h	B12h	B24h			
71952	proteína hipotética	-	-	6,2	4,0	4,2	5,9	-	-	-
125009	proteína hipotética	-	3,7	4,7	-	2,7	2,8	-	-	-
85247	proteína hipotética	1,2	2,2	2,8	-	1,9	-	-	-	-
124863	proteína hipotética	-	1,1	1,7	1,1	-	1,2	Sim	1	DUF3955
88454	proteína hipotética	1,3	1,5	1,4	-	1,4	-	Sim	-	DUF3425
126422	proteína hipotética	1,3	1,5	1,4	-	1,1	-	Sim	7	-
88640	proteína hipotética	-	4,0	-	4,6	4,1	6,1	-	-	DUF4419
6589	proteína hipotética	1,9	-	-	1,5	1,2	3,9	Sim	-	-
92691	proteína hipotética	1,5	1,2	-	2,0	1,6	-	Sim	1	-
99737	proteína hipotética	2,0	1,4	-	1,4	-	1,8	Sim	-	-
85898	proteína hipotética	2,1	-	1,6	1,6	-	-	Sim	6	-
5311	proteína hipotética	-	-	1,1	1,1	1,1	-	Sim	3	DUF4267
89722	proteína hipotética	-	-	-	1,2	1,3	1,4	Sim	1	-
6191	proteína hipotética	-	-	-	1,3	1,4	1,1	Sim	1	-
8189	proteína hipotética	-	-	-	1,3	1,3	1,2	Sim	-	-
38733	proteína hipotética	1,5	-	-	-	1,4	3,9	Sim	-	-
136540	proteína hipotética	-	-	1,3	-	-	1,3	Sim	-	-
124442	proteína hipotética	-	-	-	1,2	2,5	-	Sim	-	-
135602	proteína hipotética	-	-	-	1,2	-	1,5	Sim	-	DUF4185
92939	proteína hipotética	-	1,0	-	1,0	-	-	Sim	4	-
42618	proteína hipotética	4,3	-	-	-	-	1,8	-	-	-
122358	proteína hipotética	1,3	1,4	-	-	-	-	Sim	-	DUF1180
43045	proteína hipotética	1,2	2,0	-	-	-	-	Sim	-	CFEM domain
73201	proteína hipotética	1,9	2,1	-	-	-	-	Sim	-	-
53831	proteína hipotética	-	-	4,6	-	-	-	-	-	-



(Continuação)

ID JGI	Proteína	Log <sub>2</sub> (fold change)						SignalP	TMH	Pfam
		C6h	C12h	C24h	B6h	B12h	B24h			
80659	proteína hipotética	-	-	1,8	-	-	-	Sim	-	CFEM domain
33310	proteína hipotética	-	-	1,8	-	-	-	Sim	-	-
74442	proteína hipotética	-	-	1,3	-	-	-	Sim	-	-
74838	proteína hipotética	-	-	1,0	-	-	-	Sim	-	-
32551	proteína hipotética	-	-	-	-	1,1	-	Sim	-	-
123852	proteína hipotética	-	-	-	-	1,1	-	Sim	-	-
124417	proteína hipotética	-	-	-	-	-	1,2	Sim	-	-
100282	proteína hipotética	-	-	-	-	-	1,1	Sim	-	-
10947	proteína hipotética	-	-	-	-	-	3,7	Sim	-	-
84724	proteína hipotética	-	2,8	-	-	-	-	Sim	-	-
74517	proteína hipotética	-	1,8	-	-	-	-	Sim	-	-
92858	proteína hipotética	-	1,1	-	-	-	-	Sim	-	-
12090	proteína hipotética	-	1,1	-	-	-	-	Sim	1	-
6678	proteína hipotética	1,1	-	-	-	-	-	Sim	1	DUF3445
111165	proteína hipotética	1,2	-	-	-	-	-	Sim	-	-
102458	proteína hipotética	1,8	-	-	-	-	-	Sim	-	-
10616	proteína hipotética	2,7	-	-	-	-	-	Sim	-	-
61517	proteína hipotética	4,2	-	-	-	-	-	-	-	-



**ANEXO V.** Genes de fatores de transcrição hipotéticos sendo hiper-expressos no transcriptoma de *T. reesei* crescido em biomassa de cana-de-açúcar.

**Tabela 1.** Genes de fatores de transcrição hipotéticos sendo hiper-expressos no transcriptoma de *T. reesei* crescido em colmo (C) e bagaço (B), após 6, 12 e 24 horas (h). A identificação (ID) de cada gene é a disponível no banco de dados do JGI do *T. reesei* RUT C30 v1.0. Os domínios encontrados no banco de dados do Pfam são apresentados junto com os do banco de dados Interpro. A expressão é apresentada como o  $\log_2$  do *fold change*.

Protein ID	Pfam	InterPro	Log <sub>2</sub> (fold change)					
			C6h	C12h	C24h	B6h	B12h	B24h
97256	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,1	1,3	2,4	1,5	1,9	2,1
67471	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,2	1,6	1,7	1,7	1,9	1,8
74374	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,2	1,1	1,0	2,0	1,8	1,9
77202	Fungal specific transcription factor domain	IPR007219 Fungal specific transcription factor	1,3	1,4	1,3	3,0	3,1	2,8
128956	-	IPR007087 Zinc finger, C2H2-type	1,3	1,2	1,6	1,3	1,3	1,3
37839	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,5	1,3	3,3	3,4	3,5	3,4
85067	Homeobox domain	IPR001356 Homeobox	1,6	1,4	1,9	1,6	1,9	1,3
136533	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,6	1,4	1,5	2,2	2,2	2,0
95791	Fungal specific transcription factor domain	IPR007087 Zinc finger, C2H2-type	1,7	2,1	1,8	2,4	2,4	1,9
76250	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,7	2,1	3,4	2,3	3,1	3,4
73370	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,8	1,6	1,3	1,7	1,2	1,1
25140	-	IPR015880 Zinc finger, C2H2-like	1,8	1,6	1,7	1,7	1,2	1,8
77124	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,9	2,3	2,1	1,9	2,8	2,9
43676	Zinc finger C2H2 type domain signature.	IPR007087 Zinc finger, C2H2-type	2,6	2,2	2,9	1,8	2,3	1,6
103763	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	2,8	2,7	2,1	2,2	2,1	1,4
88058	MYND zinc finger	IPR002893 Zinc finger, MYND-type	3,0	3,4	4,3	2,0	2,5	3,2
7301	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	1,1	1,3	1,3	1,5	1,1
56646	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	2,2	-	2,5	4,6	5,0	4,0
89793	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,0	1,5	1,1	1,1	1,2	-
70232	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,3	1,0	1,6	1,1	1,6	-

**(Continuação)**

Protein ID	Pfam	InterPro	Log <sub>2</sub> (fold change)					
			C6h	C12h	C24h	B6h	B12h	B24h
90707	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,5	1,4	1,5	1,0	1,2	-
83920	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,7	1,4	1,4	1,0	1,4	-
97398	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,8	1,9	1,9	1,0	1,4	-
133808	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,8	2,0	1,5	1,3	1,8	-
77049	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,8	2,0	1,7	1,9	1,8	-
11035	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,8	1,5	1,3	1,4	1,4	-
42274	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,9	1,6	1,7	1,3	1,4	-
58034	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	2,0	1,6	1,2	1,2	1,1	-
93382	-	<i>IPR007087 Zinc finger, C2H2-type</i>	2,0	1,8	1,5	1,0	1,1	-
76745	<i>Fungal specific transcription factor domain</i>	<i>IPR021858 Protein of unknown function DUF3468</i>	2,9	3,1	2,7	1,9	1,6	-
143150	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,2	1,3	1,6	-	1,2	-
67125	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,2	1,4	1,5	-	1,2	-
91927	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,2	1,2	1,1	-	1,2	-
6520	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,3	1,6	1,1	-	1,5	-
136703	<i>Fungal specific transcription factor domain</i>	<i>IPR007219 Fungal specific transcription factor</i>	1,4	1,5	1,6	-	1,1	-
75915	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,8	1,9	2,6	-	2,0	-
134352	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	4,1	1,2	-	1,1	1,2	-
34881	<i>NmrA-like family</i>	<i>IPR008030 NmrA-like</i>	1,3	-	-	2,1	1,8	1,6
80499	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	-	-	1,0	1,4	1,6	1,3
98455	<i>Activator of cellulase expression 3, ace3</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	-	1,2	2,1	-	1,5	2,1
98788	<i>Xylanase regulator 1, xyr1</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	-	1,1	1,6	-	1,2	1,6
140537	<i>AN1 zinc finger</i>	<i>IPR000058 Zinc finger, AN1-type</i>	1,1	1,0	1,0	-	-	-
78691	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,1	1,3	1,3	-	-	-
109343	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	1,1	1,4	1,1	-	-	-

(Continuação)

Protein ID	Pfam	InterPro	Log <sub>2</sub> (fold change)					
			C6h	C12h	C24h	B6h	B12h	B24h
75696	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,3	1,0	1,0	-	-	-
83594	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,4	1,4	1,4	-	-	-
126082	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,5	1,5	1,3	-	-	-
128263	OTT_1508-like deaminase	IPR007087 Zinc finger, C2H2-type	1,8	1,8	1,6	-	-	-
128051	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	3,0	2,6	2,3	-	-	-
57466	Fungal specific transcription factor domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	-	1,1	1,0	1,2	-
52849	Zinc finger, C3HC4 type (RING finger)	IPR001841 Zinc finger, RING-type	-	-	-	1,7	1,7	1,7
141251	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,0	-	-	1,0	-	-
78617	IBR domain	IPR001841 Zinc finger, RING-type	1,0	-	-	1,0	-	-
76601	Fungal specific transcription factor domain	IPR007219 Fungal specific transcription factor	1,2	-	-	1,1	-	-
119307	AN1 zinc finger	IPR000058 Zinc finger, AN1-type	1,0	-	-	-	-	1,2
26770	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,0	-	1,2	-	-	-
10024	bZIP transcription factor	IPR004827 Basic-leucine zipper (bZIP) transcription factor	1,0	1,4	-	-	-	-
34850	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,1	1,0	-	-	-	-
128266	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,1	1,1	-	-	-	-
138343	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,2	1,1	-	-	-	-
25786	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,3	1,1	-	-	-	-
68757	Fork head domain	IPR001766 Fork head transcription factor	1,4	1,3	-	-	-	-
89409	bZIP transcription factor	IPR004827 Basic-leucine zipper domain	-	-	1,6	-	-	1,3
39271	NmrA-like family	IPR008030 NmrA-like	-	-	6,0	-	3,4	-
112768	CHY zinc finger	IPR008913 Zinc finger, CHY-type	-	-	1,3	-	1,1	-
109177	-	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	1,1	1,1	-	-	-
77614	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	1,0	1,0	-	-	-
75469	Fungal specific transcription factor domain	IPR007087 Zinc finger, C2H2-type	-	1,4	1,2	-	-	-

## (Continuação)

Protein ID	Pfam	InterPro	Log <sub>2</sub> (fold change)					
			C6h	C12h	C24h	B6h	B12h	B24h
90644	Zinc finger C-x8-C-x5-C-x3-H type (and similar)/IBR domain	IPR000571 Zinc finger, CCCH-type	-	1,0	1,3	-	-	-
100825	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	-	-	1,1	-	1,1
38522	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	-	-	2,3	2,2	-
33351	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	-	-	4,2	4,0	-
10981	bZIP transcription factor	IPR004827 Basic-leucine zipper (bZIP) transcription factor	1,0	-	-	-	-	-
92664	Fungal specific transcription factor domain	IPR007087 Zinc finger, C2H2-type	1,0	-	-	-	-	-
82207	GATA zinc finger/PAS fold	IPR000014 PAS	1,0	-	-	-	-	-
7294	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,1	-	-	-	-	-
89588	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,1	-	-	-	-	-
133387	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,4	-	-	-	-	-
69545	Homeobox domain/Zinc finger, C2H2 type	IPR001138 Fungal transcriptional regulatory protein, N-terminal	1,8	-	-	-	-	-
141272	Homeobox domain/Fungal specific transcription factor domain	IPR001356 Homeobox	2,1	-	-	-	-	-
67733	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	1,3	-	-	-	-
94420	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	1,0	-	-	-	-
76540	Homeobox domain/Fungal specific transcription factor domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	1,0	-	-	-	-
77359	Fungal specific transcription factor domain	IPR007087 Zinc finger, C2H2-type	-	1,0	-	-	-	-
137680	Helix-loop-helix DNA-binding domain	IPR001092 Basic helix-loop-helix dimerisation region bHLH	-	-	-	1,0	-	-
72675	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	-	-	1,1	-	-
139402	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	-	-	-	-	1,3
129117	Homeobox domain	IPR001356 Homeobox	-	-	-	-	-	1,6
135396	SET domain	IPR002893 Zinc finger, MYND-type	-	-	-	-	-	1,4
86314	-	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	-	-	-	1,1	-
93534	Fungal Zn(2)-Cys(6) binuclear cluster domain	IPR001138 Fungal transcriptional regulatory protein, N-terminal	-	-	-	-	1,0	-

(Continuação)

Protein ID	Pfam	InterPro	Log <sub>2</sub> (fold change)					
			C6h	C12h	C24h	B6h	B12h	B24h
89214	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	-	-	-	-	1,2	-
76518	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	-	-	1,0	-	-	-
39947	<i>Fungal Zn(2)-Cys(6) binuclear cluster domain</i>	<i>IPR001138 Fungal transcriptional regulatory protein, N-terminal</i>	-	-	1,0	-	-	-





**ANEXO VI.** Autorização da Comissão Interna de Biossegurança (CIBio) do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM) para a realização do presente projeto de mestrado.

#### DECLARAÇÃO

Declaro para os devidos fins que o conteúdo de minha dissertação de Mestrado intitulada 'Estudos genômicos da expressão gênica global do fungo filamentosso *Trichoderma reesei* crescido em bagaço e colmo de cana-de-açúcar'

( ) não se enquadra no § 4º do Artigo 1º da Informação CCPG 002/13, referente a bioética e biossegurança.

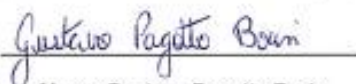
Tem autorização da(s) seguinte(s) Comissão(ões):

( X ) CIBio – Comissão Interna de Biossegurança , projeto No. JVCO 1.1, Instituição: CNPEM – Centro Nacional de Pesquisa em Energia e Materiais.

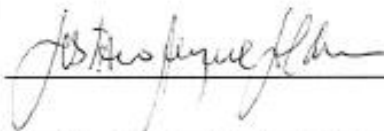
( ) CEUA – Comissão de Ética no Uso de Animais , projeto No. \_\_\_\_\_, Instituição: \_\_\_\_\_

( ) CEP - Comissão de Ética em Pesquisa, protocolo No. \_\_\_\_\_, Instituição: \_\_\_\_\_

*\* Caso a Comissão seja externa ao IBUNICAMP, anexar o comprovante de autorização dada ao trabalho. Se a autorização não tiver sido dada diretamente ao trabalho de tese ou dissertação, deverá ser anexado também um comprovante do vínculo do trabalho do aluno com o que constar no documento de autorização apresentado.*



Aluno: Gustavo Pagotto Borin



Orientador: Gustavo Henrique Goldman

Para uso da Comissão ou Comitê pertinente:

(X) Deferido ( ) Indeferido



Carimbo e assinatura

Prof. Dr. REGINA COUTINHO FRANCO DE OLIVEIRA  
Presidente  
CIBio/IB-UNICAMP

Para uso da Comissão ou Comitê pertinente:

( ) Deferido ( ) Indeferido

Carimbo e assinatura

(Continuação)

Uso exclusivo da CIBio:

Número de projeto / processo: JUCO 1.1

**Formulário de encaminhamento de projetos de pesquisa para análise pela CIBio - Comissão Interna de Biossegurança do CNPEM – Centro Nacional de Pesquisa em Energia e Materiais**

Título do projeto: Estudos genômicos da expressão gênica global do fungo filamentososo *Trichoderma reesei* crescido em bagaço e colmo de cana-de-açúcar

Pesquisador responsável: Juliana Velasco de Castro Oliveira

Experimentador: Gustavo Borin (aluno Mestrado)

Nível do treinamento do experimentador:  -Iniciação científica,  -mestrado,  -doutorado,  
 -doutorado direto,  -pós-doutorado,  -nível técnico,  -outro, especifique: \_\_\_\_\_

Resumo do projeto:

A dependência de fontes de combustíveis fósseis bem como a preocupação com o meio ambiente tem gerado mundialmente um grande interesse no desenvolvimento de outras fontes de energia. Nesse sentido, o uso de etanol como combustível é vantajoso por se tratar de uma fonte inesgotável de energia, além de ser considerada de baixo impacto ambiental. Atualmente, através da fermentação do caldo da cana-de-açúcar, o Brasil é o segundo maior produtor de bioetanol. Entretanto, estudos sugerem que se forem utilizados o bagaço e a palha da cana, o Brasil pode dobrar sua produção, sem a necessidade do aumento de áreas cultiváveis. No entanto, o uso dessa biomassa apresenta como desafio a degradação da estrutura da parede celular vegetal formada por açúcares complexos, em açúcares mais simples, fermentáveis. Para isso, são necessários diversos tipos de enzimas hidrolíticas, que naturalmente são produzidas por uma variedade de microrganismos; dentre eles destaca-se o fungo *Trichoderma reesei*, capaz de produzi-las e secretá-las em grandes quantidades. Embora diversos estudos sejam realizados com este microrganismo, ainda não são conhecidos em detalhes os mecanismos de degradação destes substratos. O presente trabalho teve como objetivo principal a análise da expressão gênica global do *T. reesei*, crescido em bagaço e colmo de cana como fontes ricas de carbono pela técnica *high-throughput* do RNA-Seq. Visa-se entender como esse fungo e seus mutantes gerados a partir de varções de deleção se comportam na degradação de substratos lignocelulósicos de interesse industrial para a produção de etanol de segunda geração. Esperamos identificar várias glicosil hidrolases e outras enzimas envolvidas direta ou indiretamente na sacarificação da biomassa, sendo expressas de modo diferencial no bagaço e no colmo em tempos diferentes. Assim, os dados que estão sendo gerados neste projeto permitirão um melhor entendimento do processo de degradação de biomassa. Todo conhecimento em âmbito molecular resultante refletirá em novas perspectivas para o desenvolvimento de coquetis enzimáticos mais eficazes a custos mais baixos, que poderá possibilitar a produção de etanol de segunda geração em escala industrial.

A CIBio analisou este projeto em reunião realizada no dia: 17.1.2013.

Parecer final:  -projeto aprovado,  -projeto recusado,  -projeto com deficiências, favor comentários anexo.

Presidente de CIBio – CNPEM-LNBio  
Jörg Kobarg

Membro da CIBio – CNPEM-LNBio  
Prof. Dr. Celso Eduardo Benedetti

Membro da CIBio – CNPEM-LNBio  
Andrea Balan

Membro da CIBio – CNPEM-SGT  
Carolina Barbosa Marini

Membro da CIBio da CNPEM-CTBE  
Fabio Squina

Membro da CIBio da CNPEM-CTBE  
Sindelia Freitas Azzoni

Membro da CIBio da CNPEM-CTBE  
Roberto Ruller