

**Laboratório de Biologia Molecular**  
**Departamento de Genética e Evolução**  
**Universidade Federal de São Carlos**

**Wilson Malagó Junior**

**“Clonagem e Estudos de Expressão de Enzimas do  
Fungo Filamentoso *Trichoderma harzianum* IOC-3844  
Envolvidas na Degradação de Biomassa“**

**São Carlos - SP, junho de 2012**

**“Clonagem e Estudos de Expressão de Enzimas do  
Fungo Filamentoso *Trichoderma harzianum* IOC-3844  
Envolvidas na Degradação de Biomassa“**

**Laboratório de Biologia Molecular**  
**Departamento de Genética e Evolução**  
**Universidade Federal de São Carlos**

**Tese de Doutorado**

**“Clonagem e Estudos de Expressão de Enzimas do  
Fungo Filamentoso *Trichoderma harzianum* IOC-3844  
Envolvidas na Degradação de Biomassa“**

**Wilson Malagó Junior**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutor em Ciências, área de concentração Bioquímica e Biologia Molecular.

**Orientador: Prof. Dr. Flávio Henrique da Silva**

São Carlos - SP, junho de 2012

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária/UFSCar**

M236ce

Malagó Junior, Wilson.

Clonagem e estudos de expressão de enzimas do fungo filamentoso *Trichoderma harzianum* IOC-3844 envolvidas na degradação de biomassa / Wilson Malagó Junior. -- São Carlos : UFSCar, 2012.

91 f.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2012.

1. Fungos. 2. *Trichoderma harzianum*. 3. Celulase. 4. Hemicelulase. 5. Transcriptoma. 6. Expressão heteróloga. I. Título.

CDD: 589.2 (20<sup>a</sup>)

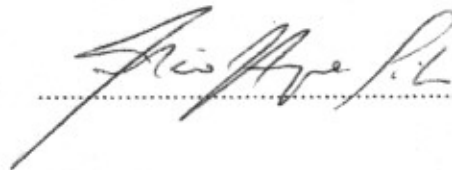
UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA MOLECULAR

"Clonagem e Estudos de Expressão de Enzimas do Fungo Filamentoso  
*Trichoderma harzianum* IOC-3844 Envolvidas na Degradação de Biomassa"

Tese de Doutorado de  
**WILSON MALAGÓ JUNIOR**

**Banca Examinadora:**

Prof. Dr. Flavio Henrique Silva



Prof. Dr. Vadim Viviani



Profa. Dra. Cristiane Sanchez Farinas



Prof. Dr. Jesus Aparecido Ferro



Prof. Dr. Roberto Ruller



## **AGRADECIMENTOS**

Aos meus pais e familiares pelo incentivo e felicidade compartilhada no doutoramento.

A minha namorada Dra. Poliana querida pelo companheirismo, incentivo e compreensão.

Ao amigo professor Dr. Flávio Henrique da Silva do Departamento de Genética e Evolução (DGE) da UFSCar, grande companheiro, pela amizade, pelos muitos anos de convivência, pela orientação, pelo apoio e pelos trabalhos desenvolvidos.

A amiga professora Dra. Andrea Soares-Costa do Departamento de Genética e Evolução (DGE) da UFSCar pela amizade, pelo apoio e por todos esses anos de convivência.

Aos amigos Dra. Ludier Kesser Santos Silva, Dr. Fernando Fonseca Pereira de Paula e MSc. Wesley Cardoso Generoso, do Laboratório de Biologia Molecular do Departamento de Genética e Evolução (DGE) da UFSCar, pela amizade e convivência.

Ao amigo Dr. Luciano Takeshi Kishi do Laboratório de Biologia Molecular do Departamento de Genética e Evolução (DGE) da UFSCar pela ajuda na submissão dos ESTs.

A todos os amigos do Laboratório de Biologia Molecular do Departamento de Genética e Evolução (DGE) da UFSCar, pela amizade e convivência.

Ao amigo Dr. César Adolfo Sommer da Seção de Gastroenterologia do Departamento de Medicina da “Boston University School of Medicine” pela amizade, convivência e pelos artigos enviados.

Aos colegas do Departamento de Genética e Evolução (DGE) da UFSCar, pela convivência.

Ao Programa de Pós-graduação em Genética e Evolutiva e Biologia Molecular (PPG-Gev) da UFSCar pelo apoio.

A todos os professores do Programa de Pós-graduação em Genética e Evolutiva e Biologia Molecular (PPG-Gev) da UFSCar pela convivência e pelo conhecimento transmitido.

As secretarias do Departamento de Genética e Evolução (DGE) e do Programa de Pós-graduação em Genética e Evolutiva e Biologia Molecular (PPG-Gev) da UFSCar pelo apoio.

Ao CNPq, FAPESP e CAPES pelo auxílio financeiro.

Ao professor Dr. Nei Pereira Junior da Escola de Química da UFRJ pela convivência e parceria.

Aos colegas do Laboratório de Bioprocessos (LADEBIO) da Escola de Química da UFRJ pela convivência e parceria.

Ao professor Dr. Igor Polikarpov do Instituto de Física da USP de São Carlos (IFSC) pela convivência e parceria.

Ao amigo MSc. Marcelo Vizoná Liberato do Laboratório de Cristalografia do Instituto de Física da USP de São Carlos (IFSC) pela amizade, convivência e parceria.

Aos amigos do Laboratório de Cristalografia do Instituto de Física da USP de São Carlos (IFSC) pela convivência e parceria.

Ao professor Dr. Marcio de Castro Silva Filho da Escola Superior de Agricultura Luiz de Queiroz (ESALQ) da USP, pela parceria.

A Dra. Lígia Hansen Arruda da Escola Superior de Agricultura Luiz de Queiroz (ESALQ) da USP, pela convivência e parceria.

Ao professor Dr. Paulo Inácio da Costa da UNESP de Araraquara pela parceria.

A Dra Thais Ferreira Isabel da UNESP de Araraquara pela convivência e parceria.

A MSc. Ana Carolina Urbaczek da UNESP de Araraquara pela convivência e parceria.

A MSc. Thalita Athiê Neo da UNESP de Araraquara pela convivência e parceria.

Ao professor Dr. Reinaldo Otávio Alvarenga Alves de Brito, do Departamento de Genética e Evolução (DGE) da UFSCar, pela convivência e parceria.

A MSc. Vanessa Regina Gonçalves do Departamento de Genética e Evolução (DGE) da UFSCar pela convivência e parceria.

A MSc. Amalia Nelci Teresita Pascual do Departamento de Genética e Evolução (DGE) da UFSCar pela convivência e parceria.

A professora Dra. Luciana Correia de Almeida Regitano do Centro de Pesquisa de Pecuária do Sudeste (EMBRAPA-CPPSE) pelo apoio.

A professora Dra. Dulce Helena Ferreira de Souza do Departamento de Química (DQ) da UFSCar pela convivência e parceria.

A MSc. Camilla Calemi Golfeto do Departamento de Química (DQ) da UFSCar pela convivência e parceria.

A professora Dra. Eleni Gomes da UNESP de São José do Rio Preto pela parceria.

A Dra. Ana Flávia Azevedo Carvalho da UNESP de Rio Claro pela convivência e parceria.



## RESUMO

A biomassa vegetal é um recurso disponível em larga escala e uma das suas mais importantes aplicações é a produção de etanol de segunda geração. No entanto, o custo das enzimas é um dos maiores entraves para a produção economicamente viável deste etanol. Neste contexto, é importante encontrar organismos produtores de grandes quantidades de enzimas que degradam a biomassa. Os objetivos deste estudo foram clonar, estudar a expressão gênica e caracterizar o conjunto de enzimas que degradam a biomassa vegetal, do fungo filamentosso *Trichoderma harzianum* IOC-3844. Um total de 1.543 seqüências de boa qualidade, geradas a partir de uma biblioteca de cDNA do *Trichoderma harzianum* IOC-3844, induzido por celulose, foi organizado em 1.002 transcritos, sendo 167 representados por mais de uma seqüência e 835 representados por apenas uma seqüência. Destes transcritos, 356 tiveram função associada e 646 não tiveram. Com isso, entre os transcritos com função associada, foram listados 20 transcritos envolvidos com degradação de biomassa vegetal. Análises de PCR em tempo real do micélio de *Trichoderma harzianum* IOC-3844, crescido por 36 e 60 horas em celulose, mostraram níveis de mRNA mais de 2.000 vezes mais representados para os transcritos *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl7* e *swol*, quando comparados com o micélio não induzido. Em alguns casos as maiores representatividades alcançaram valores superiores a 100.000 vezes. Entre os transcritos analisados o *cbh1*, o *cbh2* e o *egl7*, mostraram os mais altos níveis de expressão. O *Trichoderma harzianum* IOC-3844 exibiu um repertório com alta expressão de transcritos envolvidas na degradação de biomassa vegetal. As enzimas EGIII e Xyn2 foram expressas em sistema recombinante com uso da levedura *Pichia pastoris*, apresentando facilidade de purificação e boa atividade enzimática. Os ensaios de expressão heteróloga viabilizaram estudos posteriores que visam a aplicação industrial das enzimas. Assim, esta cepa mostrou potencial para produzir enzimas que degradam a biomassa para a produção de etanol de segunda geração, e para ser fonte de enzimas para a indústria de papel.

## PALAVRAS-CHAVE

*Trichoderma harzianum*, Transcriptoma, Celulases, Hemicelulases, Xilanases, Biomassa, Biblioteca de cDNA, Etanol de segunda geração, Expressão heteróloga.

## ABSTRACT

The plant biomass is a large-scale available resource and one of its more important applications is the second-generation ethanol production. However, the enzyme cost is one of the biggest barriers for economically viable ethanol from biomass. Therefore, it is important to identify fungal strains that can produce high concentrations of plant biomass-degrading enzymes. The aim of this work was to clone, study the gene expression and characterize the plant biomass-degrading transcript set of the filamentous fungus *Trichoderma harzianum* IOC-3844. A total of 1,543 high-quality reads from the *Trichoderma harzianum* IOC-3844 cellulose induced cDNA library were organized into 1,002 transcripts representing 167 contigs and 835 singlets. Of these 1,002 transcripts 646 had unknown functions and 356 showed associated functions. Among the transcripts with associated functions, we found 20 transcripts related to plant biomass deconstruction. The real time PCR analysis of *Trichoderma harzianum* IOC-3844 mycelia grown for 36 and 60 hours in cellulose, revealed that the levels of the following mRNAs were induced by at least 2,000-fold when compared to uninduced mycelia: *cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl7* and *swol*. In some cases, the values were higher than 100,000-fold. Among the transcripts analyzed by real time PCR, *cbh1*, *cbh2* and *egl7* exhibited the highest expression levels. The *Trichoderma harzianum* IOC-3844 exhibited a repertoire with high expression of plant biomass-degrading transcripts. The enzymes EGIII and Xyn2 were recombinantly expressed in *Pichia pastoris*, showing good quality purification and good enzymatic activity. The heterologous expression assays made possible future studies aiming at the industrial application of the enzymes. Therefore, this strain showed potential to produce biomass-degrading enzymes for second-generation ethanol production and to be a source of enzymes for the paper industry

## KEY WORDS

*Trichoderma harzianum*, Transcriptome, Cellulases, Hemicellulases, Xylanases, Biomass, cDNA library, Second-generation ethanol, Heterologous expression.

## ABREVIATURAS

Abf ou *abf*: Arabinofuranosidase

ACE ou *ace*: Fator de transcrição tipo “zinc finger”

AFEX: Método de expansão de fibra com amônia

Axe ou *axe*: Acetil xilano esterase

BGL ou *bgl*:  $\beta$ -glucosidase

BLAST: Ferramenta de busca por alinhamento

Bxl ou *bxl*:  $\beta$ -xilosidase

CBH ou *cbh*: Celobiohidrolase

CBM: Módulo de ligação de carboidrato

Cip ou *cip*: Proteína glicolítica candidata ligante de celulose

CMC: Carboximetilcelulose

CRE1 ou *cre1*: Repressor catabólico

dCAS: Sistema de anotação de cDNA para “desktop”

EG ou *egl*: Endoglucanase

EST: Etiqueta de seqüência expressa

FPase: Atividade de papel de filtro

Man1 ou *man1*:  $\beta$ -mananase

ORF: Fase aberta de leitura

PDA: Agar de batata

PDB: Banco de dados de proteínas PDB

RT-PCR: PCR com uso de transcriptase reversa

SDS-PAGE: Eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio

SHF: Hidrólise e fermentação separadas

SSF: Sacarificação e fermentação simultâneas

SWO ou *swo*: Swolenina

UniProtKB: Banco de dados de proteínas UniProtKB

Xyn ou *xyn*: Endoxilanasase

Xyr ou *xyr*: Regulador de xilanasase

## LISTA DE TABELAS

---

<b>Tabela 1:</b> Resumo sobre as clonagens de celulases e hemicelulases de <i>Trichoderma harzianum</i> IOC-3844 em plasmídeos para expressão em <i>Pichia pastoris</i> .	29
<b>Tabela 2:</b> Transcritos de <i>Trichoderma harzianum</i> IOC-3844 relacionados a desconstrução de biomassa vegetal.	37
<b>Tabela 3:</b> Comparação de abundância de transcritos, envolvidos com degradação de biomassa vegetal, entre <i>Trichoderma harzianum</i> IOC-3844 e <i>Trichoderma reesei</i> QM6a.	40
<b>Tabela 4:</b> ESTs publicamente disponíveis de <i>Trichoderma harzianum</i> associados com degradação de biomassa vegetal.	42
<b>Tabela 5:</b> Comparação entre os dados de expressão de PCR em tempo real e a contagem de ESTs de cultura de <i>Trichoderma harzianum</i> IOC-3844 induzida por 60 horas.	62

---

## LISTA DE FIGURAS

---

<b>Figura 1:</b> Gel de agarose 1 % corado com brometo de etídeo contendo RNA total de micélio de <i>Trichoderma harzianum</i> IOC-3844 cultivado por 60 horas.	32
<b>Figura 2:</b> Gel de agarose 1 % corado com brometo de etídeo contendo RNA total extraído de micélio de <i>Trichoderma harzianum</i> IOC-3844 e tratado com DNaseI.	33
<b>Figura 3:</b> Gel de agarose 1 % corado com brometo de etídeo com resultado de RT-PCR em micélio de <i>Trichoderma harzianum</i> IOC-3844 cultivado por 60 horas.	34
<b>Figura 4:</b> (A) Gel de agarose 1% corado com brometo de etídeo com resultado de extração de RNA mensageiro de <i>Trichoderma harzianum</i> IOC-3844 induzido com celulose por 60 horas, (B) Gel de agarose 1% corado com brometo de etídeo com resultado de análise de clones da biblioteca por restrição com <i>Bsr</i> GI.	35
<b>Figura 5:</b> Valores de expressão de transcritos obtidos de cultura induzida de <i>Trichoderma harzianum</i> IOC-3844 por 8, 36 e 60 horas, em relação aos transcritos obtidos de cultura não induzida de 8, 36 e 60 horas.	43
<b>Figura 6:</b> Valores de expressão de transcritos obtidos de cultura induzida de <i>Trichoderma harzianum</i> IOC-3844 por 24, 36, 48 e 60 horas, comparados com os transcritos obtidos de cultura do fungo induzido por 8 horas.	44
<b>Figura 7:</b> Valores de expressão de transcritos obtidos de cultura induzida de <i>Trichoderma harzianum</i> IOC-3844 por 8, 24, 36, 48 e 60 horas, comparados com a média dos valores de expressão do transcrito de actina nos tempos de 8, 24, 36, 48 e 60 horas de cultura do fungo induzido.	46
<b>Figura 8:</b> Sobrenadante de cultura de <i>Pichia pastoris</i> expressando EGIII de <i>Trichoderma harzianum</i> IOC-3844.	47
<b>Figura 9:</b> Frações purificadas do sobrenadante do clone B2 de <i>Pichia pastoris</i> transformada com EGIII de <i>Trichoderma harzianum</i> IOC-3844.	48
<b>Figura 10:</b> Sobrenadante de cultura de <i>Pichia pastoris</i> expressando Xyn2 de <i>Trichoderma harzianum</i> IOC-3844.	49
<b>Figura 11:</b> Teste de atividade das frações purificadas do sobrenadante da indução do clone B2 de <i>Pichia pastoris</i> transformada com EGIII de <i>Trichoderma harzianum</i> IOC-3844. Ensaio realizado em placas de CMC reveladas com Vermelho Congo.	50
<b>Figura 12:</b> Teste de atividade dos sobrenadantes de culturas induzidas de <i>Pichia pastoris</i> transformadas com Xyn2 de <i>Trichoderma harzianum</i> IOC-3844, em placas de Xilano reveladas com Vermelho Congo.	50

---

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	1
1.1 Etanol de segunda geração e biomassa vegetal	1
1.2 Processamento de biomassa vegetal	3
1.3 O gênero <i>Trichoderma</i>	5
1.4 Celulases	6
1.5 Hemicelulases	8
1.6 Ligninases	9
1.7 Enzimas acessórias	10
1.8 Principais desafios para desconstrução de biomassa vegetal e produção de etanol	11
1.9 O fungo <i>Trichoderma harzianum</i>	14
1.10 Produção científica	17
<b>2 OBJETIVOS</b>	18
<b>3 MÉTODOS</b>	19
3.1 Inóculo e cultura de <i>Trichoderma harzianum</i> IOC-3844	19
3.2 Coletas de micélio	19
3.3 Extração de RNA total	20
3.4 Extração de RNA mensageiro	20
3.5 Análises de RT-PCR	21
3.6 Construção e seqüenciamento da biblioteca de cDNA de <i>Trichoderma harzianum</i> IOC-3844	22
3.7 Análise das bibliotecas de cDNA	23
3.8 Ensaio de PCR em tempo real	25
3.9 Clonagem de enzimas de <i>Trichoderma harzianum</i> IOC-3844 em plasmídeos para expressão em <i>Pichia pastoris</i>	28
3.10 Expressão de enzimas de <i>Trichoderma harzianum</i> IOC-3844 em <i>Pichia pastoris</i>	29
3.11 Ensaio de atividade de enzimas recombinantes de <i>Trichoderma harzianum</i> IOC-3844	31
<b>4 RESULTADOS</b>	32
4.1 Extração de RNA total	32

4.2	Análise de expressão de celulasas por RT-PCR	33
4.3	Extração de RNA mensageiro e construção da biblioteca de cDNA	34
4.4	Análise de sequencias ESTs	35
4.5	Busca de seqüências de <i>Trichoderma harzianum</i> relacionadas com degradação de biomassa vegetal	41
4.6	PCR em tempo real e análises de expressão	42
4.7	Expressão de proteínas recombinantes	46
4.8	Ensaio de atividade das enzimas recombinantes	49
<b>5</b>	<b>DISCUSSÃO</b>	51
5.1	Análise inicial da expressão de celulasas por RT-PCR	51
5.2	Contagem de ESTs	51
5.3	Busca de seqüências de <i>Trichoderma harzianum</i> relacionadas com degradação de biomassa vegetal	59
5.4	PCR em tempo real e análises de expressão	60
5.5	Expressão e ensaios de atividade das proteínas recombinantes	64
<b>6</b>	<b>CONCLUSÕES</b>	69
<b>7</b>	<b>PERSPECTIVAS FUTURAS</b>	70
<b>8</b>	<b>REFERÊNCIA</b>	71

# 1 INTRODUÇÃO

## 1.1 Etanol de segunda geração e biomassa

A produção de biocombustíveis renováveis como o etanol, em especial o etanol celulósico, obtido da fermentação de açúcares oriundos da sacarificação de biomassa vegetal, apresenta potencial para substituir os combustíveis fósseis resultando na redução das emissões de gases de efeito estufa. O uso de biocombustíveis também contribui para contornar os problemas de dependência de importação, instabilidade de mercado e limitação das reservas mundiais de combustíveis fósseis (NEHRING, 2009; HILL et al., 2009; SLADE, BAUEN, SHAH, 2009; MACLEAN & LAVE, 2003). Neste contexto, além do processo tradicional de produção de etanol a partir da fermentação da sacarose presente no caldo de cana, a produção de etanol de segunda geração, a partir de biomassas residuais (material lignocelulósico) oriundas da agricultura e da agro indústria, é uma alternativa promissora para atender a crescente demanda mundial por combustíveis (BALAT & BALAT, 2009). A biomassa residual, é o composto orgânico mais abundante da biosfera e é quase totalmente composta por parede celular vegetal que, por sua vez, é constituída principalmente de celulose, de 40,6 a 51,2 %; hemicelulose, de 28,5 a 37,2 % e lignina, de 13,6 a 28,1 % (PAULY & KEEGSTRA, 2008).

A celulose é formada por cadeias polissacarídicas constituídas por moléculas de glicose, unidas entre si por meio de ligações do tipo  $\beta$ -1,4. O dissacarídeo celobiose é a unidade repetitiva da cadeia. Estas cadeias se dispõem de modo alinhado e formam um complexo de fibrilas organizadas. Aproximadamente 30 moléculas de celulose são montadas em unidades maiores constituindo as fibrilas elementares, que são empacotadas em unidades maiores chamadas microfibrilas; estas por sua vez são montadas para formar a fibra celulósica (LYND et al., 2002). As fibras celulósicas não são puramente cristalinas. O grau de cristalinidade é variável formando um gradiente desde o puramente cristalino para o puramente amorfo, com todos os graus de ordem entre os extremos (MARCHESSAULT & HOWSMON, 1957). Sustentada por ligações de hidrogênio inter e intracadeias, a celulose é altamente coesiva e possui resistência excepcional permanecendo insolúvel em água, apesar da hidrofiliidade de seus átomos de hidrogênio. Por sua estrutura rígida, a celulose apresenta resistência ao ataque hidrolítico e sua digestão é naturalmente lenta. Como exemplo, as gramíneas são



digeridas mais lentamente por animais herbívoros do que a carne o é por animais carnívoros; também, árvores mortas são degradadas muito mais lentamente por microorganismos do solo do que os animais o são (VOET & VOET, 1995). A situação não é diferente em ambientes não naturais, havendo também dificuldades para a hidrólise de celulose em escala industrial.

As cadeias de hemicelulose se associam fortemente as microfibrilas de celulose, se conectando às mesmas por múltiplas ligações de hidrogênio. Deste modo, a hemicelulose proporciona o ponto de contato na parede celular entre as microfibrilas de celulose e o restante dos polímeros da matriz (ALBERSHEIM et al., 2010). Ao contrário da celulose, a hemicelulose apresenta natureza heteropolissacarínica e considerável grau de ramificação. Isso faz com que não haja regiões cristalinas e assim as fibras são mais suscetíveis à hidrólise. Suas unidades fundamentais são basicamente os monômeros: D-xilose, D-manose, D-galactose, D-glicose, L-arabinose, D-ácido glucorônico e acetatos; destacando-se pela alta abundância a manana e principalmente o xilano (PEREIRA, COUTO, ANNA, 2008; POLIZELI et al., 2005; SAMIRIT & JAGDEEP, 2008).

A lignina, também responsável pela robustez da parede celular, é um polímero fenólico complexo e apresenta ligações covalentes com os polissacarídeos contribuindo para dificultar a hidrólise enzimática (LIYAMA, LAM, STONE, 1994; CAMPBELL & SEDEROFF, 1996). As ligações covalentes entre os polissacarídeos da parede celular têm sido reportadas como ocorrendo quase que na totalidade entre lignina e hemicelulose, porém também há evidências de que, em menor grau, existam ligações covalentes entre lignina e celulose (EK, GELLERSTEDT, HENRIKSSON, 2009; ERIKSSON, GORING, LINDGREN, 1980). Neste contexto, a cristalinidade da celulose, a cobertura da hemicelulose e a proteção da lignina proporcionam pouca acessibilidade de superfície e contribuem para a resistência da biomassa a hidrólise enzimática (MOSIER et al., 2005).

A biomassa é o único recurso renovável disponível em larga escala que pode ser convertido em combustíveis líquidos de fácil transporte e distribuição (WYMAN, 2007). Entretanto, a produção de bioetanol oriundo de biomassa residual se encontra em um nível de desenvolvimento tecnológico insatisfatório, sendo necessário melhorar os processos para permitir sua viabilidade econômica no cenário mundial (SERVICE, 2010). O Brasil se encontra em posição estratégica para investir nesta tecnologia por ter tradição na produção e distribuição de etanol e por ter disponibilidade de biomassa

residual a baixo custo, oriunda da agricultura e da agro indústria, destacando-se principalmente o bagaço de cana (PEREIRA, COUTO, ANNA, 2008). De fato o Brasil é favorecido, pois o processo da produção de etanol a partir de bagaço de cana pode ser anexado às unidades de produção de açúcar e álcool existentes, o que caracteriza baixa demanda de investimento, infra-estrutura, logística e fornecimento de energia. Além disso, o Brasil possui espaço territorial suficiente para produção de alimentos e combustíveis sem ocupar novas áreas para agricultura. Por outro lado, apesar da hidrólise de bagaço de cana para produção de etanol estar consolidada em condições de laboratório, a aplicação em larga escala ainda não é economicamente viável no Brasil (SOCCOL et al., 2010).

## **1.2 Processamento de biomassa**

Entre os pré-tratamentos mais utilizados para o início do processamento da biomassa na produção de etanol de segunda geração, se encontra o tradicional e bem aceito método de explosão a vapor ou “steam explosion” (SOCCOL et al., 2010), que desorganiza o complexo lignocelulósico dissolvendo a hemicelulose e a despolimerizando quase que totalmente (ZIMBARDI, RICCI, BRACCIO, 2002; HENDRIKS & ZEEMAN, 2009). Com isso, as pentoses, na maioria xilanas e mananas, são disponibilizadas diretamente para fermentação alcoólica, que se dá comumente com uso da levedura *Pichia stipitis* (OLOFSSON, BERTILSSON, LIDÉN, 2008; VAN VLEET & JEFFRIES, 2009). Do material restante, composto por lignina e celulose, também chamado de lignocelulose, a lignina pode ser separada (deslignificação) com uso de solventes orgânicos (ZHAO, CHENG, LIU, 2009). Permanece então a celulose, com digestibilidade aumentada, para ser submetida à hidrólise enzimática pela ação das celulasas. A glicose resultante é então submetida à fermentação alcoólica pela levedura *Saccharomyces cerevisiae*.

Um outro tipo de pré-tratamento, denominado “ammonia fibre expansion” (AFEX) pode ser usado, no entanto, este procedimento não extrai a hemicelulose, tornando necessário o uso de enzimas (hemicelulasas) para a despolimerização (GAO et al., 2011). Com isso, o custo de produção das hemicelulasas comprometeria ainda mais a viabilidade econômica do etanol de segunda geração. Neste contexto, devido à possibilidade de separação e degradação da hemicelulose no pré-tratamento, as hemicelulasas não têm tido tanta importância quanto as celulasas na produção de etanol

(PEREIRA, COUTO, ANNA, 2008). Além disso, segundo (SOCCOL et al., 2010) as melhores condições de pré-tratamento são aquelas que proporcionam máxima solubilização de açúcares de hemicelulose, juntamente com a melhor produção possível de substrato para a hidrólise enzimática e fermentação.

Alternativamente, a biomassa pode ser abordada pelo método da hidrólise ácida. Este processo separa a lignina da celulose (deslignifica a lignocelulose) e ao mesmo tempo, hidrolisa a celulose, com o uso de fortes ácidos minerais em altas temperaturas (HARRIS, 1975). No entanto, este procedimento foi abandonado por gerar substâncias tóxicas que inibem a fermentação da glicose (PEREIRA, COUTO, ANNA, 2008). Segundo (WYMAN, 2007), a utilização de celulases para hidrolisar a celulose é muito mais promissora do que a utilização de hidrólise ácida.

Outra possibilidade para degradação da lignina seria o pré-tratamento biológico. Existem organismos que degradam a lignina por meio de um complexo enzimático extracelular, e que poderiam ser usados para deslignificar a biomassa vegetal, disponibilizando a celulose. Entre eles, o mais estudado é o basidiomiceto filamentosso *Phanerochaete chrysosporium* (MARTINEZ et al., 2004; KERSTEN & CULLEN, 2007). No entanto, a expressão e manipulação das ligninases, para atender esta demanda, tem tido pouco sucesso (US-DOE, 2006). De fato, a degradação biológica de lignina é complexa e depende de fatores como a cepa de fungo utilizada, as condições de cultura, a secreção enzimática e os mecanismos oxidativos (GUILLÉN et al., 2000; WAN & LI, 2010). Além disso, nem sempre os maiores degradadores de lignina proporcionam os melhores rendimentos na conversão de biomassa vegetal em glicose (CAPELARI & TOMÁS-PEJÓ, 1997; SHI, CHINN, SHARMA-SHIVAPPA, 2008). Em estudos recentes foi analisada a viabilidade do uso de 21 basidiomicetos secretores de ligninases, entre eles o *Phanerochaete chrysosporium*, para o pré-tratamento biológico de material lignocelulósico na produção de etanol de segunda geração a partir de resíduos de trigo. Entre os organismos estudados, os autores sugeriram apenas o uso dos fungos *Irpex lacteus* e *Poria subvermispora* para o pré-tratamento biológico, porém, ressaltaram que sua aplicabilidade ainda deve ser otimizada e testada para uso em escala industrial (SALVACHÚA et al., 2011).

Neste contexto, o principal objetivo do pré-tratamento é aumentar a área de superfície acessível as celulases, ajudar a descristalinizar a celulose e remover a hemicelulose e a lignina (BRODEUR et al., 2011). Depois da remoção de hemicelulose e lignina, aplica-se um dos procedimentos mais aceitos chamado de sacarificação e

fermentação simultâneas (“Simultaneous Sacharification and Fermentation” - SSF), no qual a hidrólise enzimática da celulose ocorre simultaneamente à fermentação alcoólica da glicose. As principais vantagens deste método, em relação ao procedimento de hidrólise e fermentação separadas, (“Separate Hydrolysis and Fermentation” - SHF), se caracterizam pela diminuição de custo, por utilizar apenas um reator, e pela redução da inibição das celulasas pelos produtos finais da hidrólise enzimática (celobiose e glicose), removendo os mesmos e assim alterando o equilíbrio da reação de hidrólise (OLOFSSON, BERTILSSON, LIDÉN, 2008; PEREIRA, COUTO, ANNA, 2008). Além disso, neste processo a produção de etanol é mais rápida e o risco de contaminação é menor devido à sua própria presença e pela remoção da glicose. Por outro lado, o processo SSF apresenta dificuldade oriunda das diferenças entre as temperaturas ótimas de hidrólise enzimática (45-50 °C) e de fermentação alcoólica (28-35 °C) (SOCCOL et al., 2010).

### **1.3 O gênero *Trichoderma***

O *Trichoderma* é um gênero de fungos filamentosos com capacidade para degradar celulose e hemicelulose, e comumente encontrado em material vegetal caído no solo e na rizosfera (SCHUSTER & SCHMOLL, 2010). Os fungos deste gênero são também patógenos oportunistas de mamíferos imunodeficientes, incluindo humanos (KREDICS et al., 2003). Estes fungos são usados para aumentar a produtividade agrícola, no controle biológico de plantas, pois atuam como simbiontes avirulentos, causando mudanças significativas no proteoma e metabolismo, induzindo resistência sistêmica contra fitopatógenos, além de atuar na captação de nutrientes (YEDIDIA, BENHAMOU, CHET, 1999; HARMAN et al., 2004). Outrossim, a produção de antibióticos e enzimas hidrolíticas destes fungos também são fatores que permitem sua aplicação contra patógenos de plantas (GHISALBERTI & SIVASITHAMPARAM, 1991; GOLDMAN, HAYES, HARMAN, 1994).

O *Trichoderma reesei* (estágio assexuado anamorfo), classificado também como *Hypocrea jecorina* em seu estágio sexuado ou teleomorfo (KUHLS et al., 1996), hidrolisa a celulose por meio da produção substancial de celulasas extracelulares que podem ser recuperadas dos sobrenadantes das culturas (LYND et al., 2002). Apesar de estarem surgindo cada vez mais alternativas ao *T. reesei* (GUSAKOV, 2011; CASTRO et al., 2010a; CASTRO et al., 2010b), este fungo tem sido a principal fonte de

preparados enzimáticos utilizados no processamento de biomassa vegetal, e suas celulases têm sido foco de pesquisa nos últimos 50 anos (MANDELS & REESE, 1957; LYND et al., 2002; KABEL et al., 2006; SCHUSTER & SCHMOLL, 2010).

O gênero *Trichoderma* é também um bom produtor de hemicelulases, e tem sido usado como fonte destas enzimas por causa da aplicabilidade e importância das mesmas na indústria de papel (STALBRAND et al., 1993; XIONG et al., 2004; POLIZELI et al., 2005; WONG, TAN, SADDLER, 1988).

#### **1.4 Celulases**

Os microorganismos que obtêm energia da celulose precisam produzir celulases extracelulares, livres ou associadas às células, para hidrolisarem e metabolizarem este polímero insolúvel. Os tipos de celulases existentes atuam de modo sinérgico e apresentam a capacidade de hidrolisar ligações glicosídicas tipo  $\beta$ -1,4 entre os resíduos glicosil (LYND et al., 2002).

Estas enzimas são classificadas em três grupos: Exoglucanases, Endoglucanases e  $\beta$ -glucosidases. As exoglucanases agem na região cristalina da celulose processando as extremidades dos polissacarídeos. Elas são constituídas principalmente por dois grupos: as glucanohidrolases, que liberam glicose; e as celobiohidrolases (CBHs), que liberam celobiose. Existem dois tipos de celobiohidrolases: CBHI ou Cel7a que hidrolisa extremidades redutoras das cadeias de polissacarídeo e CBHII ou Cel6a, que hidrolisa extremidades não redutoras. As endoglucanases (EGs) clivam em sítios amorfos internos da cadeia polissacarídica da celulose (celulose não cristalina), gerando oligossacarídeos de vários tamanhos (celodextrinas ou celo-oligossacarídeos). As  $\beta$ -glucosidases (BGLs) hidrolisam oligossacarídeos solúveis (celodextrina e celobiose) e liberam glicose (LYND et al., 2002).

Quando agem em conjunto, portanto de modo sinérgico, as celulases apresentam melhor rendimento do que a soma dos rendimentos individuais. Existem pelo menos três formas de sinergia: “endo-exo”, em que as endoglucanases disponibilizam extremidades redutoras e não redutoras do polímero que sofrem a ação da CBHI e da CBHII respectivamente; “exo-exo”, entre exoglucanases processando as duas extremidades da cadeia polissacarídica; e “exo-bg”, em que as celobiohidrolases liberam celobiose que sofre ação da  $\beta$ -glucosidase (LYND et al., 2002).

A atividade celulolítica do fungo *T. reesei* é dada pela produção de pelo menos duas exoglucanases: CBHI ou Cel7a e CBHII ou Cel6a; cinco endoglucanases: EGI ou Cel7b, EGII ou Cel5a, EGIII ou Cel12a, EGIV ou Cel61a e EGV ou Cel45a; e duas  $\beta$ -glucosidases BGL1 ou Cel3a e BGL2 ou Cel1a (LYND et al., 2002). Entretanto, (FOREMAN et al., 2003) descreveram ainda mais três endoglucanases preditas: EGVI ou Cel74a, EGVII ou Cel61b e EGVIII ou Cel5b e mais cinco seqüências putativas de  $\beta$ -glucosidases: BGL3 ou Cel3b, BGL4 ou Cel3c, BGL5 ou Cel1b, BGL6 ou Cel3d e BGL7 ou Cel3e. Com exceção das enzimas BGL2, BGL4, BGL5 e BGL6, todas as outras contêm seqüência sinal para secreção celular (FOREMAN et al., 2003).

Como característica geral as CBHs e EGs apresentam uma estrutura modular que freqüentemente inclui os domínios catalíticos e os módulos de ligação de carboidratos (CBMs). O CBM permite a ligação da enzima à superfície da celulose por forças interativas não covalentes específicas, presumidamente facilitando a hidrólise pelo domínio catalítico, por aproximar este ao substrato (TEERI et al., 1998). Experimentos mostraram que o módulo CBM da CBHI se liga à superfície da celulose de modo reversível, exibindo uma taxa de re-posicionamento compatível com a taxa de hidrólise do domínio catalítico e explicando assim ação de natureza processiva das CBHs. O CBM, juntamente com o *linker* que o une ao domínio catalítico, mediam um processamento dinâmico do domínio catalítico em movimento na superfície da celulose (LINDER & TEERI, 1996; DIVNE et al., 1998).

A indução dos genes que codificam as enzimas extracelulares envolvidas na biodegradação da biomassa vegetal (celulases e hemicelulases) é dependente da fonte de carbono disponível no meio (ARO, PAKULA, PENTTILA, 2005). Por outro lado, o produto final da hidrólise da celulose (glicose), inibe a transcrição das celulases e xilanases de fungos e bactérias via repressão catabólica (NAKARI-SETALA et al., 2009; KUBICEK et al., 1993).

Em *Trichoderma reesei*, os transcritos para celulases e hemicelulases são induzidos quando a celulose e outros polímeros são usados como fonte de carbono (ARO, PAKULA, PENTTILA, 2005). Porém, a celulose é insolúvel e não pode atravessar a membrana celular e entrar na célula. Hipóteses propostas sustentaram que a indução ocorre então, devido à expressão de níveis basais de enzimas celulolíticas que produzem, a partir da celulose presente no meio, oligossacarídeos solúveis que entram na célula e agem como indutores de síntese de celulases (MANDELS & REESE, 1960). Evidências que sustentam estas hipóteses foram reportadas, mostrando que a indução do

transcrito *cbhl* pela celulose foi completamente reprimida pela adição prévia de anticorpos para as principais celulases CBHI, CBHII, EGI e EGII bloqueando a ação das mesmas na celulose. Também, a adição prévia dos mesmos anticorpos não impediu a transcrição do *cbhl* quando a sofrose, que é solúvel e entra na célula, foi usada como indutor (EL-GOGARY et al., 1989). Outro trabalho mostrou, entre outras evidências, a presença de transcritos de *cbhl* e *egl1* em células de *T. reesei* não induzidas e a atividade do promotor de *cbhl* na ausência de celulose (CARLE-URIOSTE et al., 1997).

Existem inúmeras aplicações industriais para as celulases. Estas enzimas são utilizadas na indústria têxtil no processo de “stonewashing” de tecidos “jeans” para o amolecimento da fibra e obtenção de aparência desgastada e desbotada. As celulases também são úteis na indústria de bebidas melhorando a qualidade e a produtividade em processos de fermentação para obtenção de vinho e cerveja. Na indústria de alimentos as celulases são usadas para a diminuição da viscosidade dos sucos de frutas. Na indústria de ração animal estas enzimas são utilizadas para o pré-tratamento da silagem aumentando seu valor nutricional. As celulases são utilizadas até mesmo na indústria de papel podendo auxiliar no braqueamento da polpa de celulose e na remoção de tinta em processos de reciclagem de papel (KUHAD, GUPTA, SINGH, 2011). Entretanto, (SUKUMARAN, SINGHANIA, PANDEY, 2005) ressaltou que a mais importante aplicação das celulases é para a produção de etanol de segunda geração a partir da biomassa vegetal

## 1.5 Hemicelulases

Assim como para as celulases, as hemicelulases também atuam de modo sinérgico para a degradação do substrato. São necessárias várias enzimas diferentes para hidrolisar hemicelulose devido a sua heterogeneidade (SAHA, 2003). Como o xilano é muito abundante na hemicelulose as xilanases (ou endoxilanases) apresentam grande importância na degradação deste polímero, hidrolizando ligações  $\beta$ -1,4 e produzindo oligômeros que podem ser hidrolisados em xilose pela  $\beta$ -xylosidase (não menos importante). Entretanto, enzimas adicionais são requeridas para a completa degradação da hemicelulose (SHALLOM & SHOHAM, 2003).

O fungo *Trichoderma reesei* é um dos melhores secretores de hemicelulases, produzindo em altas concentrações a maior parte das hemicelulases já descritas para

microorganismos (SHALLON & SHOHAM, 2003; MARTINEZ et al., 2008). As hemicelulases descritas para *T. reesei* (FOREMAN et al., 2003) são as endoxilanasas ou xilanasas (Xyn): Xyn1, Xyn2, Xyn3 e Xyn4, que clivam ligações  $\beta$ -1,4-xilana; a  $\beta$ -xilosidase (BXL): Bxl1, que hidrolisa xilano oligossacarídeos e xilobiose (POLIZELI et al., 2005); as acetilxilano esterases (Axe): Axe1 e Axe2, que removem grupos laterais de acetil das fibras de hetero-xilano ou acetil xilano (MARGOLLES-CLARK et al., 1996a); as arabinofuranosidases (Abf): Abf1 e Abf2, que removem resíduos L-arabinose de heteropolissacarídeos (MARGOLLES-CLARK et al., 1996b); a mananase (Man):  $\beta$ -mananase Man1, que cliva a ligação glicosídica tipo  $\beta$ -1,4 de resíduos D-manosil ligados à D-manose ou à D-glicose (TENKANEN et al., 1997); as  $\alpha$ -galactosidases (Agl): Agl1, Agl2, e Agl3, que hidrolisam resíduos  $\alpha$ -galactosídeo de galactomamans poliméricos (MARGOLLES-CLARK et al., 1996c); e a  $\alpha$ -glucuronidase: Glr1, que remove ácidos glucurônicos ligados às unidades de xilose do xilano (MARGOLLES-CLARK et al., 1996d).

Considerando a possibilidade da despolimerização da hemicelulose e de sua disponibilização direta para fermentação com o pré-tratamento, a produção de hemicelulases pode apresentar menor importância em relação à produção de celulases no processamento da biomassa vegetal para a produção de etanol de segunda geração (OLOFSSON, BERTILSSON, LIDÉN, 2008; VAN VLEET & JEFFRIES, 2009; PEREIRA, COUTO, ANNA, 2008). No entanto, ao utilizar pré-tratamentos que removem parcialmente a hemicelulose, existe a demanda de atividade xilanásica para a hidrólise da biomassa. Isto é particularmente importante tendo em vista que os xilooligômeros inibem fortemente a hidrólise da celulose (QING, YANG, WYMAN, 2010; MAEDA et al., 2011).

Por outro lado, xilanasas e mananasas apresentam grande importância na indústria de celulose e papel: auxiliando na remoção da lignina residual, que é responsável pela coloração escura da polpa de celulose, e diminuindo o uso de agentes químicos (compostos clorados) durante o processo de branqueamento da polpa de celulose (VIIKARI et al., 1994; POLIZELI et al., 2005; SAMIRIT & JAGDEEP, 2008).

## **1.6 Ligninases**

As enzimas ligninases compreendem duas famílias lignolíticas: a da fenol oxidase (laccase) e a das peroxidases. Esta última inclui a lignina peroxidase (LiP) e a



manganês peroxidase (MnP) (DASHTBAN, SCHRAFT, QIN, 2009). Alguns fungos têm sido descritos como bons produtores de ligninases (SALVACHÚA et al., 2011; DASHTBAN, SCHRAFT, QIN, 2009), entre eles o mais estudado é o *Phanerochaete chrysosporium* (MARTINEZ et al., 2004). Por outro lado, representantes do gênero *Trichoderma*, como *T. reesei* e *T. harzianum* não são produtores de ligninases (DASHTBAN, SCHRAFT, QIN, 2009). Entretanto, assim como no caso das hemicelulases, as ligninases perdem importância no processamento da biomassa vegetal em relação as celulases. Isso ocorre em parte por que a separação da lignina (deslignificação) é comumente efetuada com os pré-tratamentos normalmente utilizados, entre eles o “organosolv”, que usa solventes orgânicos (ZHAO, CHENG, LIU, 2009; PAN et al., 2005); em parte pela complexidade da degradação biológica de lignina que se caracteriza pela lentidão e baixa taxa na hidrólise (SALVACHÚA et al., 2011; BRODEUR et al., 2011); e em parte pela agregação de valor e aplicabilidade das ligninases na indústria de papel, onde estas enzimas são úteis no branqueamento da polpa (biopolpação), e no tratamento dos resíduos poluentes oriundos do branqueamento da polpa (TIEN, 1987; BAJPAI, ANAND, BAJPAI, 2006).

### **1.7 Enzimas acessórias**

Algumas enzimas denominadas acessórias têm sido descritas em microorganismos que degradam biomassa vegetal por serem capazes de diminuir a cristalinidade da celulose e dispersar suas microfibrilas (amorfogênese). Estas enzimas normalmente não apresentam atividade hidrolítica, mas são úteis para o aumento da taxa de hidrólise, uma vez que promovem o aumento da área de superfície da celulose tornando-a mais acessível para as celulases (ARANTES & SADDLER, 2010).

A Swolenina (SWOI) é uma enzima acessória descrita em *T. reesei* que apresenta similaridade com expansinas de plantas. Estas por sua vez são associadas com o relaxamento da estrutura da parede da célula vegetal durante o crescimento celular (SALOHEIMO et al., 2002). De modo análogo, a SWOI de *T. reesei*, quando usada em pré-tratamento de substratos celulósicos insolúveis (celulose cristalina) aumenta a acessibilidade das celulases e com isso acelera a hidrólise (JÄGER et al., 2011).

Outra enzima acessória, a loosenina, também relacionada com expansina de planta, foi descrita para o fungo *Bjerkandera adusta* (QUIROZ-CASTAÑEDA et al., 2011). Após realizar o pré-tratamento de substrato celulósico nativo com a loosenina,

estes autores demonstraram a capacidade desta proteína em desorganizar as fibras celulósicas e com isso promover a conseqüente aceleração das taxas de hidrólise.

Vale ainda destacar a endoglucanase EGVII de *T. reesei* que, apesar de ser classificada como celulase, exibe baixos níveis de atividade endoglucanásica e não apresenta evidências de possuir um domínio catalítico (KARKEHABADI et al., 2008). Por outro lado, apesar de não apresentar similaridade com as expansinas, a EGVII pode ser considerada como uma enzima acessória candidata, pois há indicações de que ela pode contribuir de modo sinérgico com as demais celulases, aumentando a atividade celulolítica em substratos nativos (MERINO & CHERRY, 2007).

Por último podemos citar a EGIII que apesar de possuir atividade endoglucanásica apresentou em *T. reesei* e *Phanerochaete chrysosporium*, a capacidade de relaxar a estrutura da celulose, assumindo papel de enzima acessória (HENRIKSSON et al., 1999; YUAN, WU, COSGROVE, 2001).

## **1.8 Principais desafios para desconstrução de biomassa vegetal e produção de etanol**

Apesar de ter havido significativo progresso, a viabilidade da comercialização do etanol de segunda geração tem sido comprometida não só pela heterogeneidade e recalcitrância da biomassa vegetal em si, mas também pela necessidade de múltiplos tratamentos para um processamento efetivo (US-DOE, 2006). Várias tecnologias têm sido desenvolvidas durante as últimas décadas permitindo a viabilidade técnica da conversão dos açúcares da biomassa em álcool, no entanto, o objetivo claro atualmente é fazer com que esse processo tenha um custo competitivo no mercado atual (HIMMEL et al., 2007; DASHTBAN, SCHRAFT, QIN, 2009; SERVICE, 2010).

Avanços na tecnologia de pré-tratamento para aumentar a acessibilidade do material celulósico à hidrólise enzimática, com menor uso de produtos químicos e energia, são vitais. As pesquisas que envolvem pré-tratamento podem resultar na redução do tempo de bioconversão, na diminuição da quantidade de celulase utilizada e no aumento do rendimento de etanol (MOSIER et al., 2005; KNAUF & MONIRUZZAMAN, 2004; BRODEUR et al., 2011). Além disso, (WYMAN, 2007) aponta como pontos principais para vencer a resistência natural da parede celular das plantas: o pré-tratamento, como o fator que representa o maior custo na produção de bioetanol, e a produção de enzimas para hidrólise, como o segundo fator mais custoso.

O autor ainda destaca que, a despeito de investimentos substanciais já realizados, o processo de produção das celulasas é caro e sua ação é lenta, portanto, além da demanda de preparar sólidos mais reativos com melhorias no pré-tratamento, as enzimas também precisam ser melhoradas (WYMAN, 2007).

Por outro lado, (BANERJEE et al., 2010a) apontam o custo das enzimas que degradam a biomassa vegetal como o principal gargalo para a produção industrial e economicamente viável de etanol de segunda geração. Em concordância com estas observações, (SUKUMARAN et al., 2009) apontam que o principal impedimento para a sacarificação enzimática da biomassa, para a produção de etanol, é o custo das celulasas. As abordagens para reduzir o custo das enzimas incluem: a busca por novos organismos, principalmente fungos, para obtenção de versões superiores das enzimas e para obtenção de enzimas que atuem de modo sinérgico com as preparações enzimáticas comerciais; o contínuo aperfeiçoamento das linhagens disponíveis por mutagênese convencional e molecular; a melhoria das enzimas, para as características desejáveis, por engenharia de proteína e evolução dirigida e a melhora da eficiência na produção de enzima em escala industrial (BANERJEE et al., 2010a). Neste contexto, (KNAUF & MONIRUZZAMAN, 2004) apontam as seguintes abordagens para melhorar a hidrólise enzimática: estudos através de expressão e purificação de celulasas de espécies variadas; identificação de novas proteínas por estudos de genômica, transcriptômica e proteômica, e otimização da mistura de enzimas para uma máxima degradação de biomassa vegetal. Além da mutagênese, (DASHTBAN, SCHRAFT, QIN, 2009) apontam também a importância da expressão heteróloga, que abre caminho para as técnicas de obtenção de enzimas melhoradas, e a importância dos co-cultivos de microorganismos para a obtenção de preparados enzimáticos.

Sobre a busca por novos organismos e enzimas no Brasil, (SOCCOL et al., 2010) destacaram a promissora fonte de novas enzimas lignocelulolíticas representada pelo uso da metagenômica de microorganismos não cultiváveis, a partir de diferentes biomas brasileiros, viabilizada por projetos que vem sendo conduzidos na amazônia, mata atlântica e cerrado.

Entre as abordagens possíveis para melhora das enzimas, podemos citar como exemplo o trabalho de (NAKAZAWA et al., 2009) que dentro do conceito de evolução dirigida (LIN et al., 2011), utilizaram a técnica de “error-prone PCR” para obter melhora de estabilidade e atividade da endoglucanase III de *Trichoderma reesei*. Vale citar também o uso de dados estruturais para auxiliar na engenharia de proteínas, como

no trabalho de (VOUTILAINEN et al., 2010) no qual dados de estrutura foram utilizados para orientar mutações dirigidas que culminaram na obtenção de mutantes da CBHI de *Talaromyces emersonii* com atividade e termoestabilidade melhoradas.

Outra técnica utilizada consiste na fusão de genes e, como exemplo, podemos citar (LEE et al., 2011) que fusionaram a  $\beta$ -glucosidase de *Clostridium cellulovorans* com endoglucanase e exoglucanase de *Clostridium thermocellum*, obtendo assim maior rendimento na produção de glicose em relação as enzimas não fusionadas.

Um exemplo de aperfeiçoamento de cepas é o trabalho de (DILLON et al., 2006) que fizeram mutagênese do *Penicillium echinulatum* pelo tratamento dos conídios com peróxido de hidrogênio ou 1,2,7,8-diepoxioctano, resultando na obtenção de duas cepas com atividade de hidrolítica aumentada.

Combinações de enzimas de fontes diversas podem ser utilizadas nos preparados enzimáticos para aumento de eficiência de hidrólise. Isto ocorre, por exemplo, no caso das preparações de celulases de *T. reesei*, que frequentemente apresentam baixa atividade de  $\beta$ -glucosidase, sendo necessária a suplementação do sobrenadante de cultura com adição desta enzima (LYND et al., 2002; PALLAPOLU et al., 2011). As combinações de enzima podem ser obtidas também com a co-cultura de microorganismos, como realizado por (AHAMED & VERMETTE, 2008), que reportaram melhores atividades celulolíticas de cultura mista de *Trichoderma reesei* Rut-C30 e *Aspergillus niger* LMA, quando comparadas a mono-cultura de *Aspergillus niger* LMA. Outro exemplo de obtenção de preparado enzimático para melhor desempenho de hidrólise é o trabalho de (MAEDA et al., 2011) que otimizaram uma mistura enzimática para a degradação de bagaço de cana-de-açúcar pré-tratado a partir de extratos enzimáticos de *Penicillium funiculosum* ATCC11797, *Trichoderma harzianum* IOC 3844 e de um coquetel enzimático comercial.

As características desejadas para as combinações e preparados enzimáticos variam dependendo do tipo de substrato em questão. Para adequar as combinações dos preparados enzimáticos aos diferentes tipos de biomassa e pré-tratamentos, (BANERJEE et al., 2010b) reportaram a plataforma de alto rendimento “GENPLAT” que possui manuseio automatizado de líquidos, ensaios automatizados de detecção de glicose e xilose e tratamento estatístico dos dados. A plataforma é adaptável para misturas de enzimas puras, preparados comerciais e extratos de microorganismos, permitindo a descoberta de novas enzimas além da otimização do coquetel enzimático (WALTON, BANERJEE, CAR, 2011).

As características desejadas para as celulases podem variar também dependendo do tipo de tecnologia adotada. Entre as demandas existe a busca por enzimas que apresentem melhores atividades em temperaturas mais altas do que aquelas por volta dos 45 °C tipicamente características dos preparados enzimáticos comerciais. Isto é particularmente vantajoso no caso de uso do processo de hidrólise e fermentação separadas (SHF), no qual, por conta da maior temperatura na hidrólise, ocorre maior desorganização da parede celular e conseqüente melhor penetrabilidade das enzimas (TURNER, MAMO, KARLSSON, 2007). Com isso, o tempo de reação e a quantidade de enzima necessária reduzem, possibilitando potencial redução de custo (VIAKARI et al., 2007). Por outro lado, enzimas com atividade ótima em altas temperaturas (mais do que 45 °C ou até mesmo próximo a isso), podem freqüentemente não se adequar ao processo SSF, no qual a hidrólise enzimática da celulose ocorre simultaneamente à fermentação da glicose. Isto acontece por que a temperatura do processo SSF fica em torno de 37 °C, que é a melhor temperatura possível para a atividade das celulases e a concomitante fermentação com a levedura *Saccharomyces cerevisiae* (temperatura ótima normalmente em torno de 30 °C). Nesta situação, as taxas de hidrólise se tornam substancialmente baixas e podem ser desfavoráveis por aumentar o tempo do processo. Uma possível solução para o problema seria o uso de cepas de levedura termotolerantes que fermentariam bem em temperaturas mais altas aumentando assim as taxas de hidrólise (OLOFSSON, BERTILSSON, LIDÉN, 2008; KÁDÁR, SZENGYEL, RÉCZE, 2004).

No caso das hemicelulases, principalmente as xilanases e mananases, pela sua importância e aplicabilidade na indústria de papel, busca-se incrementar a termoestabilidade. Isto por que estas enzimas auxiliam no branqueamento da polpa de celulose, que por sua vez é realizado em altas temperaturas (POLIZELI et al., 2005). Algumas técnicas já citadas acima para melhorar as celulases, também se aplicam no caso da melhoria das hemicelulases visando atender a demanda por hemicelulases termoestáveis. Estas técnicas incluem: a busca por novos organismos para identificação de novas enzimas e para obtenção de versões superiores das enzimas; a melhoria das enzimas por engenharia de proteína e evolução dirigida e a melhora da eficiência na produção de enzima em escala industrial. Como exemplo de melhoria de enzima podemos citar (RULLER et al., 2008), que obtiveram variantes termoestáveis melhoradas de xilanase de *Bacillus subtilis* por evolução dirigida combinando estudos estruturais com as técnicas de “error-prone PCR” e “DNA shuffling”.

## 1.9 O fungo *Trichoderma harzianum*

O *Trichoderma harzianum* (estágio assexuado anamorfo), classificado também como *Hypocrea lixii* em seu estágio sexuado ou teleomorfo (CHAVERRI et al., 2003), tem sido tradicionalmente usado em controle biológico, agindo como antagonista contra patógenos de plantas (LÓPEZ-MONDÉJAR, ROS, PASCUAL, 2011; SHARMA et al., 2011; BENÍTEZ et al., 2004). Entre os inúmeros estudos envolvendo este tema, alguns autores reportaram trabalhos baseados em bibliotecas de cDNA de *T. harzianum*, com objetivo de elucidar os mecanismos de biocontrole e identificar os genes envolvidos (STEINDORFF et al., 2012; LIU & YANG, 2005; VIZCAÍNO et al., 2006; VIZCAÍNO et al., 2007). No contexto do presente trabalho é importante citar que estas bibliotecas de cDNA foram construídas a partir de micélios de *T. harzianum*, cultivados sob condições representativas, ou simuladoras de situações de biocontrole, e portanto não indutoras de celulases e hemicelulases.

Neste contexto o *Trichoderma harzianum* não tem sido tradicionalmente usado para a produção de celulases, e de fato, são raros na literatura os trabalhos que abordam este tema para o fungo. No entanto, interessantemente (SADDLER, HOGAN, LOUIS-SEIZE, 1985) já haviam comparado em 1985, os sobrenadantes de culturas induzidas de *T. harzianum* E58 e de *T. reesei* RutC30, que é grande produtor de celulases (PETERSON & NEVALAINEN, 2012), encontrando atividade celulolítica ligeiramente mais alta e produção de glicose mais elevada no *T. harzianum* E58. Contrariamente, (GUILFOILE et al., 1999) relataram baixa atividade celulolítica do *T. harzianum* FP108 em relação ao *T. reesei* QM9414, e sugeriram que isto poderia ser explicado por diferenças na regulação da CBHI e da CBHII.

Mais recentemente, estudos de (CASTRO et al., 2010a) demonstraram que a linhagem do fungo filamentoso *Trichoderma harzianum* IOC-3844, é boa produtora de celulases e, por isso, apresenta potencial aplicação como fonte de enzimas para a sacarificação da celulose. Nestes estudos o extrato enzimático do fungo, ou seja, o sobrenadante da cultura obtido a partir do cultivo do micélio induzido para a produção de celulases, apresentou atividades celulolíticas bem balanceadas, altos conteúdos de atividade de endoglucanases, e níveis significativos de atividade de papel de filtro (FPase) e de  $\beta$ -glucosidase. Estes autores também demonstraram que, comparativamente, o *Trichoderma harzianum* IOC-3844 mostrou maiores atividades

celulolíticas (FPase, endoglucanase e  $\beta$ -glucosidase) em relação ao *Trichoderma reesei* RutC30 (CASTRO et al., 2010a). Além disso, os estudos de (MAEDA et al., 2011) demonstraram a aplicabilidade do extrato enzimático do *T. harzianum* IOC-3844, na hidrólise do bagaço de cana-de-açúcar.

Vale citar ainda o trabalho de (DELABONA et al., 2012) onde a capacidade de hidrólise dos fungos *T. harzianum* P49P11 e *T. reesei* Rut-C30 é comparada. Os autores reportaram capacidade de produção de glicose similar entre os fungos, na sacarificação de bagaço de cana pré-tratado e atividades de xilanase equivalentes.

Por outro lado, o *T. harzianum* tem sido estudado há tempos como fonte produtora de xilanases (WONG, TAN, SADDLER, 1988; POLIZELI et al., 2005; MIN et al., 2009) e até mesmo mananases (TORRIE, SENIOR, SADDLER, 1990). Interessantemente, os estudos de (SADDLER, HOGAN, LOUIS-SEIZE, 1985) que indicam diferenças na regulação no sistema de xilanases entre *T. reesei* e *T. harzianum*, reportando: o *T. reesei* Rut-C30 com suas atividades xilanásicas mais altas quando induzido por xilano; e o *T. harzianum* E58 produzindo suas atividades xilanásicas e celulolíticas mais altas quando induzido apenas por celulose. Neste contexto, (SADDLER, HOGAN, LOUIS-SEIZE., 1985) reportaram atividade xilanásica bastante aumentada de *T. harzianum* E-58 em relação ao *T. reesei* Rut-C30, quando ambos foram induzidos por celulose. Também, mais recentemente, (MAEDA et al., 2011) reportaram importante atividade xilanásica do *T. harzianum* IOC-3844 induzido por celulose.

Apesar do potencial como fonte produtora de enzimas hidrolíticas, o conjunto de seqüências relacionadas com a desconstrução de biomassa vegetal ainda não foi reportado para o *Trichoderma harzianum*. De fato, o transcriptoma envolvido com a degradação de biomassa deste fungo é pouco conhecido, visto que, após buscas detalhadas (ver mais adiante em Resultados e em Discussão), além do depósito da seqüência da EGIII clonada e seqüenciada no presente trabalho [GenBank:AFK32784], constatamos que existem publicamente disponíveis para *Trichoderma harzianum* apenas as seqüências das celulasas CBHI: [GenBank:Q9P8P3] e [GenBank:ADH04808] e BGL2: [GenBank:ABN50090] e [GenBank:ABN50091]; da xilanase Xyn2: [GenBank:ACF40831] e [GenBank:P48793]; e do fator de transcrição CRE1 [GenBank:CAA64656] (BENSON et al., 2011).

Estes fatos nos encorajaram a construir uma biblioteca de cDNA para estudar o transcriptoma do *Trichoderma harzianum* IOC-3844, induzido para a produção de celulasas e hemicelulasas, com foco nas enzimas envolvidas na degradação da biomassa

vegetal e também, para clonar os genes que codificam para estas enzimas possibilitando ensaios de expressão heteróloga. Para isso, este fungo foi induzido por celulose, uma vez que ele produz boa atividade celulolítica e xilanásica na presença deste indutor (SADDLER, HOGAN, LOUIS-SEIZE, 1985). Vale destacar que o *Trichoderma harzianum* IOC-3844 não sofreu qualquer modificação genética e que este é o primeiro estudo que descreve o conjunto de transcritos de *Trichoderma harzianum* relacionados à degradação de biomassa.

Para estudar a expressão dos transcritos encontrados, a contagem de ESTs (Etiquetas de sequencias expressas ou “Expression sequence tags”) da biblioteca induzida de *T. harzianum* IOC-3844 foi comparada com a contagem de ESTs de uma biblioteca induzida de *T. reesei* QM6a, disponível publicamente (DIENER et al., 2004). Além disso, foram realizadas análises de PCR em tempo real em transcritos representativos do conjunto de enzimas relacionadas à desconstrução de biomassa vegetal do *T. harzianum* IOC-3844, confirmando o potencial deste fungo para produzir estas enzimas e mostrando sua possível aplicabilidade na produção de etanol de segunda geração e na indústria de papel.



## 1.10 Produção científica

Os resultados deste trabalho permitiram a redação de três artigos para publicação. Com os resultados das análises de expressão reportadas aqui um manuscrito foi redigido e submetido para publicação (ver apêndice A). Outro manuscrito foi publicado no periódico “Genetics and Molecular Research” (GENEROSO et al., 2012), reportando a expressão heteróloga (descrita aqui) e a caracterização enzimática (realizada pela equipe do Laboratório de Biologia Molecular DGE-UFSCar) da endoglucanase III do *T. harzianum* IOC-3844 (ver apêndice B). O terceiro artigo foi publicado no periódico “Acta Crystallographica” (LIBERATO et al., 2012) e reporta a cristalização e análise preliminar de difração de raios-X da endoglucanase III de *T. harzianum* IOC-3844 (ver apêndice C), a qual foi identificada no presente trabalho. Ainda, um quarto artigo, não relacionado com o assunto desta tese (MALAGÓ-JR et al., 2009), foi publicado como resultado das atividades desenvolvidas na disciplina cursada durante o doutoramento, para atender o “Programa de Estágio Supervisionado de Capacitação de Docente” (PESCD), originado por iniciativa da “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior” (CAPES). Este artigo foi publicado no periódico “Biochemistry and Molecular Biology Education” (BAMBED) e trata de uma proposta inovadora de abordagem para aula prática com objetivo de demonstrar o DNA como material genético em experimentos inéditos (ver apêndice D).

## 2 OBJETIVOS

Os objetivos originais do projeto foram:

- 1- Construir uma biblioteca de cDNA de *Trichoderma harzianum* induzido para a produção de celulases e seqüenciar os clones.
- 2- Identificar e caracterizar os genes que codificam para celulases por meio de comparação com celulases conhecidas de outros organismos.
- 3- Isolar e clonar os genes que codificam para as celulases de *Trichoderma harzianum*.
- 4- Desenvolver sistemas de expressão e purificação das celulases do *Trichoderma harzianum* em *Pichia pastoris*.
- 5- Estudar o transcriptoma de *Trichoderma harzianum* induzido para a produção de celulases comparando com o transcriptoma do fungo não induzido.

Outras informações sobre os objetivos propostos e sobre o cumprimento do cronograma podem ser vistas no apêndice E.

### 3 MÉTODOS

#### 3.1 Inóculo e cultura de *Trichoderma harzianum* IOC-3844

A linhagem de *Trichoderma harzianum* IOC-3844 foi obtida da coleção de culturas da Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro. A linhagem foi mantida a 28 °C em placas contendo PDA - “Potato Dextrose Agar” (Difco - 213400). O meio utilizado para cultura líquida foi descrito por (MANDELS, PARRISH, REESE, 1962) e complementado com celulose microcristalina (Avicel), em caso de indução de celulases e hemicelulases, ou glicerol, em caso de não indução. O meio é constituído por: Uréia 0,03 %; (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 0,14 %; MgSO<sub>4</sub> 0,03 %; Peptona 0,2 %; Tween-80 0,2 %; CaCl<sub>2</sub> 0,03 % e solução de metais (FeSO<sub>4</sub> 0,5 %, MnSO<sub>4</sub> 0,16 %, ZnSO<sub>4</sub> 0,16 % CoCl<sub>2</sub> 0,2 %) a 0,1 % em tampão fosfato de potássio 0,1 M pH 5,6. Para obtenção do micélio, 10<sup>7</sup> esporos foram coletados do cultivo em placas contendo meio PDA sólido e foram sincronizados em 10 mL de meio contendo glicerol 0,08 % por 16 horas a 200 RPM e 28 °C. Foram adicionados 10 mL da suspensão de esporos em 500 mL de meio contendo 0,8 % de glicerol. A amostra foi então dividida em duas alíquotas de 250 mL para as culturas de indução e de não indução. O fungo cresceu por 17 h a 200 RPM e 28 °C, depois disso, cada amostra foi centrifugada a 12000 RPM (21850 g) por 18 minutos e 4 °C. O sobrenadante foi descartado e os micélios foram lavados duas vezes em 100 mL de tampão fosfato pH 6,0, sob as mesmas condições de centrifugação. Após as lavagens, os micélios foram incubados por 1 h em 100 mL de tampão fosfato pH 6,0 a 200 RPM e 28 °C. O material foi então centrifugado novamente e ressuspensionado em 250 mL de meio contendo 0,8 % de glicerol para o caso de não indução e em 250 mL de meio contendo 1 % de celulose microcristalina (Avicel) no caso de indução de celulases e hemicelulases. Antes do uso, a celulose microcristalina foi lavada uma vez com HCl 1 mM, lavada uma vez com água deionizada, autoclavada por 15 minutos, lavada uma vez com água deionizada e lavada duas vezes com tampão fosfato. As culturas permaneceram a 200 RPM e 28 °C até a coleta.

#### 3.2 Coletas de micélio

Os cultivos de 250 mL, induzido e não induzido permaneceram a 200 RPM e 28 °C e tiveram 10 pontos coletados: 8, 10, 12, 14, 16, 20, 24, 36, 48 e 60 horas. Em cada

coleta, 10 mL do micélio foram colocados em aparato de filtração esterilizado (“Stericup Filter Units” - Millipore) contendo papel de filtro comum. As amostras foram filtradas para a remoção do sobrenadante com auxílio de bomba de vácuo. Após a filtragem o micélio foi removido rapidamente do papel de filtro com auxílio de pinças, e imediatamente embrulhado em papel alumínio e congelado em nitrogênio líquido. As amostras foram então mantidas em freezer -80 °C.

### **3.3 Extração de RNA total**

A extração de RNA total foi realizada com uso do reagente Trizol (Invitrogen - 15596-026) conforme protocolo proposto pelo fabricante com poucas alterações. Resumidamente, foi adicionado 1 mL de Trizol para cada 120 mg de micélio, após a maceração e pulverização do micélio em nitrogênio líquido com auxílio de almofariz e pistilo. O material foi então centrifugado a 11400 RPM (12100 g) por 10 minutos a 4 °C, e o sobrenadante foi recuperado. Foram adicionados 200 µL de clorofórmio, a amostra foi então homogenizada com agitação moderada por 15 segundos e centrifugada a 11400 RPM (12100 g) por 15 minutos a 4 °C. A fase aquosa foi recuperada e submetida a duas lavagens com adição de 1 volume de fenol:clorofórmio:álcool isoamílico 25:24:1 (Invitrogen - 15593-031) e centrifugação de 11400 RPM (12100 g) por 15 minutos a 4 °C. O sobrenadante foi recuperado e o RNA foi precipitado pela adição de 500 µL isopropanol, homogenização por agitação moderada por 10 segundos, incubação por 10 minutos a temperatura ambiente e centrifugação a 11400 RPM (12100 g) por 10 minutos a 4 °C. O sobrenadante foi descartado e o precipitado foi lavado com adição de 1 mL de etanol 70 %, agitação moderada por 10 segundos e centrifugação a 7000 RPM (4500 g) por 5 minutos a 4 °C. O sobrenadante foi descartado e o precipitado foi ressuspenso em 30 µL de água deionizada livre de RNase. As amostras permaneceram por 10 minutos a 60 °C antes de serem armazenadas a -80 °C. A quantificação, pureza e integridade das amostras de RNA total foram checadas através eletroforese em gel de agarose 1 % corado com brometo de etídeo, e com medidas de densidade ótica no espectrofotômetro “Nanodrop 1000”.

### 3.4 Extração de RNA mensageiro

O RNA mensageiro foi extraído a partir de 230 µg de RNA total com uso do “kit” “PolyATtract mRNA Isolation System III” (Promega - Z5300), segundo recomendações do fabricante. Para capturar os RNAs mensageiros foram utilizados 150 pMol de sonda de oligo-dT biotilado. Depois foram adicionados 600µL da solução contendo partículas paramagnéticas de estreptavidina, que se ligam à biotina. O material foi então colocado em estante magnética para atração das partículas de estreptavidina. As moléculas de RNA mensageiro ficaram então aderidas à parede do tubo e o sobrenadante contendo RNA total foi desprezado. O RNA mensageiro foi lavado conforme recomendações do fabricante do “kit”, e depois eluído em 250 µL de água deionizada livre de RNase. O eluato contendo o RNA mensageiro foi recuperado, permanecendo o oligo-dT biotilado preso à parede do tubo para ser desprezado. O material foi então precipitado com adição de 1/10 do volume de acetato de sódio 3M pH 5,2, adição de 1 volume de isopropanol 100 %, incubação a -20 °C durante a noite e centrifugação a 11400 RPM por 40 minutos a 4 °C. O sobrenadante foi desprezado e a amostra foi lavada com adição de 1 mL de etanol 75 % e centrifugação a 11400 RPM por 10 minutos a 4 °C. O sobrenadante foi desprezado e o precipitado foi ressuspenso em 13 µL de água deionizada livre de RNase. A quantificação, pureza e integridade do RNA mensageiro foram checadas através de eletroforese em gel de agarose 1 % corado com brometo de etídeo, e com medidas de densidade ótica no espectrofotômetro “Nanodrop 1000”.

### 3.5 Análises de RT-PCR

A partir das seqüências das ORFs (Fases abetras de leitura ou “Open reading frames”) dos genes *cbh1* (GenBank: AF223252) e *bgl2* (GenBank: EF426298) do *Trichoderma harzianum*, disponíveis publicamente (BENSON et al., 2011), foram desenhados oligonucleotídeos para a detecção dos RNAs mensageiros por PCR com uso de transcriptase reversa “Reverse transcriptase PCR” (RT-PCR). Para o *cbh1* os oligonucleotídeos usados foram: Fw-RT-PCR-*cbh1*-harz 5'-aatggtcttggtcatgagcttg-3' e Rv-RT-PCR-*cbh1*-harz 5'-gttgagtagataacctggagttgg-3'. Para o *bgl2* os oligonucleotídeos usados foram: Fw-RT-PCR-*bgl2*-harz 5'-ttcagtggggtttgacgac-3' e Rv-RT-PCR-*bgl2*-harz 5'-accaggagatggagaagcggtac-3'. Para o controle do RT-PCR foram

desenhados oligonucleotídeos para a ORF da actina de *Trichoderma harzianum* com base no fragmento de EST deste gene (GenBank: AJ901127), disponível publicamente (BENSON et al., 2011). Os oligonucleotídeos para o controle foram: Fw-RT-PCR-actina-harz 5'-tctacaatgagctgcgtgtgcc-3' e Rv-RT-PCR-actina-harz 5'-tcgtaggacttccaagctggag-3'. As reações de transcriptase reversa (RT) foram realizadas com uso do “kit” “ImProm-II Reverse Transcription System” - (Promega - A3800). Para a síntese da primeira fita de cDNA foram adicionados 100 ng de RNA total e 10 pMol de cada oligonucleotídeo reverso específico em uma reação de 5 µL. As amostras foram incubadas a 70 °C por 5 minutos e depois incubadas imediatamente no gelo por mais 5 minutos. Os seguintes reagentes foram adicionados ao RNA total: 6,5 µL de água deionizada livre de nuclease, 4 µL de tampão da transcriptase reversa (5x Impron II), MgCl<sub>2</sub> 2,5 mM final, dNTPs 0,5 mM final, 0,5 µL de inibidor de RNase “RNAsin” (20u) e 1 µL de transcriptase reversa (Promega). As amostras foram incubadas a 25 °C por 5 minutos para hibridização, 42 °C por 1 hora para síntese do cDNA e 70 °C por 15 minutos para a inativação da enzima transcriptase reversa. Além disso, foram feitos controles sem cDNA molde e sem transcriptase reversa para monitorar possíveis contaminações. As PCRs foram feitas com reagentes da Fermentas - EP0402, utilizando 1 µL da reação de RT como molde. As reações de PCR foram realizadas em volume final de 25 µL com adição de 16,3 µL de água deionizada, 2,5 µL de tampão 10X, MgCl<sub>2</sub> 2 mM final, dNTPs 0,4 mM final, 10 pMol de cada oligonucleotídeo e 0,2 µL (1 unidade) da enzima *Taq*. As condições de ciclagem foram: 94 °C por 1 minuto; 35 ciclos de 94 °C por 20 segundos, 56 °C por 20 segundos e 72 °C por 45 segundos; extensão final de 72 °C por 5 minutos.

### **3.6 Construção e seqüenciamento da biblioteca de cDNA de *Trichoderma harzianum* IOC-3844**

O micélio do *Trichoderma harzianum* IOC-38844 induzido por 60 horas foi utilizado para a construção da biblioteca de cDNA, que foi realizada com a utilização do “kit” “CloneMiner cDNA Library Construction Kit” (Invitrogen - 18249-029), conforme recomendações do fabricante. Resumidamente, 1,3 µg de RNA mensageiro serviu de molde para a síntese da primeira fita de cDNA, com uso da enzima transcriptase reversa “Superscript II Reverse Transcriptase” (Invitrogen - 18064-022). Para isto foi utilizado um iniciador oligo-dT contendo o sítio de recombinação *attB2*

biotinilado. A síntese da segunda fita de cDNA e a eliminação do RNA mensageiro foram realizadas simultaneamente com a utilização das enzimas, *E. coli* DNA polimerase I (Invitrogen - 18010-017), *E. coli* RNase H (Invitrogen) e *E. coli* DNA ligase (Invitrogen - 18052-019). O material foi então extraído com igual volume de fenol:clorofórmio:álcool isoamílico 25:24:1 (Invitrogen - 15593-031), e o sobrenadante foi precipitado com 0,5 volume de acetato de amônio 7,5 M; 20 µg de glicogênio e 2,5 volumes de etanol 100 %. Depois as amostras foram lavadas com etanol 70 % e ressuspendidas em água deionizada livre de RNase. O adaptador contendo o sítio de recombinação *attB1* foi então ligado às extremidades 5' do cDNA dupla fita por 16 horas a 16 °C, com uso da enzima T4 DNA ligase. A biotina presente no adaptador *attB2* protege e evita a ligação do adaptador *attB1* à extremidade 3' do cDNA. Após inativação da ligase com incubação a 70 °C por 10 minutos o material foi precipitado como descrito acima. A amostra foi então submetida a fracionamento para a seleção de tamanho dos cDNAs dupla fita com uso da coluna “cDNA size fractionation column” (Invitrogen - 18092-15). As primeiras frações contendo os fragmentos maiores de cDNA foram selecionadas seqüencialmente até totalizar 300 ng de cDNA que foram precipitados como descrito acima. Depois da precipitação o material foi quantificado e 100 ng foram recombinados e em plasmídeo pDONR-222 (Invitrogen) por 16 h a 25 °C, com uso da mistura de enzimas “BP Clonase Enzyme Mix” (Invitrogen - 11789-013). A mistura de clonases foi inativada pela adição de proteinase K (Invitrogen) e incubação a 37 °C por 15 minutos. A proteinase K foi inativada a 75 °C por 10 minutos e o material foi então precipitado como descrito acima. Os plasmídeos foram usados para transformar células de *E. coli* eletrocompetentes “ElectroMax DH10B T1 Phage Resistant Cells” (Invitrogen - 12033-015). As células foram submetidas a diluições seriadas e plaqueadas em meio LB-Agar contendo 50 µg/ml de Kanamicina (USB). O título da biblioteca foi determinado pela contagem de colônias. As colônias foram picadas em 1 mL de meio LB-broth contendo 50 µg/ml de Kanamicina (USB) em placas de 96 clones e incubadas por 22 horas a 300 RPM e 37 °C. As culturas foram utilizadas para realizar minipreparações de DNA plasmidial que foram realizadas, conforme protocolo descrito por (VETTORE et al., 2003). O tamanho médio dos clones foi verificado por clivagem com uso da enzima de restrição *BsrGI* (New England Biolabs - R0575S). Os clones foram seqüenciados a partir da extremidade 5' com a utilização do oligonucleotídeo M13-Fw 5'-gtaaaacgacggccag-3' no seqüenciador “MegaBACE 1000

Flex” (GE Healthcare), com utilização do “kit” “DYEnamic ET Terminator Cycle Sequencing Kit” (GE Healthcare).

### 3.7 Análise das bibliotecas de cDNA

As seqüências da biblioteca de *Trichoderma harzianum* IOC-3844 foram analisadas com uso da plataforma de “softwares” dCAS - “Desktop cDNA Annotation System” (GUO et al., 2009). “Um total de 2016 seqüências em formato de arquivos brutos com extensão “.esd” foram analisadas. Foram atribuídos valores de qualidade para cada base das seqüências brutas com uso do “software” “phred” (EWING et al., 1998; EWING & GREEN, 1998). As seqüências de baixa qualidade foram então removidas com utilização dos critérios de seleção de qualidade “default” da plataforma dCAS. Depois, as seqüências correspondentes a oligonucleotídeos e plasmídeos foram removidas através de buscas no banco de dados UniVec (<http://www.ncbi.nlm.nih.gov/VecScreen/UniVec.html>). Restaram 1544 seqüências de boa qualidade que foram agrupadas com uso da ferramenta de busca por alinhamento BLAST - “Basic Local Alignment Search Tool” (ALTSCHUL et al., 1990) e do programa de agrupamento de sequencias Cap3 (HUANG & MADAN, 1999). Com isso as seqüências foram arranjadas em 1003 transcritos, sendo 167 “contigs” ou agrupamentos, formados por mais de uma seqüência, e 836 ESTs que permaneceram isolados em seqüência única ou “singlet”. Os 1003 transcritos, ou seja, as 167 seqüências consenso obtidas dos “contigs” e as 836 seqüências únicas, foram identificados através de buscas via “blastx” (ALTSCHUL et al., 1990) contra o banco de dados de seqüências de proteínas não redundantes (nr). Os dados analisados foram compilados em uma planilha “Excel” contendo “hyperlinks” para os arquivos de dados gerados e para os localizadores de recurso padrão, ou “URLs”, associados a cada transcrito.

Além da biblioteca de cDNA de *T. harzianum* IOC-3844, produzida neste estudo, as seqüências de cDNA de uma biblioteca induzida de *T. reesei* QM6a (DIENER et al., 2004), disponíveis publicamente, também foram analisadas na plataforma dCAS (GUO et al., 2009). No entanto, para estas análises o passo de seleção de qualidade não foi realizado no dCAS, uma vez que as seqüências foram disponibilizadas publicamente em formato “FASTA”, já tendo havido seleção de qualidade com uso do software “phred”, aproveitando seqüências com mais de 100



bases com qualidade  $\geq 20$  ou com média de qualidade  $\geq 12$  (DIENER et al., 2004). Para mapear os transcritos relacionados com degradação de biomassa vegetal, os dados que retornaram da busca via “blastx” das duas bibliotecas de cDNA, foram cuidadosamente examinados considerando todas as nomenclaturas possíveis dos genes descritas em <http://www.ncbi.nlm.nih.gov/gene> (NCBI), e levando em consideração os principais genes já reportados (FOREMAN et al., 2003; SHALLON & SHOHAM, 2003; POLIZELI et al., 2005; DASHTBAN, SCHRAFT, QIN, 2009; ARANTES & SADDLER, 2010). Os fatores de transcrição envolvidos na desconstrução da biomassa também foram mapeados levando em consideração os principais já reportados (ARO, PAKULA, PENTTILA, 2005; KUBICEK et al., 2009). Os transcritos que foram listados para *T. harzianum* IOC-3844 e *T. reesei* QM6a foram obtidos a partir dos alinhamentos entre os consensos de cDNA e os melhores “hits” que retornaram do “blastx” (ALTSCHUL et al., 1990). Estes alinhamentos foram analisados e atingiram um critério mínimo de confiabilidade com “E-values”  $< 1e-10$ , identidade  $> 40\%$  e tamanho de alinhamento  $> 100$  aminoácidos. A maior parte das ORFs de *T. harzianum* IOC-3844, envolvidas com degradação de biomassa vegetal (Tabela 2) foi seqüenciada completamente e alinhada com as proteínas correspondentes de *T. reesei*. Os alinhamentos foram feitos com uso da ferramenta “blastp” (ALTSCHUL et al., 1997) e podem ser vistos no apêndice G.

Com base na abundância de cDNA, os níveis de transcritos envolvidos com degradação de biomassa foram comparados entre o *T. harzianum* IOC-3844 (biblioteca construída neste trabalho) e o *T. reesei* QM6a (cDNAs disponíveis publicamente) (DIENER et al., 2004). As porcentagens dos transcritos foram calculadas com relação ao número total de transcritos para cada biblioteca de cDNA.

As diferenças de expressão foram calculadas com base na razão entre as porcentagens de cada transcrito, ou de cada categoria de transcritos, entre os dois fungos. As significâncias estatísticas foram calculadas com uso do teste exato de “Fisher”. Para isto a contagem de cada transcrito, ou de cada categoria de transcritos, foi comparada entre as duas espécies, em uma abordagem par a par levando em conta o número total de seqüências das bibliotecas correspondentes.

### 3.8 Ensaios de PCR em tempo real

Para os ensaios com o micélio não induzido, cultivado em meio contendo 0,8 % de glicerol como fonte de carbono, foram usadas as coletas de 8, 36 e 60 horas de cultura. Para os ensaios com o micélio induzido, cultivado em meio contendo 1 % de Avicel como fonte de carbono, foram usadas as coletas de 8, 24, 36, 48 e 60 horas. As amostras de RNA total de cada ponto de cultura, contendo de 3 a 4 µg, foram tratadas com “DNase I Amplification Grade” (Invitrogen - 18068015) antes da síntese de cDNA. Os tratamentos com DNase I foram realizados em volume final de 10 µL para cada 1 µg de RNA total, com adição de 1 unidade de DNaseI (1 u/µL), 1 µL do tampão 10x da DNaseI e água deionizada livre de nuclease. As amostras foram incubadas a 25 °C por 15 minutos e a DNaseI foi inativada pela adição de EDTA 2,5 mM e incubação a 65 °C por 10 minutos. As sínteses de primeira fita de cDNA foram então realizadas a partir das amostras de RNA total com uso do “kit” “Improm II Reverse Transcription System” (Promega - A3800). Para isso foram montadas reações com volume final de 5 µL contendo 200 ng de RNA total, 1 µL de oligo-dT (500 ng/µL) e água deionizada livre de nuclease. As reações foram incubadas a 70 °C por 5 minutos e então imediatamente colocadas em gelo por 5 minutos. Totalizando um volume final de 20 µL, os seguintes reagentes foram misturados e adicionados às reações de 5 µL: 4 µL do tampão 5x Improm II; 2 µL de MgCl<sub>2</sub> 25 mM; 1 µL de dNTPs 10 mM; 0,5 µL de RNAsin (40 u/µL) e 1 µL da enzima reverse transcriptase. As amostras foram então incubadas a 25 °C por 5 minutos; 42 °C por 1 hora e 70 °C por 15 minutos.

Os oligonucleotídeos específicos usados nos experimentos de PCR em tempo real foram desenhados com auxílio do programa Primer 3 versão 4.0 disponível em <http://frodo.wi.mit.edu/primer3>. As seqüências dos oligonucleotídeos foram: Fw-qPCR-*cbh1*-harz 5'-tcggctcccgtctttactt-3'; Rv-qPCR-*cbh1*-harz 5'-atcggegtccatgaaac-3'; Fw-qPCR-*cbh2*-harz 5'-cgcatccttttggtatcg- 3'; Rv-qPCR-*cbh2*-harz 5'-tgggcgttgagcacttg-3'; Fw-qPCR-*egl1*-harz 5'-cttgaagccaactcaagagc-3'; Rv-qPCR-*egl1*-harz 5'-gatggtgaagaccttgagg-3'; Fw-qPCR-*egl2*-harz 5'-gattttggctgtgggactg-3'; Rv-qPCR-*egl2*-harz 5'-agggaggcggaaaatggt-3'; Fw-qPCR-*egl3*-harz 5'-cactgcgtcaaatcccaac-3'; Rv-qPCR-*egl3*-harz 5'-gggccaatgtctccgtatc-3'; Fw-qPCR-*egl7*-harz 5'-tgcacgactgtggacaaga-3'; Rv-qPCR-*egl7*-harz 5'-atcctgctgagcccagact-3'; Fw-qPCR-*swo*-harz 5'-ctggtcttgcaagcctta-3'; Rv-qPCR-*swo*-harz 5'-ccggatccattactc-3'; Fw-

qPCR-*act-harz* 5'-agaagttgctgcctcgtc-3' e Rv-qPCR-*act-harz* 5'-ggaaaacagctcggggag-3'.

Os ensaios de PCR em tempo real foram realizados no termociclador “PIXO Real-time PCR System” (Helixis). Todas as reações foram feitas com 5 µL do reagente “Platinum SYBR Green qPCR Supermix UDG” (Invitrogen - 11733-038) e 4 pMol de cada oligonucleotídeo em 10 µL de volume final. Para determinar as eficiências dos pares de oligonucleotídeos foram usadas diluições seriadas a partir de 1 µL das reações de síntese da primeira fita de cDNA. As eficiências determinadas para cada par de oligonucleotídeos foram: *cbh1* - 90,98%, *cbh2* - 97,83%, *egl1* - 91,62%, *egl2* - 91,65%, *egl3* - 90,13%, *egl7* - 99,56%, *swo* - 95,44% e *act* - 108,22%.

Para as análises de expressão foram usados 2 µL das reações de síntese da primeira fita de cDNA. As condições de ciclagem para todas as reações foram: 50 °C por 2 minutos, 95 °C por 2 minutos e 40 ciclos de 95 °C por 30 segundos, 56 °C por 30 segundos e 72°C por 40 segundos. As curvas de “melting” foram obtidas por incubação a 95 °C por 15 segundos, 55 °C por 15 segundos e 95 °C por 15 segundos. Foram realizadas também reações controle sem cDNA e sem a enzima transcriptase reversa. Cada amostra de RNA foi ensaiada com reações de PCR em tempo real em triplicata e as amostras com Cts mais discrepantes foram desprezadas (uma de três). As análises de expressão foram feitas com uso do método de  $2^{-\Delta\Delta Ct}$  (LIVAK & SCHMITTGEN, 2001) com modificações que permitiram calibrar pequenas diferenças entre as eficiências de amplificação dos pares de oligonucleotídeos utilizados, como proposto por (PFAFFL, 2001). Para obtenção dos valores de expressão diferencial as análises de expressão foram realizadas com uso do software “REST 2009 v2.0.13” (QIAGEN) disponível em <http://www.qiagen.com/REST>. Este software corrige diferenças de eficiência de oligonucleotídeos conforme o método proposto por (PFAFFL, 2001). Para a obtenção dos gráficos os valores de expressão diferencial gerados no programa REST foram transportados para o programa “GraphPad Prism v.5”, obtido em <http://www.graphpad.com/prism/>.

As diferenças de expressão foram calculadas a partir de comparações de valores de Cts realizadas par a par. O valor de  $\Delta\Delta Ct$  foi assim determinado:  $2^{-\Delta\Delta Ct}$ , onde  $\Delta\Delta Ct = (Ct \text{ do alvo} - Ct \text{ da actina}) - (Ct \text{ do alvo calibrador} - Ct \text{ da actina})$ . Foram realizadas três tipos de análises de expressão diferencial com diferentes alvos calibradores. Estes alvos calibradores foram: (i) os valores de Ct dos transcritos de micélio não induzido coletado em 8, 36 e 60 horas, para as comparações de níveis de expressão dos transcritos entre

micélio induzido e não induzido nos mesmos pontos de coleta; (ii) os valores de Ct dos transcritos de micélio induzido coletado em 8 horas, para comparações de níveis de expressão entre os transcritos de micélios induzidos de 24, 36, 48 e 60 horas e os transcritos de micélio induzido por 8 horas e (iii) a média dos valores de Ct de actina induzida em todos os tempos de coleta (8, 24, 36, 48 e 60 horas), para comparações de níveis de expressão entre os transcritos induzidos de 8, 24, 36, 48 e 60 horas e a média da actina induzida de 8, 24, 36, 48 e 60 horas.

### **3.9 Clonagem de enzimas de *Trichoderma harzianum* IOC-3844 em plasmídeos para expressão em *Pichia pastoris***

As seqüências das celulases *cbh1* (GenBank: AF223252) e *bgl2* (GenBank: EF426298) de *Trichoderma harzianum* IOC-3844, disponíveis publicamente (BENSON et al., 2011), foram usadas para o desenho de oligonucleotídeos para obtenção das ORFs via RT-PCR com uso da enzima “Platinum Taq DNA Polymerase High Fidelity” (Invitrogen - 11304-011). A PCR foi realizada a partir da síntese da primeira fita de cDNA com uso do RNA total isolado de micélio induzido. A reação de RT foi feita como descrito acima no item “Análises de RT-PCR”, exceto pela utilização do oligonucleotídeo reverso oligo-dT, ao invés de oligonucleotídeos reversos gene específicos. As demais ORFs foram isoladas por PCR, com uso da enzima “Platinum Taq DNA Polymerase High Fidelity” (Invitrogen - 11304-011), utilizando os plasmídeos pDONR-222 (Invitrogen) recombinantes da biblioteca de cDNA como molde. Todas as ORFs para expressão foram amplificadas com oligonucleotídeos específicos contendo sítios de enzima de restrição para posterior clonagem nos plasmídeos de expressão (Tabela 1). Os produtos de PCR foram digeridos com enzimas de restrição e clonados nos plasmídeos de expressão. As fusões 5’ e 3’ das ORFs clonadas foram confirmadas por seqüenciamento com uso do seqüenciador “MegaBACE 1000 Flex” (GE Healthcare). As informações resumidas sobre as clonagens realizadas podem ser vistas na tabela 1. Algumas enzimas tiveram suas ORFs clonadas com o peptídeo sinal nativo nos plasmídeos pGAP-Z e pPIC-Z. Outras foram clonadas em fusão com o fator “ $\alpha$ ”, que é o peptídeo sinal presente nos plasmídeos de expressão pGAP-Z- $\alpha$  e pPIC-Z- $\alpha$ , que, assim como o peptídeo sinal nativo, também faz com que a proteína recombinante seja secretada para o meio de cultura. As ORFs fusionadas com o fator “ $\alpha$ ” tiveram o peptídeo sinal removido depois do mapeamento

do ponto de clivagem com uso da ferramenta “SignalP 4.0 Server” disponível em <http://www.cbs.dtu.dk/services/SignalP/>. Para a CBHI, além da ORF, o seu domínio catalítico (DC) também foi clonado.

**Tabela 1: Resumo sobre as clonagens de celulases e hemicelulases de *Trichoderma harzianum* IOC-3844 em plasmídeos para expressão em *Pichia pastoris*.** As seqüências sublinhadas mostram os sítios das enzimas de restrição usadas nas clonagens.

Proteína	Pep. sinal nativo	Oligonucleotídeo Fw (5'-3')	Oligonucleotídeo Rv (5'-3')	Plasmídeo de Expressão
CBHI	Sim	<u>gggaattc</u> cacgatgtatcggaagttggcc	ggggtc <u>gaccaggc</u> actgagagtagaatg	pPIC-Z-A
CBHI	Não	<u>gggaattc</u> cagcaggtctgcaccca	ggggtc <u>gaccaggc</u> actgagagtagaatg	pPIC-Z- $\alpha$ -A
CBHI (DC)	Não	<u>gggaattc</u> cagcaggtctgcaccca	ggggtc <u>gaccaggc</u> actgagagtagaatg	pPIC-Z- $\alpha$ -A
EGI	Não	<u>ccgaattc</u> cagcaacctggaaccagca	<u>ccgcggccgct</u> aggcattgagtagtaaatcg	pPIC-Z- $\alpha$ -A
EGIII	Não	<u>cggaaattc</u> cagaccagctgcgaaca	<u>cgcctc</u> gacgctgatagatgcggctcc	pPIC-Z- $\alpha$ -A
EGIV	Sim	<u>ccttcgaaat</u> ggctcaaaagctttcc	<u>cgtctaga</u> aatctggaacgcactgagc	pPICZ- $\alpha$ -C*
EGIV	Não	<u>ccatcgat</u> ccatggacatgtcaaca	<u>cgtctaga</u> aatctggaacgcactgagc	pPICZ- $\alpha$ -C
EGIV	Sim	<u>ccttcgaaat</u> ggctcaaaagctttcc	<u>cgtctaga</u> aatctggaacgcactgagc	pGAPZ-C
EGIV	Não	<u>ccatcgat</u> ccatggacatgtcaaca	<u>cgtctaga</u> aatctggaacgcactgagc	pGAP-Z- $\alpha$ -C
EGVII	Não	<u>cggaaattc</u> catggacagatccaaaac	<u>gctctaga</u> aaagccttgccacagggct	pPIC-Z- $\alpha$ -A
BGL2	Não possui	<u>gggaattc</u> atgttgcceaagacttt	<u>gctctaga</u> aaagcctttgcctctctt	pPIC-Z- $\alpha$ -A
Xyn2	Não	<u>ccgaattc</u> gctcccactgaatcctcc	<u>cccgtc</u> gacgctgacagtgatggaagca	pPIC-Z- $\alpha$ -A
Xyn3	Sim	<u>ccttcgaaat</u> gaaagcaaatataatatttgc	<u>ccgtc</u> gacttgtaaatgcttacaatactg	pPICZ-C
Xyn3	Não	<u>ccatcgat</u> cctcccggccgaattcattcga	<u>ccgtc</u> gacttgtaaatgcttacaatactg	pPICZ- $\alpha$ -C
Xyn3	Sim	<u>ccttcgaaat</u> gaaagcaaatataatatttgc	<u>ccgtc</u> gacttgtaaatgcttacaatactg	pGAPZ-C
Xyn3	Não	<u>ccatcgat</u> cctcccggccgaattcattcga	<u>ccgtc</u> gacttgtaaatgcttacaatactg	pGAP-Z- $\alpha$ -C

\*O plasmídeo pPICZ- $\alpha$ -C usado para clonar a EGIV fica sem o fator “ $\alpha$ ” devido ao uso da enzima *Bst*BI que cliva no sítio “ttcga”. Com isso o plasmídeo fica equivalente ao pPICZ-C.

### 3.10 Expressão de enzimas de *Trichoderma harzianum* IOC-3844 em *Pichia pastoris*

Foram realizados ensaios de expressão para as construções EGIII em pPIC-Z- $\alpha$ -A e Xyn2 em pPICZ- $\alpha$ -A. Ensaios de expressão em *Pichia pastoris* de outras enzimas de *Trichoderma harzianum* IOC-3844 foram ou estão sendo realizados por outros estudantes do Laboratório de Biologia Molecular do DGE-UFSCar. Os resultados resumidos destes ensaios podem ser vistos no apêndice F. Nos plasmídeos pPIC a expressão heteróloga é dirigida pelo promotor do gene da enzima álcool oxidase (AOX) que responde ao estímulo de metanol. O protocolo de indução foi realizado conforme as recomendações do manual de expressão “Easy Select Pichia Expression Kit, A Manual of Methods for Expression of Recombinant Proteins Using pPICZ and pPICZ $\alpha$  in

*Pichia pastoris*”, versão G - K1740-01 - Invitrogen. Os principais passos estão resumidos a seguir: de 1 a 5 µg de plasmídeos recombinantes foram linearizados com a enzima de restrição *PmeI* e utilizados para transformar a cepa KM71H muts de *P. pastoris* (Invitrogen) por eletroporação. As leveduras transformadas foram selecionadas por plaqueamento em meio contendo zeocina (Invitrogen) nas concentrações de 100 a 500 µg/mL. As colônias selecionadas foram analisadas através de PCR de colônia para confirmar a presença dos plasmídeos recombinantes no genoma. A busca por clones que expressam as enzimas clonadas em pPICs foi realizada com o crescimento das colônias a 250 RPM e 30 °C por dois dias em 3 mL de meio BMGY. Depois deste período o material foi centrifugado e ressuspendido em 2 mL de meio de indução BMMY, depois incubado a 300 RPM e 30 °C por até 6 dias. Estes ensaios tiveram como controle negativo, a indução de leveduras transformadas com os mesmos plasmídeos usados para a expressão, porém sem clones recombinantes, ou “vazios”. Os controles positivos foram leveduras transformadas com outras proteínas do nosso laboratório, sabidamente expressas nas condições utilizadas. Os tempos de indução foram ajustados conforme os clones ensaiados e os pHs dos meios de cultura variaram de 6,0 a 6,5. As induções foram confirmadas por eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio ou “SDS-PAGE” corado com “Coomassie Brilliant Blue” (SAMBROOK & RUSSELL, 2001) com uso alíquotas de 15 µL dos sobrenadantes das culturas. Após a verificação do sítio de clivagem de peptídeo nos plasmídeos de expressão, o peso molecular das enzimas recombinantes foi calculado com uso do programa “GENE RUNNER” versão 3.01, disponível em <http://www.generunner.net/>. Os clones selecionados em pPIC foram induzidos em maior escala com inóculos em 500 mL de meio BMGY a 250 RPM e 30 °C por 24 horas. Depois o material foi centrifugado e ressuspendido em 100 mL de meio BMMY a 300 RPM e 30 °C para indução por 1 a 6 dias.

Os sobrenadantes dos meios de cultura foram recuperados por centrifugação ao término do período de indução. Para purificação adicional, as proteínas recombinantes foram capturadas pela passagem dos sobrenadantes induzidos em colunas contendo a resina de níquel “Ni-NTA Superflow” (QIAGEN). As colunas contendo 5 mL de resina, que haviam sido mantidas em etanol 20 %, foram previamente lavadas com 3 volumes de água deionizada e equilibradas com 3 volumes de tampão de lise sem imidazol (Tris HCl 10 mM; NaH<sub>2</sub>PO<sub>4</sub> 50 mM; NaCl 100 mM) pH 8,0 ajustado com NaOH. Após a passagem do sobrenadante de cultura de *Pichia pastoris* transformante, os eluatos foram

recuperados para análise e as colunas contendo as proteínas foram lavadas com tampão de lise sem imidazol. As proteínas recombinantes foram então eluídas da coluna com a passagem de volumes de 5 mL de concentrações crescentes de imidazol, no tampão de lise conforme segue: 10 mM-a; 10 mM-b; 25 mM-a; 25 mM-b; 50 mM-a; 50 mM-b; 75 mM-a; 75 mM-b; 100 mM-a; 100 mM-b; 250 mM-a; 250 mM-b e 250 mM-c. A captura das proteínas recombinantes foi realizada pela afinidade do níquel aos anéis imidazólicos presentes nos aminoácidos, que formam a cauda de histidinas da proteína recombinante, oriunda da fusão dos clones com os plasmídeos pPIC. A eluição se deu devido à competição do imidazol pelo níquel. Os purificados foram analisados por SDS-PAGE corado com “Coomassie Brilliant Blue” (SAMBROOK & RUSSELL, 2001). As quantificações de proteína recombinante foram realizadas com uso do “kit” “Pierce BCA Protein Assay Kit” - Thermo Scientific (SMITH et al., 1985).

### **3.11 Ensaios de atividade de enzimas recombinantes de *Trichoderma harzianum* IOC-3844**

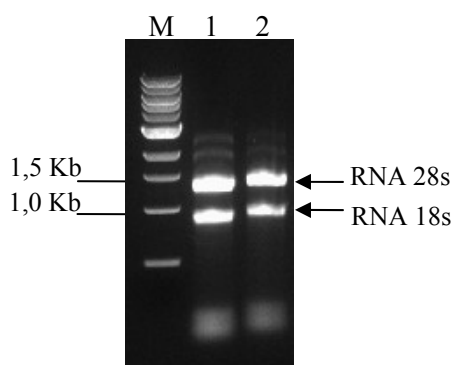
As enzimas EGIII e Xyn2 de *Trichoderma harzianum* IOC-3844 produzidas em *Pichia pastoris* foram analisadas quanto à atividade com uso do corante Vermelho Congo, que possui afinidade aos polissacarídeos solúveis e que é sensível à mudança de pH que ocorre devido à produção de açúcares redutores oriundos da hidrólise enzimática. Ensaios de atividade de outras enzimas de *Trichoderma harzianum* IOC-3844 produzidas em *Pichia pastoris* foram ou estão sendo realizados por outros estudantes do Laboratório de Biologia Molecular do DGE-UFSCar e os resultados estão resumidos no apêndice F. Para os ensaios de atividade 10 µL dos sobrenadantes de cultura induzida de *Pichia pastoris* ou 10 µL das frações purificadas a partir dos sobrenadantes, foram adicionados à superfície do substrato de carboximetilcelulose ou CMC (Sigma-C5678) para a enzima EGIII, ou de Xilano (X0502 - Sigma) para a enzima Xyn2. Os substratos foram preparados com 0,5 % de CMC (para EGIII) ou 0,5 % de Xilano (para Xyn2) e 1,5 % de ágar bacteriológico (Himedia - RM026). O material foi incubado por 20 minutos a 28 °C para a enzima EGIII em CMC, e por uma hora a 30 °C para a enzima Xyn2 em Xilano. As placas foram então banhadas por 30 minutos com Vermelho Congo 0,2 % (Vetec - 600) e reveladas com NaCl 1M por 10 minutos.



## 4 RESULTADOS

### 4.1 Extração de RNA total

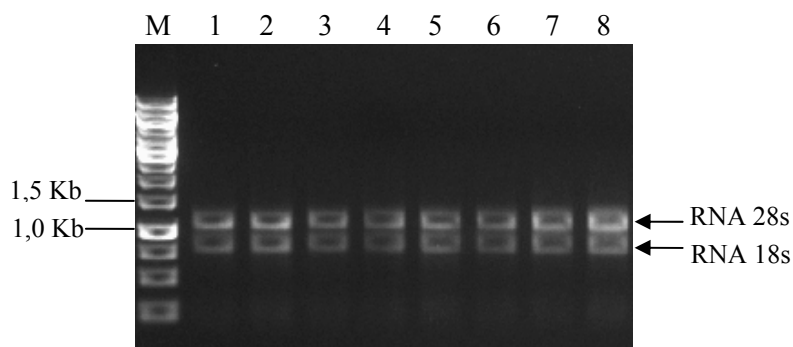
Foram realizadas extrações de RNA total, a partir de micélio induzido cultivado por 60 horas com celulose microcristalina (Avicel) e a partir de micélio não induzido cultivado por 60 horas, para análise por RT-PCR e posterior construção da biblioteca de cDNA de *Trichoderma harzianum* IOC-3844. Nestas extrações cerca de 50 µg de RNA total foram obtidos a partir de 120 mg de micélio. A qualidade do material foi confirmada por meio de medidas de densidade óptica em espectrofotômetro, e pela análise de eletroforese em gel de agarose 1 % (Figura 1). As medidas de densidade óptica no espectrofotômetro também foram realizadas para quantificar as amostras de RNA.



**Figura 1: Gel de agarose 1 % corado com brometo de etídeo contendo RNA total extraído de micélio de *Trichoderma harzianum* IOC-3844 cultivado por 60 horas. M: Marcador de massa molecular 1kb (Biolabs); 1: RNA de micélio não induzido; 2: RNA de micélio induzido. As setas indicam os RNAs 28S e 18S.**

Para as análises por PCR em tempo real foram realizadas extrações adicionais de RNA total de *Trichoderma harzianum* IOC-3844, a partir de micélio induzido por Avicel e cultivado por 8, 24, 36 e 48 horas; e a partir de micélio não induzido cultivado por 8, 36 e 60 horas. Nestas extrações, em média, cerca de 50 µg de RNA total foram obtidos a partir de 120 mg de micélio. Os RNAs foram tratados com DNaseI para garantir a pureza do material eliminando a possibilidade de interferências de DNA genômico nos ensaios posteriores. Após o tratamento com DNaseI a qualidade dos

RNAs foi verificada por análise de eletroforese em gel de agarose 1 % (Figura 2) e pelas medidas de densidade óptica em espectrofotômetro. As medidas de densidade óptica no espectrofotômetro também foram realizadas para quantificar as amostras de RNA.



**Figura 2: Gel de agarose 1 % corado com brometo de etídeo contendo RNA total extraído de micélio de *Trichoderma harzianum* IOC-3844 e tratado com DNaseI.** M: Marcador de massa molecular 1kb “GeneRuler” (Fermentas); 1, 2, 3, 4 e 5: RNA extraído de micélio induzido por 8, 24, 36, 48, e 60 horas; 6, 7 e 8: RNA extraído de micélio não induzido por 8, 36 e 60 horas. As setas indicam os RNAs 28S e 18S.

#### **4.2 Análise de expressão de celulases por RT-PCR**

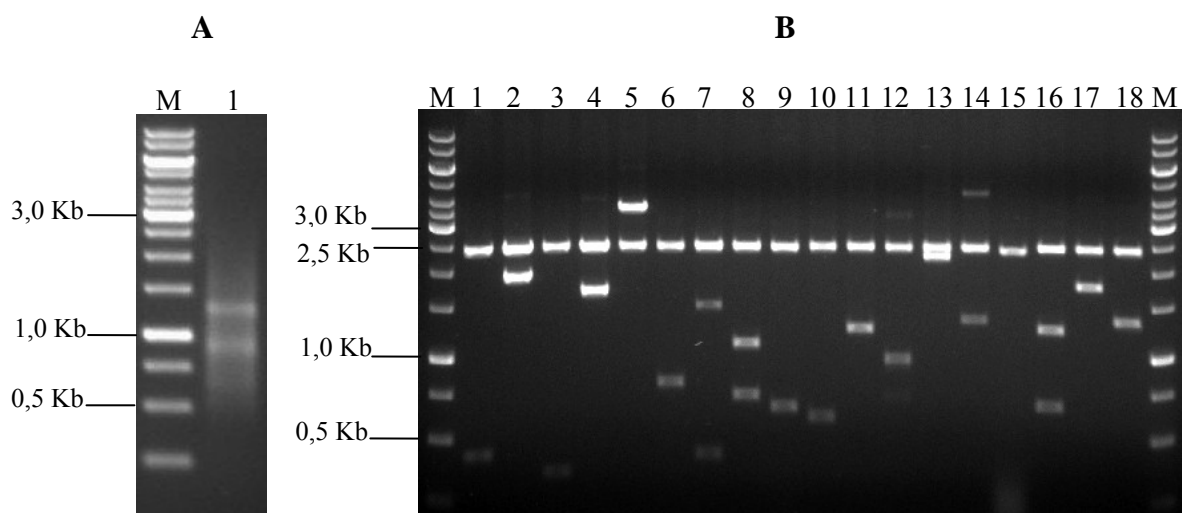
As amostras de RNA coletadas em 60 horas de cultivo de micélio induzido e não induzido foram avaliadas quanto à expressão de celulases por RT-PCR. A análise em gel de agarose 1 % do revelou expressão aumentada de *cbh1* e *bgl2* em micélio induzido em comparação ao micélio não induzido (Figura 3).



**Figura 3: Gel de agarose 1 % corado com brometo de etídeo com resultado de RT-PCR em micélio de *Trichoderma harzianum* IOC-3844 cultivado por 60 horas.** M: Marcador de massa molecular 1kb “GeneRuler” (Fermentas). 1: *cbh1* não induzido; 4: *cbh1* induzido; 7: *bgl2* não induzido; 10: *bgl2* induzido; 13: *actina* não induzido; 16: *actina* induzido; 2, 5, 8, 11, 14 e 17: Controles sem transcriptase reversa; 3, 6, 9, 12, 15 e 18: Controles sem molde. As setas indicam os fragmentos amplificados de *actina* (455pb), *bgl2* (220pb) e *cbh1* (185pb).

#### 4.3 Extração de RNA mensageiro e construção da biblioteca de cDNA

Uma quantidade 1,3 µg de RNA mensageiro foi obtida a partir de 240 µg de RNA total extraído de micélio induzido por 60 horas. A qualidade do RNA mensageiro foi confirmada por gel de agarose 1 % (Figura 4 A) e pelas medidas de densidade óptica em espectrofotômetro, que serviram também para quantificar o mRNA. O título da biblioteca foi de  $1,2 \times 10^8$  clones, o que corresponde a grande representatividade. Após análise de restrição de 18 plasmídeos pDONR-222 (invitrogen) recombinantes, com a enzima *Bsr*GI, a média do tamanho dos clones de cDNA calculada foi de 1,7 kb (Figura 4 B).



**Figura 4:** (A) Gel de agarose 1% corado com brometo de etídeo com resultado de extração de RNA mensageiro de *Trichoderma harzianum* IOC-3844 induzido com celulose por 60 horas, (B) Gel de agarose 1% corado com brometo de etídeo com resultado de análise de clones da biblioteca por restrição com *Bsr*GI. (A) M: Marcador de massa molecular 1kb “GeneRuler” (Fermentas); 1: RNA mensageiro. (B) M: Marcador de massa molecular 1kb “GeneRuler” (Fermentas); 1 a 18: Clones digeridos.

#### 4.4 Análise de sequencias ESTs

Depois do seqüenciamento dos clones da biblioteca induzida de cDNA de *T. harzianum* IOC-3844, 2016 seqüências brutas foram analisadas na plataforma de softwares dCAS (GUO et al., 2009). Após a filtragem das seqüências e a remoção de seqüências de oligonucleotídeo e de plasmídeos, um total de 1544 seqüências de boa qualidade foram organizadas em 1003 transcritos, sendo 167 “contigs”, formados por mais de uma seqüência, e 836 ESTs que permaneceram isolados em seqüência única ou “singlet”. As 1544 seqüências de boa qualidade estão sendo submetidas ao banco de dados “GenBank” (BENSON et al., 2011). Após analisar os resultados das buscas via “blastx” (ALTSCHUL et al., 1990) dos 1003 transcritos obtidos, 357 tiveram funções associadas e 646 não apresentaram funções conhecidas. Entre os transcritos com função associada, 20 foram mapeados com funções relacionadas à desconstrução de biomassa vegetal (Tabela 2).

Como o seqüenciamento da biblioteca foi realizado a partir das extremidades 5' dos clones de cDNA, por meio da análise do alinhamento resultante do “blastx” os transcritos que estão com as ORFs completas nos plasmídeos pDONR-222 (Invitrogen) puderam ser identificados. Quase todos os transcritos listados na tabela 2 tiveram suas ORFs completamente clonadas, provavelmente devido ao tamanho médio elevado dos clones da biblioteca que foi de 1,7 kb (Figura 4). As exceções foram os transcritos correspondentes às proteínas Cip1 (proteína glicolítica candidata ligante de celulose) e ACEI (fator de transcrição tipo “zinc finger”) que alinharam a partir dos aminoácidos 14º e 6º respectivamente. Além disso, com exceção dos transcritos *egl2*, *bgl2*, *bgl5*, *bxl1*, *man1*, *cip1*, *cip2*, *xyl1* e *acel*, os demais transcritos tiveram suas ORFs completamente seqüenciadas com auxílio de oligonucleotídeos adicionais. O apêndice F mostra os transcritos de *T. harzianum* IOC 3844, listados na tabela 2, que tiveram as ORFs completamente seqüenciadas. Foram realizados alinhamentos da tradução das ORFs destes transcritos com os depósitos das proteínas correspondentes em *T. reesei*. Estes alinhamentos e os valores de similaridade de seqüência entre *T. harzianum* e *T. reesei* podem ser vistos no apêndice G.

**Tabela 2: Transcritos de *Trichoderma harzianum* IOC-3844 relacionados à desconstrução de biomassa.** As contagens de ESTs estão mostradas em ordem decrescente dentro de cada categoria.

Categoria	Proteínas	Número de ESTs	Organismos do melhor resultado do “blastx” nr	Número de depósito [GenBank]	Tamanho do alinhamento (aa)	Tamamho da Proteína “blastx” (aa)	Identidade do alinhamento %	E-valor
Celulases	CBHI	34	<i>H. lixii</i>	[Q9P8P3]	172	505	98	3E-98
	CBHII	28	synthetic construct	[CBX74420]	295	473	93	1E-156
	EGVII	16	<i>H. rufa</i>	[ACD36971]	249	249	86	1E-130
	BGL1	13	<i>T. reesei</i>	[AAA18473]	396	744	85	0
	EGIII	6	<i>H. schweinitzii</i>	[AAM77711]	173	234	82	1E-81
	BGL2	6	<i>H. lixii</i>	[ABN50090]	144	453	100	4E-81
	BGL5	4	<i>T. reesei</i>	[EGR49111]	196	484	91	1E-109
	EGI	4	<i>H. jecorina</i>	[ADM08177]	238	459	86	1E-120
	EGIV	3	<i>Trichoderma sp.</i>	[ACH92573]	250	347	81	1E-123
	EGII	3	<i>T. viride</i>	[BAA36216]	142	418	60	3E-47
	EGV	2	<i>H. rufa</i>	[AAQ21385]	100	247	82	6E-42
	Hemicelulases	Xyn2	27	<i>H. lixii</i>	[ACF40831]	221	221	100
Man1		8	<i>T. longibrachiatum</i>	[ADN93457]	202	437	87	3E-99
Xyn3		7	<i>H. jecorina</i>	[BAA89465]	168	347	86	2E-77
Bxl1		1	<i>H. jecorina</i>	[CAA93248]	155	797	92	2E-77
Acessórias	SWOI	1	<i>H. pseudokoningii</i>	[ABV57767]	157	494	82	4E-69
Putativas	Cip1	2	<i>H. jecorina</i>	[EGR52696]	173	316	83	2E-84
	Cip2	2	<i>H. jecorina</i>	[EGR44948]	143	460	76	5E-58
Fatores de Transcrição	Xyr1	2	<i>T. reesei</i>	[EGR48040]	188	920	87	1E-94
	ACEI	2	<i>T. reesei</i>	[EGR51484]	195	733	94	2E-98

Entre as 17 celulases descritas para *T. reesei* (ver introdução), foram clonados neste trabalho 11 transcritos homólogos do *T. harzianum* IOC-3844 (Tabela 2). Como só havia as ORFs das proteínas CBHI e BGL2 de *T. harzianum* publicamente disponíveis na época do início deste trabalho, estão sendo reportadas aqui 9 seqüências adicionais de celulases para este organismo. Entre estas 9 celulases a EGIII do *Trichoderma harzianum* IOC-3844 está disponível publicamente GenBank:AFK32784 (BENSON et al., 2011).

Entre as 5 xilanases descritas para *T. reesei* (ver introdução) foram clonados 3 transcritos homólogos do *T. harzianum* IOC-3844 (Tabela 2). Como só existe a ORF da proteína Xyn2 de *T. harzianum* publicamente disponível, estão sendo reportadas aqui 2 seqüências adicionais de xilanases para este organismo. A seqüência da  $\beta$ -mananase Man1, previamente descrita para *T. reesei* (ver introdução), também foi clonada na biblioteca de *T. harzianum* IOC-3844 e está sendo reportada aqui pela primeira vez para este fungo (Tabela 2).

A enzima acessória swolenina, previamente descrita em *T. reesei*, e co-regulada com celulasas neste fungo (SALOHEIMO et al., 2002; FOREMAN et al., 2003) está sendo reportada para *T. harzianum* pela primeira vez neste trabalho (Tabela 2).

As enzimas candidatas à interação com celulose e prováveis glicolíticas Cip1 e Cip2, previamente descritas para *T. reesei* (FOREMAN et al., 2003), também estão sendo pela primeira vez reportadas aqui para *T. harzianum* (Tabela 2).

Finalmente, a tabela 2 também mostra o regulador de xilanase Xyr1 e o fator de transcrição tipo “zinc finger” ACEI, ambos também já descritos em *T. reesei* (RAUSCHER et al., 2006; SALOHEIMO et al., 2000), e sendo reportados pela primeira vez neste trabalho para *T. harzianum*. A proteína Xyr1 é o principal ativador geral da expressão de celulasas e hemicelulasas (STRICKER et al., 2006; MACH-AIGNER et al., 2008) enquanto que a proteína ACEI é repressora de atividades celulásicas e xilanásicas (ARO et al., 2003; RAUSCHER et al., 2006).

Uma biblioteca induzida de *T. reesei* QM6a, disponível publicamente (DIENER et al., 2004) também foi analisada neste trabalho com uso da plataforma dCAS (GUO et al., 2009). Após analisar 22819 seqüências de boa qualidade, os transcritos de *T. reesei* QM6a envolvidos com degradação de biomass vegetal totalizaram 28, os quais foram separados em 24 enzimas e 4 fatores de transcrição (Tabela 3). Assim como feito para *T. harzianum* IOC-3844 (Tabela 2) os transcritos relacionados à degradação de biomassa listados para *T. reesei* QM6a (Tabela 3) também tiveram seus alinhamentos com os melhores “hits” do “blastx” (ALTSCHUL et al., 1990) verificados para satisfazerem um critério mínimo de confiabilidade com “E-values” < 1e-10, identidade > 40 % e tamanho de alinhamento > 100 aminoácidos (dados não mostrados). Com isso, uma comparação de níveis de expressão entre *T. harzianum* IOC-3844 e *T. reesei* QM6a foi realizada com base na contagem de ESTs (Tabela 3). Alguns transcritos da tabela 3 apareceram em *T. reesei* QM6a e não apareceram em *T. harzianum* IOC-3844. São eles: as celulasas EGVI, EGVIII, BGL3, BGL6 e BGL7; as hemicelulasas Axe1, Axe2 e

Abf2; e os fatores de transcrição ACE2 e CRE1. Isto pode ter ocorrido por que a biblioteca de *T. reesei* QM6a teve um maior número de clones analisados em relação à biblioteca de *T. harzianum* IOC-3844 (22819 clones e 1544 respectivamente). No entanto, sem um trabalho exaustivo de seqüenciamento de cDNAs, não há como afirmar se estes transcritos apresentam baixa expressão relativa ou se até mesmo não ocorrem em *T. harzianum* IOC-3844. Por outro lado, as celulases EGIII e EGV, presentes em *T. harzianum* IOC-3844, não revelaram transcritos em *T. reesei* QM6a (Tabela 3). Isto indica expressão aumentada destes transcritos em *T. harzianum* IOC-3844, apesar de não haver como calcular a razão de expressão entre os dois fungos nestes casos (Tabela 3).

De modo geral a tabela 3 sugere maiores níveis de expressão dos transcritos de *Trichoderma harzianum* IOC-3844 em relação aos transcritos *Trichoderma reesei* QM6a. Estes resultados corroboram com os resultados de (CASTRO et al., 2010a) que mostraram maiores atividades celulolíticas (FPase, endoglucanase e  $\beta$ -glucosidase) de *Trichoderma harzianum* IOC-3844 em relação ao *Trichoderma reesei* RutC30. As diferenças entre as contagens de ESTs listadas na tabela 3 para os transcritos destes dois fungos estão comentadas mais adiante na discussão.



**Tabela 3: Comparação de abundância de transcritos, envolvidos com degradação de biomass vegetal, entre *Trichoderma harzianum* IOC-3844 e *Trichoderma reesei* QM6a.** Nos casos com contagem de ESTs de valor zero, os valores estatísticos foram determinados, mas as razões de expressão não puderam ser calculadas.

Categoria	Proteína	<i>T. harzianum</i> IOC-3844		<i>T. reesei</i> QM6a		Razão	p-valor
		N	%	N	%	<i>harz/reesei</i>	
Celulases	CBHI	34	2,2	134	0,59	3,75	< 0,0001
	CBHII	28	1,81	56	0,25	7,39	< 0,0001
	EGI	4	0,26	10	0,04	5,91	0,0006
	EGII	3	0,19	12	0,05	3,69	0,03
	EGIII	6	0,39	0	0	-	< 0,0001
	EGIV	3	0,19	9	0,04	4,93	0,008
	EGV	2	0,13	0	0	-	< 0,0001
	EGVI	0	0	19	0,08	-	< 0,0001
	EGVII	16	1,04	20	0,09	11,82	< 0,0001
	EGVIII	0	0	12	0,05	-	< 0,0001
	BGL1	13	0,84	3	0,01	64,04	< 0,0001
	BGL2	6	0,39	5	0,02	17,73	< 0,0001
	BGL3	0	0	4	0,02	-	< 0,0001
	BGL5	4	0,26	3	0,01	19,71	< 0,0001
	BGL6	0	0	1	0	-	< 0,0001
BGL7	0	0	5	0,02	-	< 0,0001	
Total		119	7,71	293	1,28	6	< 0,0001
Hemicelulases	Xyn2	27	1,75	8	0,04	49,88	< 0,0001
	Xyn3	7	0,45	6	0,03	17,24	< 0,0001
	Bxl1	1	0,06	7	0,03	2,11	0,4745*
	Axe1	0	0	7	0,03	-	< 0,0001
	Axe2	0	0	2	0,01	-	< 0,0001
	Abf2	0	0	2	0,01	-	< 0,0001
	Man1	8	0,52	4	0,02	29,56	< 0,0001
	Total		43	2,78	36	0,16	17,65
Acessória	SWOI	1	0,06	6	0,03	2,46	0,3882*
Putativa	Cip1	2	0,13	18	0,08	1,64	0,5017*
	Cip2	2	0,13	4	0,02	7,39	0,0067
	Total		4	0,26	22	0,1	2,69
Fatores Transcr.	ACE1	2	0,13	12	0,05	2,46	0,2225*
	ACE2	0	0	3	0,01	-	< 0,0001
	Xyr1	2	0,13	3	0,01	9,85	< 0,0001
	CRE1	0	0	6	0,03	-	< 0,0001
ESTs restantes		1374	88,99	22450	98,38	0,9	0,006
Total de ESTs		1544	100	22819	100	1	-

\* Valores estatísticos não significativos

#### 4.5 Busca de seqüências de *Trichoderma harzianum* relacionadas com degradação de biomass vegetal

Para checar a existência de seqüências disponíveis publicamente de *T. harzianum*, relacionadas com a degradação de biomass vegetal, foi realizada uma busca via “blastn” com parâmetros “default” (ALTSCHUL et al., 1990), contra a coleção de nucleotídeos (nr/nt), usando todas as ORFs descritas para *T. reesei*, correspondentes às proteínas listadas na tabela 3. A maior parte dos números de acesso destas ORFs está descrita em (FOREMAN et al., 2003). Os demais números de acesso usados foram: *ace1* [GenBank:AF190793], *ace2* [GenBank:AF220671] *xyl1* [GenBank:AF479644], *cbh1* [GenBank:AY368686] e *cre1* [GenBank:X95368] (BENSON et al., 2011). Esta busca retornou as ORFs de *T. harzianum* para *cbh1*: [GenBank:AF223252], para *bgl2*: [GenBank:EF426298] e [GenBank:EF426299] e para *cre1* [GenBank:X95369]; além do nosso depósito da ORF de *egl3* [GenBank:JQ923478] (BENSON et al., 2011). Usando as mesmas seqüências foram realizadas buscas via “blastx” com parâmetros “default” (ALTSCHUL et al., 1990), contra o banco de seqüências não redundantes de proteínas (nr). Com isso, foram encontradas as seqüências de *T. harzianum* da proteína CBHI: [GenBank:Q9P8P3] e [GenBank:ADH04808]; da BGL2: [GenBank:ABN50090] e [GenBank:ABN50091]; da Xyn2: [GenBank:ACF40831] e [GenBank:P48793] e da CRE1 [GenBank:CAA64656]; além do nosso depósito da EGIII [GenBank:AFK32784], (BENSON et al., 2011).

Ainda no intuito de mapear seqüências de *T. harzianum* disponíveis publicamente, e relacionadas com degradação de biomass, as ORFs de *T. reesei* correspondentes às proteínas listadas na tabela 3 foram usadas para buscar ESTs contra o banco de dados “expressed sequence tags (est)” via “blastn” usando os parâmetros “default” (ALTSCHUL et al., 1990). Com isso foram encontrados 8 ESTs para *T. harzianum* (Tabela 4). Para identificar estes ESTs, os mesmos foram usados para realizar buscas nos bancos de coleção de nucleotídeos (nr) e de seqüências não redundantes de proteínas (nr), via “blastn” e “blastx” respectivamente usando os parâmetros “default” (ALTSCHUL et al., 1990). Depois disso, foram encontrados as sequencias de nucleotídeo e de aminoácido correspondentes a 7 diferentes proteínas relacionadas à degradação de biomass vegetal (Tabela 4). Comentários adicionais sobre estas buscas se encontram na discussão.

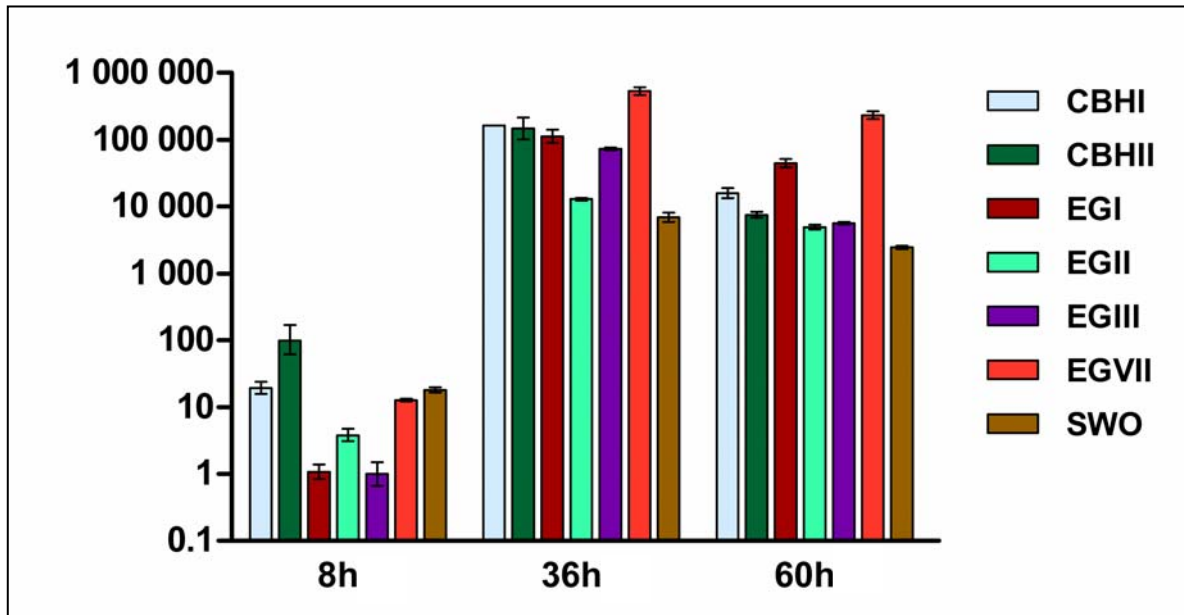
**Tabela 4: ESTs publicamente disponíveis de *Trichoderma harzianum* associados com degradação de biomass vegetal.**

ESTs de <i>T. harzianum</i> [GenBank]	Proteína	Depósitos de DNA [GenBank]	Depósitos de Proteína [GenBank]	Alinhamento do “blastx”	Tamanho do alinhamento (aa)	Tamanho da proteína (aa)	Identidade %	E-valor
[CK907860]	BGL2	[EF426299]	[ABN50090]	1°	200	453	96	4E-140
[AJ899631]	BGL5	[AY281377]	[EHK23361]	3°	222	484	91	9E-152
[HS574263]	BGL3	[AY281374]	[EGR48517]	2°	73	874	96	2E-43
[AJ896429]	Axe2	[AY281376]	[EHK24306]	2°	131	299	79	1E-68
[FG342168]	ACE1	[AF190793]	[EHK49917]	3°	146	733	86	1E-67
[DY761707]	Xyr1	[AF479644]	[EGR48040]	1°	42	920	100	1E-18
[AJ895295]	BGL2	[EF426299]	[ABN50090]	1°	181	453	100	8E-131
[AJ904356]	CRE1	[X95369]	[CAA64656]	2°	113	409	96	3E-66

Finalmente, as últimas buscas por sequencias de *T. harzianum*, relacionadas à desconstrução de biomassa, foram realizadas usando o nome da espécie contra os bancos de dados “Protein Data Bank - PDB” (BERMAN et al., 2000) e o “Proteinknowledgebase - UniProtKB” (WU et al., 2006). Com isso foram encontrados apenas: a estrutura da CBHI (2Y9L e 2Y9N) e da Xyn2 (1XND) no “PDB”, e as ORFs das proteínas CBHI, BGL2, Xyn2 e CRE1 no “UniProtKB”.

#### 4.6 PCR em tempo real e análises de expressão

Comparando os valores de expressão entre o micélio induzido e não induzido de *T. harzianum* IOC-3844, nos pontos de coleta de 8, 36 e 60 horas de cultivo, a maioria dos transcritos analisados se mostrou superexpressa em todos os pontos de amostragem, com exceção das endoglucanases *egl1* e *egl3* em 8 horas de indução. Entre os demais transcritos superexpressos em 8 horas de indução, o de *cbh2* mostrou o maior valor, seguido por *cbh1*, *swol1*, *egl7* e *egl2* (Figura 5).

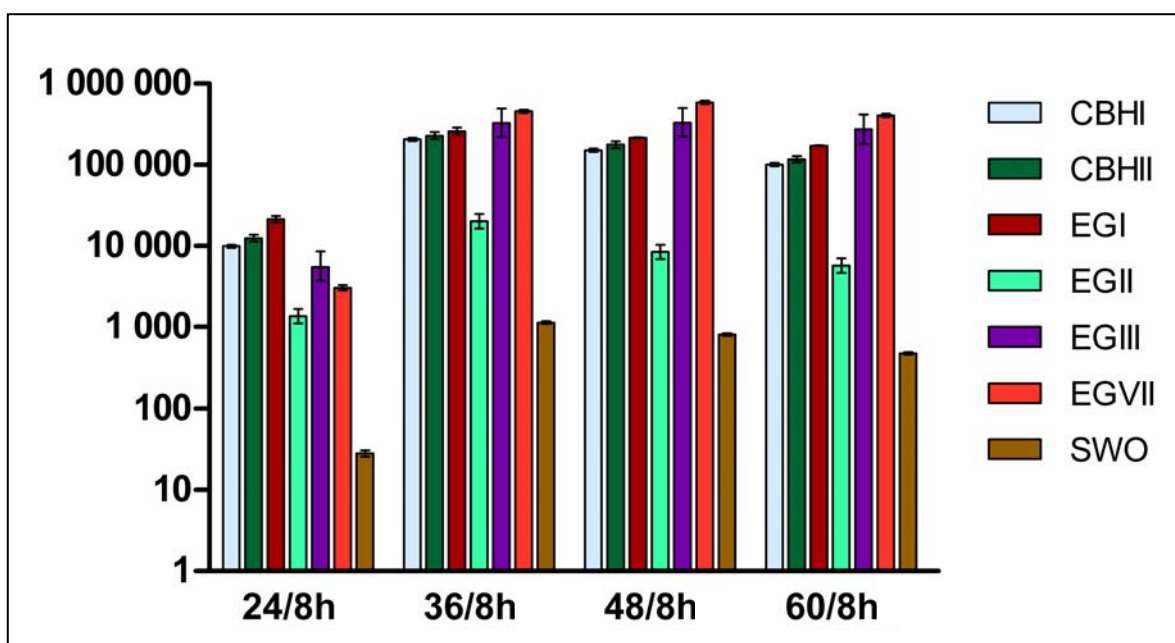


**Figura 5:** Valores de expressão de transcritos obtidos de cultura induzida de *Trichoderma harzianum* IOC-3844 por 8, 36 e 60 horas, em relação aos transcritos obtidos de cultura não induzida de 8, 36 e 60 horas. Os valores foram normalizados pela expressão do transcrito de actina. Os alvos calibradores foram os transcritos de cultura não induzida de 8, 36 e 60 horas. O eixo das ordenadas está em escala logarítmica.

A figura 5 também mostra valores massivos de superexpressão em 36 horas de indução com alguns dos valores acima de centenas de milhares de vezes mais expressos em relação aos transcritos correspondentes não induzidos, como calculado para *cbh1* (164.498,23), *cbh2* (147.746,79), *egl1* (112.798,46) e *egl7* (535.978,98). Todos os valores de expressão e as significâncias estatísticas correspondentes à figura 5 podem ser vistos no apêndice H.

Em 60 horas de indução, todos os valores de expressão mostraram diminuição relativa comparados com 36 horas de indução, mas ainda mantendo valores de superexpressão na casa dos milhares, destacando-se o transcrito de *egl7* que mostrou o valor de 232.243,76 vezes mais expresso em relação ao *egl7* do micélio de 60 horas não induzido (Figura 5). Vale destacar que os valores de superexpressão de *egl7* em 36 e 60 horas de indução, foram os mais altos encontrados. Finalmente, o valor de superexpressão de *egl1* também chama atenção por ser mais alto do que os valores de *cbh1* e *cbh2* em 60 horas de indução (Figura 5).

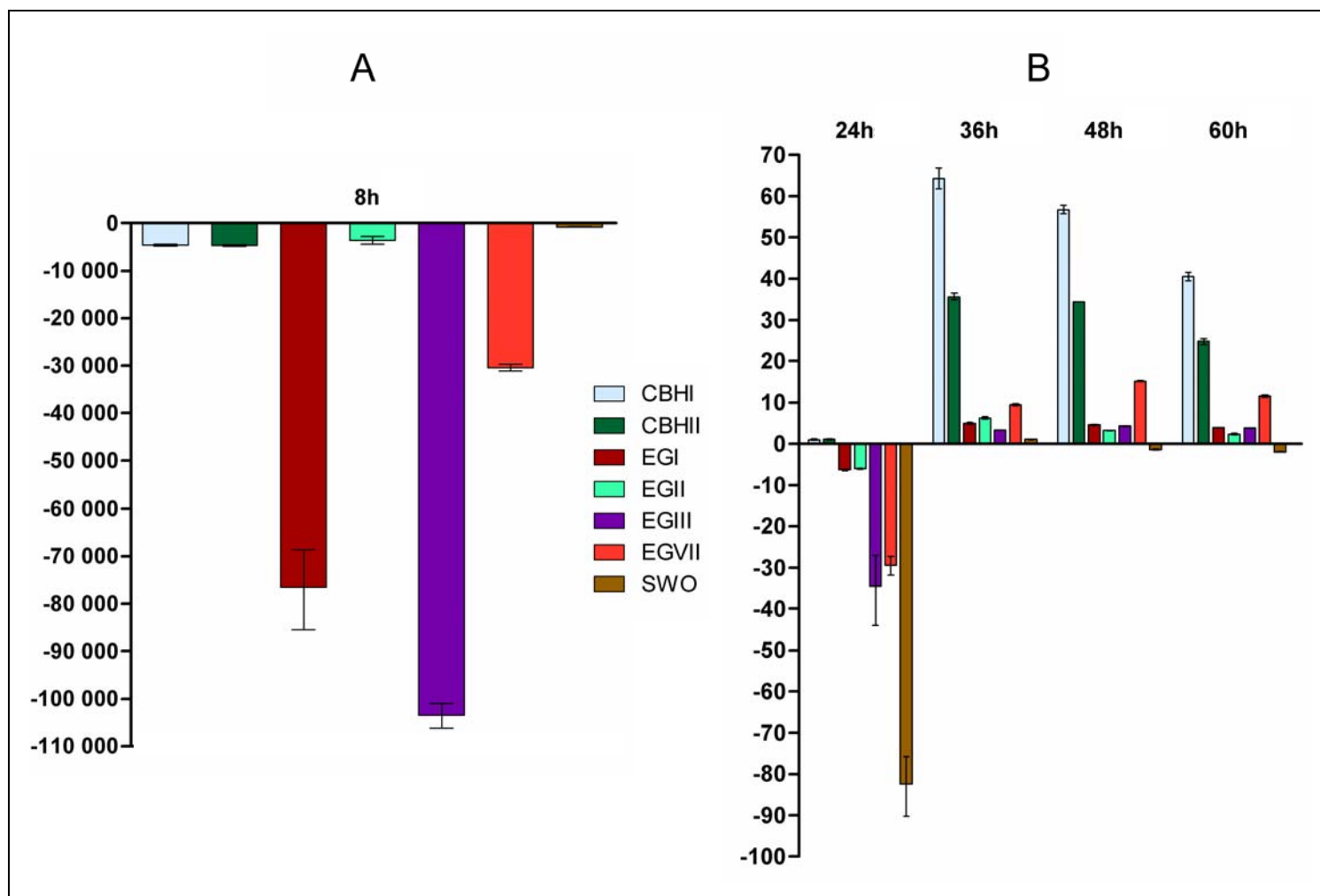
A comparação de expressão dos transcritos obtidos de 24, 36, 48, e 60 horas de indução, em relação aos transcritos de 8 horas de indução, mostrou um aumento geral de níveis de mRNA de 24/8 a 36/8 horas (Figura 6). Um leve decréscimo de 36/8 a 60/8 horas foi verificado para a maior parte dos transcritos, com exceção de *egl3* e *egl7* que mostraram maiores valores em 48/8 horas quando comparados aos valores de 36/8 horas (Figura 6). Todos os valores de expressão e as significâncias estatísticas correspondentes à figura 6 podem ser vistos no apêndice H.



**Figura 6:** Valores de expressão de transcritos obtidos de cultura induzida de *Trichoderma harzianum* IOC-3844 por 24, 36, 48 e 60 horas, comparados com os transcritos obtidos de cultura do fungo induzido por 8 horas. Os valores foram normalizados pela expressão do transcrito de actina. Os alvos calibradores foram os transcritos de cultura induzida por 8 horas. O eixo das ordenadas está em escala logarítmica.

A comparação dos valores de expressão dos transcritos obtidos de micélio induzido, em relação ao valor de expressão médio do transcrito de actina em todos os pontos induzidos (ver métodos), mostra as abundâncias relativas entre os transcritos em cada tempo de indução avaliado (Figura 7). Os transcritos *egl1*, *egl3* e *egl7* se mostraram excessivamente subexpressos em comparação com o nível de mRNA de actina (ou em relação aos demais transcritos), no tempo de 8 horas de indução (Figura 7 A). Isto está em concordância com os valores mostrados nas figuras 5 e 6, em que os

mesmos transcritos mostraram os mais altos incrementos nos níveis de expressão, em comparação com condições não indutoras (Figura 5) e com o tempo de indução de 8 horas (Figura 6). Além disso, os valores de expressão dos demais transcritos, ou seja de *cbh1*, *cbh2*, *egl2* e *swo 1* podem sugerir uma resposta mais rápida à indução por celulose (Figura 7 A). A figura 7 B confirma a tendência sugerida na figura 6, mostrando os maiores valores de expressão em 36 horas para todos os transcritos com uma ligeira diminuição de 36 para 48 e finalmente de 48 para 60 horas de indução. Vale também destacar que os valores de expressão dos transcritos *cbh1* e *cbh2*, especialmente o *cbh1*, mostraram maiores diminuições de 36 para 60 horas em comparação aos demais transcritos (Figura 7B). Todos os valores de expressão e as significâncias estatísticas correspondentes à figura 7 podem ser vistos no apêndice H.



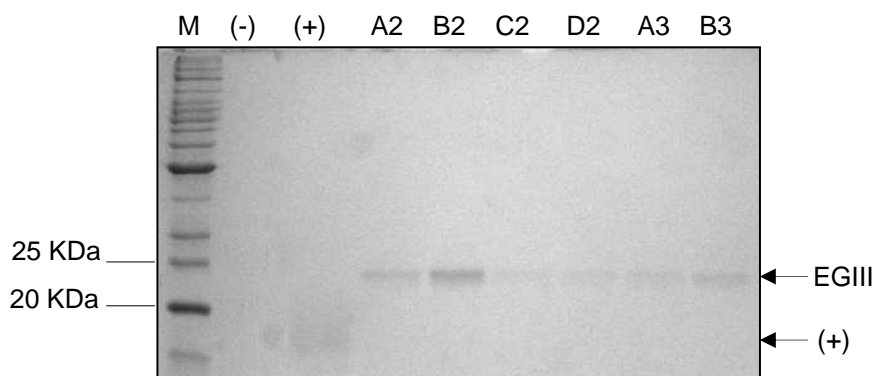
**Figura 7: Valores de expressão de transcritos obtidos de cultura induzida de *Trichoderma harzianum* IOC-3844 por 8, 24, 36, 48 e 60 horas, comparados com a média dos valores de expressão do transcrito de actina nos tempos de 8, 24, 36, 48 e 60 horas de cultura do fungo induzido. Os valores foram normalizados pela expressão do transcrito de actina. Os alvos calibradores foram a media do valor de expressão do transcrito de actina em todos os tempos de indução. O valor zero no eixo das abscissas indica o nível de expressão da actina de micélio induzido. A: Valores de expressão de 8 horas de indução mostrados separadamente por razões de escala. B: Valores de expressão para 24, 36, 48 e 60 horas de indução.**

#### 4.7 Expressão de proteínas recombinantes

Os ensaios de indução da enzima EGIII de *T. harzianum* IOC-3844 foram realizados em pH 6,5. O peso molecular esperado para a EGIII (sem peptídeo sinal) foi

calculado a partir da seqüência de DNA da ORF recombinante, em plasmídeo pPIC-Z-Alpha-A, com uso do programa “GENE RUNNER” versão 3.01. O valor encontrado foi de 23,8 KDa. O tempo de indução analisado foi de 72 horas e 15 µL dos sobrenadantes de cultura foram analisados por SDS-PAGE (Figura 8).

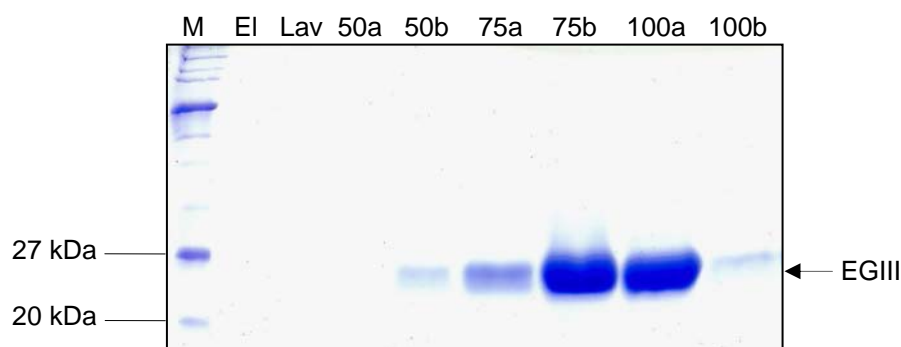
Foram ensaiadas 22 colônias de *P. pastoris* recombinantes para EGIII de *T. harzianum* IOC-3844 clonada em pPIC-Z-Alpha-A. As análises por SDS-PAGE mostraram que a maioria das colônias de *P. pastoris* secretou a proteína recombinante no sobrenadante de cultura. A figura 8 mostra 6 dos 22 clones ensaiados, entre eles se encontra o clone B2 que apresentou a banda mais intensa, tendo sido por isso escolhido para os ensaios subseqüentes de expressão.



**Figura 8: Sobrenadante de cultura de *Pichia pastoris* expressando EGIII de *Trichoderma harzianum* IOC-3844.** SDS-PAGE mostrando em M: Marcador de massa molecular “BenchMark” (Invitrogen). (-): Controle negativo de *P. pastoris* transformada com pPIC-Z-Alpha-A. (+): Controle positivo com *P. pastoris* transformada com proteína recombinante de 15 KDa clonada em pPIC-Z-Alpha-A. A2, B2, C2, D2, A3 e B3: 6 dos 22 clones de EGIII de *P. pastoris* ensaiados para expressão. As setas indicam as bandas de expressão da EGIII e da proteína usada como controle positivo.

A indução da enzima EGIII a partir do clone B2 de *Pichia pastoris* (figura 8) foi otimizada para o tempo de indução de 24 horas mantendo o pH 6,5. A figura 9 mostra 20 µL das principais frações de 5 mL do sobrenadante de cultura, obtidas após a passagem para purificação e eluição da coluna de níquel “Ni-NTA Superflow” (QIAGEN). A quantificação da proteína recombinante mostrou rendimento de 60 mg/L em meio de indução BMMY, o que corresponde a 300mg/L em meio de cultivo BMGY.

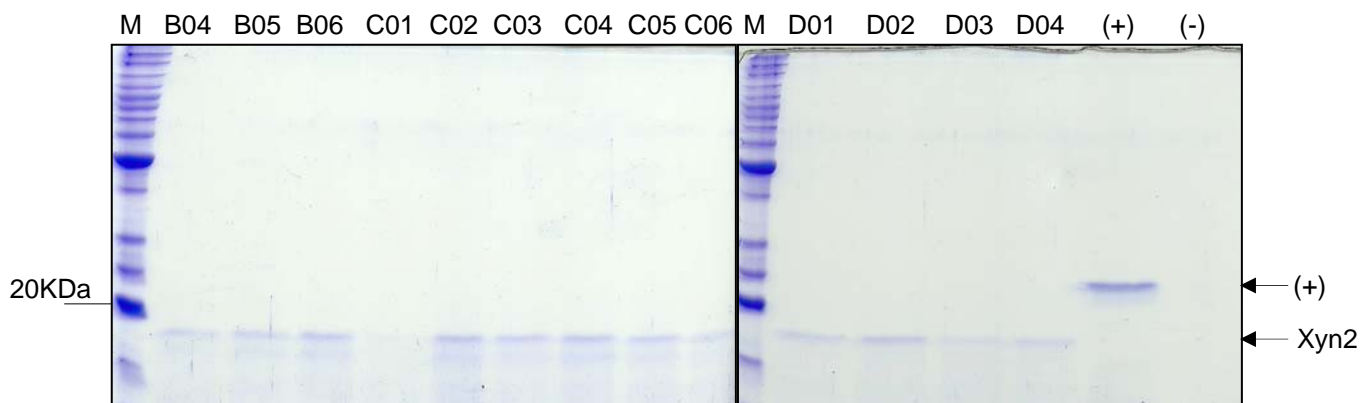




**Figura 9: Frações purificadas do sobrenadante do clone B2 de *Pichia pastoris* transformada com EGIII de *Trichoderma harzianum* IOC-3844.** SDS-PAGE mostrando em M: Marcador de massa molecular “Broad Range” (Biolabs). El: 20  $\mu$ L do eluato do sobrenadante de cultura (meio BMMY). Lav: 20  $\mu$ L da lavagem da coluna de níquel com tampão de lise sem imidazol. 50a a 100b: frações de 5 mL contendo concentrações crescentes de imidazol (expressas em mM) em tampão de lise. A seta indica a banda de expressão da proteína EGIII.

Os ensaios de indução da enzima Xyn2 de *T. harzianum* IOC-3844 foram realizados em pH 6,0. O peso molecular esperado para a Xyn2 foi calculado a partir da seqüência de DNA da ORF recombinante com uso do programa “GENE RUNNER” versão 3.01. O valor encontrado foi de 22 KDa.

O tempo de indução analisado foi de 72 horas e 15  $\mu$ L dos sobrenadantes de cultura foram analisados por SDS-PAGE (Figura 10). Foram ensaiadas 22 colônias de *P. pastoris* recombinantes para Xyn2 de *T. harzianum* IOC-3844 clonada em pPIC-Z-Alpha-A. As análises por SDS-PAGE mostraram que a maioria das colônias de *P. pastoris* secretou a proteína recombinante no sobrenadante de cultura. A figura 10 mostra 13 dos 22 clones ensaiados.

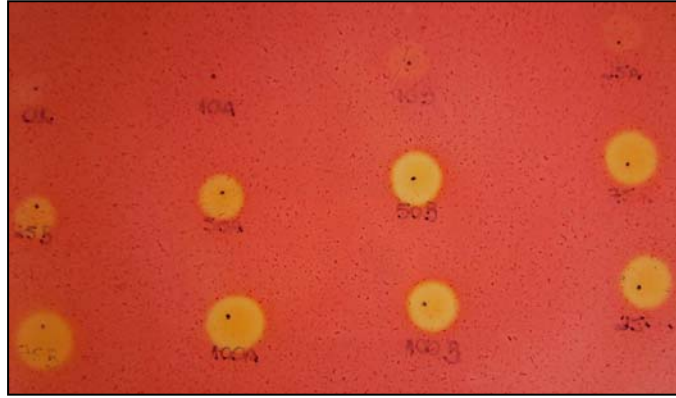


**Figura 10: Sobrenadante de cultura de *Pichia pastoris* expressando Xyn2 de *Trichoderma harzianum* IOC-3844.** SDS-PAGE mostrando em M: Marcador de massa molecular “BenchMark” (Invitrogen). B04 a D04: 13 dos 22 clones ensaiados. (+): Controle positivo com *P. pastoris* transformada com proteína recombinante de 23,8 KDa clonada em pPIC-Z-Alpha-A. (-): Controle negativo de *P. pastoris* transformada com pPIC-Z-Alpha-A. As setas indicam as bandas de expressão da Xyn2 e da proteína usada como controle positivo.

Ensaaios de expressão em *Pichia pastoris* de outras enzimas de *Trichoderma harzianum* IOC-3844 foram ou estão sendo realizados por outros estudantes do Laboratório de Biologia Molecular do DGE-UFSCar e os resultados estão resumidos no apêndice F.

#### 4.8 Ensaaios de atividade das enzimas recombinantes

As frações purificadas de EGIII de *T. harzianum* IOC-3844 (Figura 9) sugeriram forte atividade desta enzima em placas contendo CMC e ensaiadas com Vermelho Congo (Figura 11).



**Figura 11:** Teste de atividade das frações purificadas do sobrenadante da indução do clone B2 de *Pichia pastoris* transformada com EGIII de *Trichoderma harzianum* IOC-3844. Ensaio realizado em placas de CMC reveladas com Vermelho Congo.

Os sobrenadantes de cultura de *P. pastoris* contendo Xyn2 de *T. harzianum* IOC-3844 sugeriram alta atividade xilanásica (Figura 12).



**Figura 12:** Teste de atividade dos sobrenadantes de culturas induzidas de *Pichia pastoris* transformadas com Xyn2 de *Trichoderma harzianum* IOC-3844, em placas de Xilano reveladas com Vermelho Congo. A seta indica o controle negativo de *P. pastoris* transformada com pPIC-Z-Alpha-A.

Ensaio de atividade de outras enzimas de *Trichoderma harzianum* IOC-3844 produzidas em *Pichia pastoris* foram ou estão sendo realizados por outros estudantes do Laboratório de Biologia Molecular do DGE-UFSCar e os resultados estão resumidos no apêndice F.

## 5 DISCUSSÃO

### 5.1 Análise inicial da expressão de celulases por RT-PCR

As análises de RT-PCR foram realizadas para verificar o micélio quanto à indução da expressão de celulases e validar seu uso na construção da biblioteca induzida por celulose. A expressão aumentada dos transcritos de *cbh1* e *bgl2* em micélio induzido por 60 horas, em comparação ao micélio não induzido, pode indicar a indução simultânea das outras celulases e das hemicelulases (SADDLER, HOGAN, LOUISSEIZE, 1985; MARGOLLES-CLARK, ILMÉN, PENTTILA, 1997). Em relação ao tempo de indução, estudos prévios em *T. reesei* (ABRAHÃO-NETO et al., 1995) reportaram alta expressão de celulases entre 16 e 60 horas de indução, com atividade celulolítica crescente em até 80 horas. Neste contexto, o RNA de 60 horas de indução foi validado para ser usado na construção da biblioteca induzida de *Trichoderma harzianum* IOC-3844.

### 5.2 Contagem de ESTs

A biblioteca de cDNA de *T. reesei* QM6a, disponível publicamente, foi obtida a partir de micélios cultivados em diferentes condições indutoras, diferentes fontes de celulose e variados pontos de coleta. As culturas líquidas foram induzidas: com 2 % de Avicel, 2 % de “Solka-Floc” e 2 % de farelo de trigo, com os micélios coletados depois de 3 e 6 dias e com 2 % de polpa de beterraba com o micélio coletado depois de 6 dias. As culturas sólidas foram induzidas: com farelo de trigo e “Solka-Floc”, com os micélios coletados depois de 6 e 7 dias e com polpa de beterraba e “Solka-Floc”, com os micélios coletados depois de 9 dias (FOREMAN et al., 2003; DIENER et al., 2004).

Estes protocolos de indução representam uma abordagem experimental distinta da usada no presente trabalho, no qual a construção da biblioteca de cDNA do *T. harzianum* IOC-3844 foi feita a partir de micélio coletado depois de 60 horas de indução em meio líquido com 1 % de Avicel. Estas diferenças podem ser mais significativas quanto ao tempo de coleta, uma vez que todas as induções, para ambos os fungos, foram realizadas na presença de celulose. No entanto, há indicações de que o *T. reesei* RutC30 leva mais tempo do que o *T. harzianum* IOC-3844 para atingir os níveis

máximos de atividades de celulolíticas (CASTRO et al., 2010a; CASTRO et al., 2010b), o que minimizaria as diferenças de tempos de coleta entre os dois ensaios citados.

Neste contexto se justifica a comparação, baseada em contagem de ESTs, dos níveis de expressão dos transcritos listados na tabela 3. Interessantemente, as diferenças de expressão mostradas na tabela 3 sugerem que o *T. harzianum* IOC-3844 é melhor produtor de enzimas relacionadas com a desconstrução de biomassa em comparação ao *T. reesei* QM6a. Isso pode ser visualizado com base nas razões calculadas para os transcritos e para as categorias de transcritos (Tabela 3). Mais adiante serão discutidas as comparações dos níveis de expressão dos transcritos de interesse para as duas espécies de fungos, com base na contagem de ESTs e nos experimentos de PCR em tempo real. De qualquer maneira, os dados mostrados nas tabelas 2 e 3 sugerem que o *T. harzianum* IOC-3844, induzido em cultura líquida por 60 horas, é capaz de produzir suas principais celulases em grande quantidade.

As mais altas representatividades dos transcritos *cbh1* e *cbh2* de *T. harzianum* IOC-3844 (Tabela 2), em comparação com suas demais celulases, também tem sido reportadas para *T. reesei* (ILMÉN et al., 1997; LYND et al., 2002). Os dados mostrados na tabela 3 também confirmam níveis mais elevados de *cbh1* e *cbh2* para *T. reesei* QM6a em relação as suas demais celulases. No entanto, os transcritos, *cbh1* e *cbh2* de *T. harzianum* IOC-3844 parecem ter níveis mais equilibrados de expressão entre si, em comparação com o *T. reesei* QM6a, no qual o transcrito de *cbh1* é muito mais representado do que o *cbh2* (Tabela 3). De fato, (ILMÉN et al., 1997) mostraram expressão bem mais elevada de *cbh1* em comparação a *cbh2* e as demais celulases, em *T. reesei* QM9414 cultivado sob uma variedade de condições indutoras com celulose. Mais considerações a respeito da expressão de *cbh1* e *cbh2* de *T. harzianum* IOC-3844 serão discutidas mais adiante.

Entre as endoglucanases de *T. harzianum* IOC-3844, os transcritos de *egl7* e *egl3* mostraram contagem mais elevada (Tabela 2). Por outro lado, os transcritos de endoglucanases mais abundantes para *T. reesei* QM6a foram o *egl7* e o *egl6* (Tabela 3).

Análises estruturais reportadas por (KARKEHABADI et al., 2008) revelaram que a proteína EGVII de *T. reesei* possui similaridade com a proteína não catalítica ligante de quitina CBP21 de *Serratia marcescens*. Interessantemente, apesar de a EGVII não possuir domínio típico de ligação a carboidrato (CBM), as similaridades estruturais com CBP21 incluem a região CBM desta enzima e levaram os autores a sugerir função similar de ligação a carboidrato na EGVII (KARKEHABADI et al., 2008). Estes

autores reportaram também baixa atividade endoglucanásica da EGVII de *T. reesei* e além disso, as análises estruturais não permitiram a identificação de um centro catalítico nesta enzima. Com isso, levando em consideração que a CBP21 de *Serratia marcescens* atua indiretamente aumentando a degradação das quitinases (VAAJE-KOLSTAD et al., 2005), é possível sugerir uma função sinérgica ou acessória para a EGVII (KARKEHABADI et al., 2008). Corroborando com estas observações (MERINO & CHERRY, 2007) reportaram que em substratos nativos com hemicelulose e lignina (cellulose não purificada) a EGVII pode contribuir para a sinergia entre as celulases aumentando a atividade celulolítica. Desta forma, o maior nível de expressão de EGVII, sugerido pela contagem de ESTs para os dois fungos analisados (Tabela 3), pode indicar importante função desta proteína como enzima acessória. Isto pode ser particularmente importante em *T. harzianum* IOC-3844 no qual o transcrito *egl7* apresenta representatividade de 11,82 vezes maior em relação ao *egl7* de *T. reesei* QM6a (Tabela 3). Mais adiante há comentários adicionais em relação aos níveis de expressão do transcrito *egl7* de *T. harzianum* IOC-3844.

O transcrito *egl6* pode ser importante para o sistema celulolítico de *T. reesei* QM6a, apresentando uma das maiores contagens de ESTs entre as endoglucanases deste fungo, mas pode ser desnecessário para o sistema celulolítico de *T. harzianum* IOC-3844, já que não foi detectado neste fungo (Tabela 3). No entanto, mais esforços de sequenciamento poderiam revelar uma eventual presença deste transcrito em *T. harzianum* IOC-3844 induzido por Avicel.

Por outro lado, o transcrito *egl3* parece ser mais importante para o *T. harzianum* IOC-3844, considerando que ele é bastante representado em relação aos demais transcritos para endoglucanases deste fungo, e uma vez que ele não aparece na biblioteca analisada de *T. reesei* QM6a (Tabela 3). Corroborando com estas informações (ULKER & SPREY, 1990) reportaram que em *T. reesei* a representatividade da EGIII é menor do que 1 % do total de celulases expressas neste fungo. Interessantemente, esta enzima não apresenta domínio típico CBM (FOREMAN et al., 2003; SANDGREN, STAHLBERG, MITCHINSON, 2005) o que proporciona atividade preferencial em celulose amorfa ou solúvel, como descrito para *T. reesei* (SPREY & BOCHEM, 1992; KARLSSON et al., 2002). Esta atividade preferencial em celulose solúvel também foi reportada para EGIII de *T. harzianum* IOC-3844, em uma publicação recente, fruto do trabalho desenvolvido no Laboratório de Biologia Molecular do DGE-UFSCar em colaboração com o laboratório de Cristalografia do

Instituto de Física da USP de São Carlos (IFSC) (LIBERATO et al., 2012). Este trabalho envolve ensaios com a EGIII recombinante de *T. harzianum* IOC-3844 (ver mais adiante o item “Expressão e ensaios de atividade das proteínas recombinantes”). Vale destacar também que a EGIII apresenta atividade tipo expansina desorganizando as fibras de celulose e relaxando a estrutura da parede celular, e com isso podendo deixar a celulose mais acessível a outras celulasas (HENRIKSSON et al., 1999; YUAN, WU, COSGROVE., 2001). Desta forma, a EGIII, sendo mais representada em *T. harzianum* IOC-3844 (Tabela 3; ULKER & SPREY, 1990) pode ter importante papel neste fungo, por um lado proporcionando uma possível maior eficiência em abordar celulose amorfa, em comparação ao *T. reesei*, e por outro lado atuando de modo sinérgico, desorganizando a celulose e promovendo aumento de acessibilidade de outras celulasas ao substrato. Com isso, apesar de a EGIII apresentar menor eficiência hidrolítica, em substratos celulósicos, em comparação a EGI e EGII em *T. reesei* (KARLSSON et al., 2002), estas observações podem ajudar a explicar as altas atividades de endoglucanase reportadas em *T. harzianum* IOC-3844 por (CASTRO et al., 2010a). Neste contexto, os altos níveis de expressão de *egl3* e *egl7* entre as endoglucanases de *T. harzianum* IOC-3844, indicados pela contagem de ESTs (Tabela 3), podem contribuir significativamente para o sistema celulolítico deste fungo.

O transcrito *egl8* também chama a atenção pela provável maior importância para a atividade celulolítica de *T. reesei* QM6a, já que a contagem de ESTs o indica como um dos transcritos de endoglucases mais expressos neste fungo e, uma vez que ele não aparece entre os ESTs de *T. harzianum* IOC-3844 (Tabela 3). A ausência dos transcritos de *egl6* e *egl8* na biblioteca de *T. harzianum* IOC-3844 pode sugerir uma compensação por suas outras endoglucanases bastante expressas, incluindo aquelas que aparecem apenas na sua biblioteca e também aquelas que aparecem para os dois fungos. Além disso, todas estas endoglucanases em conjunto apresentam maiores contagens de ESTs em *T. harzianum* IOC-3844 em comparação ao *T. reesei* QM6a (Tabela 3).

Os transcritos *egl1*, *egl2* e *egl4* possuem abundância similar dentro do transcriptoma de cada um dos fungos mas apresentam contagem de ESTs de 3 a 6 vezes maiores em *T. harzianum* IOC-3844 (Tabela 3). Interessantemente, apesar de ter domínio típico de CBM e sítio catalítico (SALOHEIMO et al., 1997), a proteína EGIV de *T. reesei* apresenta atividade endoglucanásica significativamente diminuída em comparação a EGI (KARKEHABADI et al., 2008; KARLSSON et al., 2001). No entanto, diferentemente da EGVII, não há evidência na literatura sobre uma eventual

importância da EGIV como enzima acessória. Finalmente, o transcrito *egl5* pode possuir maior importância na atividade endoglucanásica de *T. harzianum* IOC-3844, visto que este transcrito não aparece na biblioteca de *T. reesei* QM6a, que é 14,78 vezes maior (Tabela 2). Em conjunto, as representatividades mais altas dos transcritos de endoglucanases de *T. harzianum* IOC-3844 (Tabela 3) corroboram com as maiores atividades endoglucanásicas deste fungo, em relação ao *T. reesei* Rut-C30, reportadas por (CASTRO et al., 2010a). Mais adiante há comentários adicionais sobre a expressão dos transcritos *egl1*, *egl2*, *egl3* e *egl7* de *T. harzianum* IOC-3844.

O transcrito de *bgl1* pode ser bastante importante para a atividade  $\beta$ -glucanásica de *T. harzianum* IOC-3844 visto que, comparado ao *T. reesei* QM6a, este transcrito foi 64.04 vezes mais abundante (Tabela 3), e considerando que foi o transcrito que apresentou maior contagem entre as  $\beta$ -glucosidases de *T. harzianum* IOC-3844 (Tabela 2). Os transcritos *bgl2* e *bgl5* também se apresentaram bem mais abundantes em *T. harzianum* IOC-3844 com contagens de 17,73 e 19,71 vezes maiores, respectivamente (Tabela 3). Por outro lado, os transcritos *bgl3*, *bgl6* e *bgl7* apareceram apenas na biblioteca de *T. reesei* QM6a (Tabela 3). Para saber se estes transcritos ocorrem em *T. harzianum* induzido por Avicel, mais esforços de seqüenciamento são necessários. No entanto, apesar disto, os transcritos de  $\beta$ -glicosidases *bgl1*, *bgl2* e *bgl5* de *T. harzianum* IOC-3844 são bem mais representados neste fungo em relação ao *T. reesei* QM6a (Tabela 3), e isto corrobora com os resultados de (CASTRO et al., 2010a) que reportaram níveis significativamente elevados de atividade de  $\beta$ -glucosidásica de *T. harzianum* IOC-3844.

O transcrito de *xyn2* apresentou maiores contagens de ESTs entre as hemicelulases de *T. harzianum* IOC-3844 e também de *T. reesei* QM6a. No entanto, o *xyn2* mostrou representatividade bastante aumentada em *T. harzianum* IOC-3844, comparado ao *T. reesei* QM6a, sendo 49,88 vezes mais abundante (Tabela 3). Os transcritos *xyn3* e *man1* de *T. harzianum* IOC-3844 também apresentaram abundância maior em relação ao *T. reesei* QM6a, com contagens de 17,24 e 29,56 vezes mais representados respectivamente (Tabela 3). Por outro lado, o transcrito de *bxll* mostra equilíbrio do nível de expressão entre os dois fungos (Tabela 3). Os demais transcritos de hemicelulases *axe1*, *axe2* e *abf2* não apareceram em *T. harzianum* IOC-3844 (Tabela 3), no entanto mais esforços são necessários para saber se estes transcritos ocorrem neste fungo induzido por celuose. De qualquer maneira mesmo com a ausência destes transcritos, os transcritos de hemicelulases de *T. harzianum* IOC-3844 em conjunto se



mostraram 17,65 vezes mais expressos do que os transcritos de hemicelulases de *T. reesei* QM6a (Tabela 3). As observações acima sugerem o *T. harzianum* IOC-3844 como fonte potencial de xilanase e mananase para aplicações industriais. Estes resultados corroboram com o trabalho de (SADDLER, HOGAN, LOUIS-SEIZE, 1985) que apesar de mostrar que o *T. harzianum* E58 produziu atividades xilanásicas similares em comparação ao *T. reesei* Rut-C30, quando estes fungos foram crescidos com xilano, mostrou também que este fungo produziu atividades xilanásicas bastante aumentadas, também em relação ao *T. reesei* Rut-C30, quando estes fungos cresceram com celulose “Solka-Floc”, ou com substrato nativo (biomass vegetal submetida a tratamento de “Steam explosion”). Além disso, (MAEDA et al., 2011) confirmaram importante atividade xilanásica induzida por celulose microcristalina (Avicel) para o *T. harzianum* IOC-3844.

A expressão de *xyn2* em *T. reesei* QM6a (Tabela 3) está em concordância com o trabalho de (ZEILINGER et al., 1996) que reporta, para *T. reesei* QM9414, a expressão deste transcrito induzido por xilano ou por celulose. Além disso, a detecção do transcrito de *xyn3* em *T. reesei* QM6a (Tabela 3) está em concordância com o trabalho de (XU et al., 2000) que reportam que esta enzima, em *T. reesei* PC- 3-7, é induzida por Avicel mas não é induzida por xilano e seus derivados. Já a ausência de *xyn1* em *T. reesei* QM6a não era esperada (Tabela 3), considerando que (ZEILINGER et al., 1996) reportaram a expressão de *xyn1* na presença de xilano e que (TENKANEN, PULS, POUTANEN, 1992) reportaram a expressão deste transcrito na presença de substrato nativo contendo xilano. Isto por que a biblioteca de *T. reesei* QM6a analisada aqui foi obtida de cultivos com vários tipos de substratos, incluindo substratos nativos contendo xilano (DIENER et al., 2004). No entanto, não temos informação de qual proporção de RNA, extraído de micélio induzido com xilano, foi usada na construção desta biblioteca de *T. reesei* QM6a. Com isso, uma eventual pouca representatividade deste RNAs poderia explicar a ausência do transcrito de *xyn1* na biblioteca analisada de *T. reesei* QM6a (Tabela 3). Por outro lado, a indução de expressão das xilanases em *T. reesei* parece ser controversa. O trabalho de (MARGOLLES-CLARK, ILMÉN, PENTTILA, 1997) reportou expressão equivalente de *xyn1* e *xyn2*, em *T. reesei* QM9414 induzido por celulose “Solka-Floc”, no entanto, (ZEILINGER et al., 1996) não detectaram a expressão de *xyn1* nesta mesma cepa induzida por celulose purificada. Além disso, (XU et al., 2000) detectaram a expressão de *xyn3* em *T. reesei* PC- 3-7 induzido por Avicel

mas não detectaram a expressão deste transcrito em *T. reesei* QM9414 na presença de substratos de celulose ou de xilano.

As expressões de *xyn2* e de *xyn3* de *T. harzianum* IOC-3844 induzidas por celulose (Tabela 2), estão de acordo com evidências que reportaram a expressão de três xilanases de massa de 20, 22 e 29 KDa, presentes no sobrenadante de cultura de *T. harzianum* E58 induzido por celulose “Solka-Floc” (WONG et al., 1986; TAN et al., 1985). A proteína de 22 KDa encontrada por estes autores provavelmente é a *xyn2* encontrada aqui, uma vez que a massa predita da proteína madura Xyn2 de *T. harzianum* IOC-3844 foi de 22 KDa, e que a Xyn2 purificada do sobrenadante de cultura do *T. harzianum* C4 apresentou massa de 22,5 KDa (LEE et al., 2009). Vale destacar que com o alinhamento das seqüências de aminoácidos, a Xyn2 de *T. harzianum* IOC-3844 mostrou 100 % de identidade em relação à seqüência da Xyn2 de *T. harzianum* C4 publicamente disponível (dados não mostrados) [“GenBank”ACF40831] (BENSON et al., 2011).

A proteína de 29 KDa encontrada em *T. harzianum* E58 por (WONG et al., 1986) possivelmente é a Xyn3 descrita aqui para *T. harzianum* IOC-3844 (Tabela 2), apesar da massa predita a partir da seqüência madura de aminoácidos desta última ter sido de 36,6 KDa. Esta observação é válida uma vez que ocorre situação similar em *T. reesei*: para a Xyn3 de *T. reesei* PC-3-7 a massa predita a partir da seqüência madura de aminoácidos é de 36,45 KDa, e no entanto, a massa da enzima isolada do sobrenadante de cultura é de 32 KDa (OGASAWARA et al., 2006).

A proteína de 20 KDa encontrada por (WONG et al 1986; TAN et al., 1985) em sobrenadante de *T. harzianum* E58 induzido por celulose “Solka-Floc”, provavelmente é uma homóloga da *xyn1* de 19 KDa encontrada em sobrenadante de *T. reesei* (TENKANEN, PULS, POUTANEN, 1992). A não detecção do transcrito de *xyn1* em *T. harzianum* IOC-3844 (Tabela 2) estaria em desacordo com a expressão da xilanase de 20 KDa encontrada por (WONG et al 1986; TAN et al., 1985) em *T. harzianum* E58, e com a expressão da *xyn1* de *T. reesei* QM9414 induzido por celulose “Solka-Floc” reportada por (MARGOLLES-CLARK, ILMÉN, PENTTILA, 1997). Por outro lado, uma possível ausência de *xyn1* em *T. harzianum* IOC-3844 induzido com celulose purificada, estaria em concordância com os resultados de (ZEILINGER et al., 1996), que não detectaram este transcrito em *T. reesei* QM9414 induzido com celulose purificada. Um maior esforço de seqüenciamento poderia ajudar a revelar eventual presença de *xyn1* em *T. harzianum* IOC-3844 induzido por celulose.

O transcrito da enzima acessória *swol*, relacionada à expansina, mostra abundância equilibrada em *T. harzianum* IOC-3844 e *T. reesei* QM6a (Tabela 3). Adiante há mais comentários sobre a expressão de *swol* em *T. harzianum* IOC-3844.

Entre as prováveis enzimas glicolíticas Cip1 e Cip2 previamente descritas para *T. reesei*, e com potencial papel na degradação de biomassa vegetal (FOREMAN et al., 2003), o transcrito de *cip2* pode ser mais importante para *T. harzianum* IOC-3844 por apresentar maior abundância de ESTs (7,39 vezes mais representado) em relação ao *T. reesei* QM6a. Por outro lado, o transcrito *cip1* apresenta contagem de ESTs equilibrada entre os dois fungos (Tabela 3). Outrossim, os dados da tabela 3 indicam que os transcritos *cip1* e *cip2* de *T. harzianum* IOC-3844 são co-regulados com as celulases, assim como reportado para *T. reesei*, por (FOREMAN et al., 2003).

Analisando os fatores de transcrição listados na tabela 3, o transcrito para o repressor de expressão de celulases e xilanases *ace1* (ARO et al., 2003) não apresenta diferenças de abundância entre os dois fungos induzidos. Por outro lado, o transcrito do ativador de celulases e hemicelulases *xyl1* mostra maior abundância em *T. harzianum* IOC-3844 com 9,85 vezes mais ESTs do que em *T. reesei* QM6a. Já o ativador de transcrição *ace2* não foi detectado na biblioteca de *T. harzianum* IOC-3844, o que inviabiliza a comparação do número de ESTs deste transcrito entre os fungos.

A importância do *ace2* foi reportada por (ARO et al., 2001) que mostraram que durante o crescimento em celulose “Solka-Floc”, a deleção do *ace2* em *T. reesei* levou a uma redução de 30 % a 70 % na atividade de celulases e diminuiu a expressão de *xyn2*. No entanto, uma vez que o *T. harzianum* IOC-3844 mostrou maior abundância de transcritos de enzimas envolvidas em atividades celulolíticas do que em *T. reesei* QM6a (Tabela 3 e discussões mais adiante) e apresentou também representatividades altas de transcritos para *Xyn2* (Tabelas 2 e 3), é possível especular que o *T. harzianum* IOC-3844 pode apresentar algum nível de expressão de ACE2 não detectado devido ao tamanho reduzido da biblioteca de cDNA analisada (1544 clones analisados).

Por outro lado, a maior representatividade do transcrito de *Xyl1* por si, em *T. harzianum* IOC-3844, comparado ao *T. reesei* QM6a (Tabela 3), poderia superar a repressão por ACE1 e até uma eventual ausência de ACE2, contribuindo para a alta produtividade de celulases e hemicelulases do primeiro fungo, em especial as endoglucanases, como reportado por (CASTRO et al., 2010a) e discutido neste trabalho. Estas observações são sustentadas pelos resultados de (ARO et al., 2003) que reportaram que uma linhagem de *T. reesei* contendo os genes *ace1* e *ace2* deletados

expressou celulasas e xilanases em níveis similares aos de uma linhagem de *T. reesei* contendo apenas a deleção do gene *aceI*. De acordo com (KUBICEK et al., 2009) estes resultados podem ser explicados pela atividade permanente de Xyr1 na célula.

Por último, o fator de transcrição *creI* (repressor de catabolismo de carbono) não foi detectado na biblioteca de *T. harzianum* IOC-3844 (Tabela 3). No entanto, isto não significa que este transcrito não ocorra ou seja pouco representado neste fungo em relação ao *T. reesei* QM6a. Além disso, este gene já foi descrito para o *T. harzianum* T3 e foi o primeiro regulador a ter função associada ao controle de expressão de celulasas em fungo filamentosos (ILMÉN, THRANE, PENTTILA, 1996). O *creI* é um importante regulador de assimilação de carbono (PORTNOY et al., 2011) e reprime a expressão de celulasas e hemicelulasas em *T. reesei* na presença de glicose (ILMÉN et al., 1997; NAKARI-SETALA et al 2009). Sua função foi primeiramente descrita por (ILMÉN, THRANE, PENTTILA, 1996) que mostraram que a linhagem mutante de *T. reesei* Rut-C30 expressa uma forma truncada do *creI* e é capaz de produzir celulasas em meio contendo glicose. Além disso, (NAKARI-SETALA et al 2009) comprovaram que a forma truncada de *creI* em *T. reesei* é uma mutação nula que por si só inviabiliza a repressão da expressão de celulasas e hemicelulasas na presença de glicose. Outrossim, estes autores mostraram aumento de atividade celulásica e xilanásica, em condições indutoras, em *T. reesei* QM6a transformado com o *creI* truncado ou com o *creI* anulado, quando comparado ao mesmo fungo contendo o gene selvagem. Isto torna a obstrução do *creI* uma importante estratégia para aumentar a produção de celulasas e hemicelulasas em *Trichoderma* (NAKARI-SETALA et al 2009). Estas observações nos encorajam a planejar ensaios futuros para, uma vez identificada a ORF de *creI* em *Trichoderma harzianum* IOC-3844, investigar as atividades celulásicas e xilanásicas deste fungo com o *creI* obstruído .

### **5.3 Busca de seqüências de *Trichoderma harzianum* relacionadas com degradação de biomassa vegetal**

Os ESTs envolvidos com degradação de biomassa vegetal, listados na tabela 4, apresentam baixas abundâncias relativas em comparação ao número total de transcritos das bibliotecas que os originaram. Com exceção dos ESTs dos fatores de transcrição ACEI, Xyr1 e CRE1, os ESTs das enzimas listadas na tabela 4, podem ter sido expressos por que os micélios usados para a obtenção dos cDNA das bibliotecas de

origem, apesar de não terem sido tipicamente induzidos para celulases ou hemicelulases, foram crescidos na presença de parede celular (informações obtidas a partir das descrições dos depósitos dos ESTs). O repressor de celulase e xilanase ACEI (Tabela 4) normalmente aparece em variadas condições de crescimento uma vez que ele também pode regular genes não relacionados com a desconstrução da biomassa (ARO et al., 2003). O aparecimento de Xyr1 (Tabela 4) pode ser explicado pela sua disponibilidade permanente da célula, tendo sua transcrição aparentemente não induzida por celulose (MACH-AIGNER et al., 2008). Já o CRE1 é normalmente expresso em fontes de carbono indutoras e não indutoras (ILMÉN, THRANE, PENTTILA, 1996).

Neste contexto, os resultados destas buscas indicam que estamos reportando neste trabalho, o primeiro estudo que descreve e analisa o conjunto de transcritos relacionados com a desconstrução da biomassa vegetal, obtidos de micélio de *T. harzianum* induzido por celulose microcristalina (Avicel).

#### **5.4 PCR em tempo real e análises de expressão**

A figura 5 mostra valores de expressão diferencial bastante elevados para os transcritos de micélio de *T. harzianum* IOC-3844 induzido com celulose em comparação aos transcritos de micélio não induzido (crescido em meio contendo glicerol como fonte de carbono). Estes valores são muito maiores do que os valores já reportados para *T. reesei*. Com base em experimentos de “Northern blot” realizados em até 24 horas de indução com celulose microcristalina (Avicel), foi proposto por (CARLE-URIESTE et al., 1997) aumentos de pelo menos 1.100 vezes na transcrição de *cbh1* e *egl1* de *T. reesei* QM9414 induzido, em comparação ao fungo crescido em meio contendo glicerol como fonte de carbono (não induzido). Também foi sugerido por (ILMÉN et al., 1997) que o transcrito de *cbh1* de *T. reesei* QM9414 seria pelo menos 6.000 vezes mais abundante em 20 horas de cultivo induzido com celulose “Solka-Floc”, quando comparado com condições de cultivo não induzido, com glicose como fonte de carbono. Além disso, estes autores também sugeriram que os transcritos de *cbh2* e *egl1* de *T. reesei* QM9414 são pelo menos 1.000 vezes mais abundantes em condições indutoras (com celulose) em comparação a condições não indutoras.

Mais recentemente (VERBEKE et al., 2009) analisaram a linhagem hiper celulolítica de *T. reesei* CL847 via PCR em tempo real para avaliar os níveis de expressão dos seguintes transcritos de celulasas e genes relacionados com expansinas: *cbh1*, *egl1*, *egl4*, *egl8*, *bgl1*, *bgl2*, *eel2* e *swol*. Estes autores reportaram os valores de expressão dos transcritos alvos em micélio induzido de 2 a 48 horas com soforose, celulose “Solka-Floc” e outros indutores, comparando com micélio não induzido crescido com glicerol como fonte de carbono. Com isto, interessantemente, os autores reportaram todos os seus transcritos induzidos, e superexpressos em relação aos não induzidos, com valores abaixo de 1.000 vezes mais expressos. Estes valores estão bem abaixo dos valores encontrados no nosso trabalho (Figura 5). Os valores de superexpressão do transcrito de *cbh1* de *T. harzianum* IOC-3844, obtido a partir do micélio induzido por 36 e 60 horas, em comparação ao micélio não induzido, foram de 164.498,23 e 15.929,41 vezes mais expressos respectivamente (Figura 5). Por outro lado (VERBEKE et al., 2009) mostraram valores de superexpressão de cerca de 200 e 60 vezes mais expresso para *cbh1* de *T. reesei* CL847 em ensaio semelhante com tempos de indução de 24 e 48 horas respectivamente. Para o transcrito de *egl1* de *T. harzianum* IOC-3844 o nível de mRNA é superexpresso com o valor de 112.798,46 e 44.421,61 vezes respectivamente para 36 e 60 horas de indução em comparação a condições de não indução (Figura 5). Isto também mostra expressão muito mais elevada do que a reportada por (VERBEKE et al., 2009) para *egl1* induzido de *T. reesei* CL847, que mostrou valores de 80 e 30 vezes mais expresso em 24 e 48 horas de indução respectivamente, quando comparado a micélio não induzido. A mesma observação é válida para o transcrito *swol* de *T. harzianum* IOC-3844 que mostrou valores de superexpressão de 6.909,49 e 2.456,22 vezes mais expresso em 36 e 60 horas de indução respectivamente (Figura 5), enquanto que (VERBEKE et al., 2009) reportaram valores por volta de 100 vezes mais expressos em 24 horas e 10 vezes mais expresso em 48 horas de indução de *swol* em *T. reesei* CL847. Além destes resultados de PCR em tempo real revelando alta expressão relativa de *swol* de *T. harzianum* IOC-3844, os resultados mostrados na tabela 2 e na figura 7, estão em concordância com os resultados de (FOREMAN et al., 2003) que mostraram que a *swol* de *T. reesei* é co-regulada com as celulasas, reforçando sua provável importância como enzima acessória.

A figura 6 ilustra os níveis de mRNA dos transcritos envolvidos com degradação de biomassa induzidos por 24, 36, 48 e 60 horas, em comparação com os mRNAs induzidos por 8 horas. Com isso, pode ser observado um aumento geral no número de

transcritos de 24/8 para 36/8 horas, com valores de aumento na superexpressão variando a partir de 12,29 vezes para *egl1*, com 20,7 vezes para *cbh1* e até 148,95 vezes para *egl7*, calculados pela razão entre os valores de expressão de 36/8 horas e os valores de 24/8 horas (valores listados no apêndice H). Depois, de 36/8 para 60/8 horas há uma ligeira diminuição na superexpressão (Figura 6). O trabalho de (VERBEKE et al., 2009) mostra um ensaio similar com *T. reesei* CL847 no qual os níveis de mRNA dos transcritos *cbh1*, *egl1*, *egl4* e *bgl1*, induzidos com celulose por 4 até 48 horas, foram comparados com os níveis dos transcritos em 2 horas de indução com celulose. Estes autores também encontraram um aumento geral do número de transcritos de 4 a 10 horas de indução, com valores entre cerca de 4 e 6 vezes mais expressos, e uma diminuição de expressão aparentemente mais abrupta entre 10 e 48 horas de indução. No entanto, este aumento geral de transcritos não foi tão agudo quanto em *T. harzianum* IOC-3844. Apesar dos ensaios do presente trabalho e dos ensaios de (VERBEKE et al., 2009) terem sido realizados com tempos de coletas diferentes e terem em comum somente os transcritos *cbh1* e *egl1*, é possível especular que a expressão de genes celulolíticos em *T. harzianum* IOC-3844 é mais forte e duradoura do que em *T. reesei* CL847. De fato, os níveis de mRNA mostrados neste trabalho e discutidos até aqui, indicam o *T. harzianum* IOC-3844 como potencial produtor do complexo celulolítico, como previamente descrito por (CASTRO et al., 2010a) em suas análises com o sobrenadante de cultura induzida do mesmo fungo.

Os dados da figura 7 permitem comparações entre os níveis absolutos de mRNA de cada transcrito de *T. harzianum* IOC-3844, em cada tempo de indução estudado (8, 24, 36, 48 e 60 horas). Vale destacar que para o *T. harzianum* IOC-3844, as abundâncias relativas dos transcritos de 60 horas de indução, ensaiados com PCR em tempo real (Figura 7), correspondem às abundâncias relativas encontradas na contagem de ESTs para os mesmos transcritos, obtidos também do mesmo micélio induzido por 60 horas. Há apenas a exceção dos transcritos *egl1* e *egl3* que invertem a posição, mas que, no entanto, apresentaram níveis de expressão similar (Tabela 5). Estas observações legitimam as suposições comentadas acima a respeito das diferenças de expressão entre *T. harzianum* IOC-3844 e *T. reesei* QM6a indicadas pela contagem de ESTs (Tabela 3).

**Tabela 5: Comparação entre os dados de expressão de PCR em tempo real e a contagem de ESTs de cultura de *Trichoderma harzianum* IOC-3844 induzida por 60 horas.** Ambos os experimentos foram realizados com RNA extraído de um mesmo micélio induzido por 60 horas.

Os valores de expressão relativa a actina foram obtidos do ensaio de PCR em tempo real, como descrito para a figura 7. A contagem de ESTs para o transcrito da *actina* na biblioteca de *T. harzianum* IOC-3844 foi 1 (não mostrado).

PCR em tempo real			Biblioteca de cDNA	
Transcrito	Expressão relativa à actina		Transcrito	Numero de ESTs
<i>cbh1</i>	40.5	superexpresso	<i>cbh1</i>	34
<i>cbh2</i>	24.82	superexpresso	<i>cbh2</i>	28
<i>egl7</i>	11.51	superexpresso	<i>egl7</i>	16
<i>egl1</i>	3.92	superexpresso	<i>egl3</i>	6
<i>egl3</i>	3.81	superexpresso	<i>egl1</i>	4
<i>egl2</i>	2.29	superexpresso	<i>egl2</i>	3
<i>swol</i>	1.9	subexpresso	<i>swol</i>	1

Analisando os resultados mostrados na figura 7; e os dados da tabela 3, é possível assumir que os níveis de mRNA de *cbh1* e *cbh2* são mais equivalentes em *T. harzianum* IOC-3844 do que em *T. reesei* QM6a, que mostra uma maior proporção de transcritos de *cbh1*. De fato, níveis bem maiores de expressão de *cbh1* têm sido reportados em relação à *cbh2* e às demais celulasas (GRITZALI & BROWN, 1979; CARLE-URIESTE et al., 1997; ILMÉN et al., 1997; FOREMAN et al., 2003). Os níveis mais equilibrados de mRNA de *cbh1* e *cbh2* em *T. harzianum* IOC-3844, e portanto a maior expressão relativa de *cbh2*, podem ajudar a explicar as altas atividades de FPase deste fungo, em comparação ao *T. reesei* Rut-C30, reportadas por (CASTRO et al., 2010a).

Analisando as endoglucanases *egl1*, *egl2*, *egl3* e *egl7* de *T. harzianum* IOC-3844 em condições indutoras de 36 a 60 horas (Figura 7), e tendo por base a alta expressão relativa da *egl1*, é possível propor que o *T. harzianum* IOC-3844 possui uma expressão geral de endoglucanases mais elevada em relação ao *T. reesei* CL847 estudado por (VERBEKE et al., 2009). Ao observar os dados da tabela 3 as mesmas observações são válidas e indicam expressão aumentada de endoglucanases de *T. harzianum* IOC-3844 em relação ao *T. reesei* QM6a, mesmo levando em consideração que os transcritos *egl6* e *egl8* não apareceram na biblioteca de cDNA de *T. harzianum* IOC-3844. A expressão mais elevada de *egl7* entre as endoglucanases induzidas de *T. harzianum* IOC-3844 (Figura 7 e tabela 2), juntamente com a expressão relativamente elevada das demais endoglucanases (tabela 3), podem explicar os altos conteúdos de atividade endoglucanásica do *T. harzianum* IOC-3844 em comparação ao *T. reesei* Rut-C30, encontrados por (CASTRO et al., 2010a).



O elevado número de transcritos de *T. harzianum* IOC-3844 envolvidos com degradação de biomass vegetal, descrito neste trabalho, pode estar relacionado com as altas produtividades de atividades celulolíticas descritas por (CASTRO et al., 2010a) para este fungo, em comparação ao *T. reesei* Rut-C30. Neste contexto, é possível supor que existem diferentes mecanismos de regulação da transcrição no *T. harzianum* IOC-3844 que culminam em uma elevada abundância de mRNAs, e que conseqüentemente resultam em alto número relativo de proteínas envolvidas com a degradação de biomass vegetal.

Paradoxalmente, no entanto, apesar de haver variações entre as medidas, de um modo geral o conteúdo de proteína total presente nos sobrenadantes de cultura induzida do *T. harzianum*, tem sido reportado como significativamente menor (CASTRO et al., 2010a; DELABONA et al., 2012) em comparação ao conteúdo de proteína total de *T. reesei* (SADDLER, HOGAN, LOUIS-SEIZE, 1985; HERPOËL-GIMBERT et al., 2008; PETERSON & NEVALAINEN, 2012). Neste contexto, é possível sugerir que o *T. harzianum* IOC-3844 apresenta maiores atividades específicas das enzimas que degradam biomass vegetal, como explicação para as altas atividades celulolíticas reportadas neste fungo por (CASTRO et al., 2010a), em relação ao *T. reesei* Rut-C30. Isto já havia sido sugerido por (SADDLER, HOGAN, LOUIS-SEIZE, 1985) para o *T. harzianum* E58, em comparação com o *T. reesei* Rut-C30. Por outro lado, estes autores também haviam sugerido que o *T. reesei* Rut-C30 poderia apresentar baixa percentagem de celulasas no sobrenadante, enquanto que estas enzimas constituiriam a maior parte das proteínas presentes no sobrenadante de cultura do *T. harzianum* E58 (SADDLER, HOGAN, LOUIS-SEIZE, 1985). Isto explicaria como a menor quantidade de proteína total encontrada no sobrenadante de *T. harzianum* E58 estava relacionada com a maior atividade específica de celulasas reportada para este fungo, em comparação ao *T. reesei* Rut-C30 (SADDLER, HOGAN, LOUIS-SEIZE, 1985). Contrariamente, o trabalho de (HERPOËL-GIMBERT et al., 2008) não sugere baixas percentagens de celulasas nos sobrenadantes de culturas induzidas de *T. reesei* Rut-C30 e de *T. reesei* CL847. No entanto, o trabalho de (JUN, KIESELBACH, JÖNSSON, 2011) sugere menores percentagens relativas de celulasas no sobrenadante induzido de *T. reesei* Rut-C30, e isso mostra contradição nas informações acerca da abundância de celulasas no secretoma de *T. reesei*.

Neste contexto, para explicar as diferenças entre *T. harzianum* IOC-3844 e *T. reesei*, descritas neste trabalho e reportadas por (CASTRO et al., 2010a), se fazem

necessários estudos comparativos adicionais com intuito de esclarecer eventuais diferenças nos mecanismos reguladores da transcrição, entender as diferenças da quantidade de mRNAs e suas relações com as enzimas envolvidas com degradação de biomassa, presentes nos sobrenadantes de cultura (secretoma), e mapear diferenças de atividades enzimáticas específicas.

## **5.5 Expressão e ensaios de atividade das proteínas recombinantes**

Como organismo eucarioto a levedura *Pichia pastoris* apresenta vantagens como hospedeira para expressão heteróloga, proporcionando processamento, enovelamento e modificações pós traducionais apropriadas para produzir proteínas ativas, originadas de organismos eucariotos, incluindo os eucariotos superiores. Também, a *Pichia pastoris* é fácil de manipular e é capaz de produzir grandes quantidades de proteína em relação a outros hospedeiros eucariotos, além de apresentar opções com sistemas que secretam as proteínas recombinantes, facilitando a purificação (BUCKHOLZ & GLEESON, 1991; ROMANOS, SCORER, CLARE, 1992). Além disso, a *Pichia pastoris* tem sido usada para a expressão de celulasas e hemicelulasas em estudos que visam solucionar as demandas das aplicações comerciais destas enzimas (Banerje et al., 2010c; AHMED, RIAZ, JAMIL, 2009).

A produção heteróloga de celulasas é importante para estudos que visam otimizar a atividade enzimática (NAKAZAWA et al., 2009; VOUTILAINEN et al., 2010) e para a suplementação de extratos de enzimas, obtidos de sobrenadantes de culturas de fungos celulolíticos ou disponíveis comercialmente (LYND et al., 2002; WANG et al., 2012). Além disso, existem abordagens que usam misturas sintéticas de celulasas, exclusivamente obtidas com enzimas produzidas via expressão heteróloga, para ensaios de degradação de biomassa vegetal (BANERJEE, SCOTT-CRAIG, WALTON, 2010). Interessantemente, (MEYER, ROSGAARD, SØRENSEN, 2009) defendem que enzimas recombinantes produzidas em sistemas heterólogos podem ser usadas para compor misturas sintéticas, para degradação de biomassa, que já se mostram competitivas em relação ao custo, quando comparadas a misturas obtidas por expressão nativa.

Entre os trabalhos que envolvem a expressão heteróloga da EGIII de *Trichoderma*, encontra-se o de (OKADA et al., 1998) que reporta a expressão da EGIII de *T. reesei* QM9414 em *Saccharomyces cerevisiae* e *Schizosaccharomyces pombe*.

Estes autores mostraram forte atividade celulolítica de EGIII recombinante de *T. reesei* QM9414, através de ensaios em placas com CMC coradas com Vermelho Congo. Neste trabalho a expressão da EGIII também foi realizada em *E. coli*, porém a maior parte da enzima produzida neste hospedeiro se mostrou inativa.

Outro trabalho envolvendo a expressão heteróloga da EGIII de *Trichoderma* foi publicado por (WANG et al., 2005). Estes autores, interessados na aplicação industrial de enzimas mais tolerantes em ambiente alcalino, realizaram a evolução dirigida da EGIII de *T. reesei*, por meio da técnica de “error-prone PCR”. O trabalho reporta a expressão heteróloga da EGIII de *T. reesei* em *Saccharomyces cerevisiae* H158. Após os ensaios em placas com CMC coradas com Vermelho Congo, uma variante da EGIII recombinante apresentou mudança da atividade ótima de pH 4,8 para pH 5,4 em relação à enzima nativa (WANG et al., 2005).

Mais recentemente, usando expressão heteróloga em *Escherichia coli* “RosettaBlue DE3”, (NAKAZAWA et al., 2009) fizeram evolução dirigida da EGIII de *T. reesei* QM9414. Após ensaiar as variantes obtidas através de “error-prone PCR” em placas com CMC coradas com Vermelho Congo, estes autores obtiveram uma variante da EGIII com maior atividade específica, com ampla estabilidade entre pHs 4,4 e 8,8 e com termo-tolerância aumentada em relação a variante nativa.

Os ensaios bem sucedidos de expressão heteróloga e de atividade da EGIII de *T. harzianum* IOC-3844 em *Pichia pastoris* KM71H, mostrados neste trabalho (Figuras 8, 9 e 11), foram continuados por outro estudante do Laboratório de Biologia Molecular DGE-UFSCar. Isto proporcionou a caracterização enzimática da EGIII recombinante. A produção heteróloga desta enzima também permitiu estudos estruturais que foram conduzidos por nossos parceiros da equipe do professor Dr. Igor Polikarpov, do Instituto de Física da USP de São Carlos.

Estes estudos proporcionaram a publicação de dois artigos envolvendo a EGIII do *T. harzianum* IOC-3844. Um deles, publicado no periódico “Genetics and Molecular Research” (ver apêndice B), reporta os ensaios bem sucedidos de expressão heteróloga e purificação da enzima EGIII do *T. harzianum* IOC-3844, como mostrado nas figuras 8 e 9; e os ensaios qualitativos em placas contendo CMC coradas com Vermelho Congo, como mostrado na figura 11. Além disso, o manuscrito reporta ensaios adicionais que revelaram: a melhor atividade da enzima em pH 5,5 e temperatura de 48,2 °C; estabilidade em temperaturas de até 40 °C; valor de *K<sub>m</sub>* de 21,35 g/L e valor de *K<sub>cat</sub>* de 1880 min<sup>-1</sup> (GENEROSO et al., 2012). O outro manuscrito é o de (LIBERATO et al.,

2012) que foi publicado no periódico “Acta Crystallographica” (ver apêndice C) e reporta a cristalização e análise preliminar de difração de raio-X da EGIII do *T. harzianum* IOC-3844, além de apontar a preferência desta enzima pelo substrato de CMC em detrimento de substratos contendo celulose microcristalina.

A expressão recombinante de EGIII de *T. harzianum* IOC-3844 também abriu possibilidade para futuros testes de combinação desta enzima com as demais celulasas. Isto permitirá a avaliação de interações sinérgicas para esclarecer a possível contribuição da EGIII de *T. harzianum* IOC-3844 no rendimento de hidrólise durante a degradação de biomass vegetal, uma vez que esta enzima pode apresentar importante função acessória e hidrólise preferencial em celulose amorfa (ver item anterior “Contagem de ESTs” da Discussão).

A produção de xilanases apresenta importância devido à sua aplicabilidade na indústria de papel onde estas enzimas apresentam potencial para auxiliar na remoção da lignina residual, diminuindo o uso de agentes químicos durante o processo de branqueamento da polpa de celulose (POLIZELI et al., 2005). Por outro lado, a produção de xilanases também é importante para a degradação de biomass vegetal, particularmente nos processos que utilizam pré-tratamento que não removem, ou que removem parcialmente a hemicelulose. De fato, (GAO et al., 2011) mostraram a importância de suplementar as misturas, ou preparados de celulasas, com hemicelulasas, incluindo as xilanases, no processo de degradação de biomassa pré-tratada com o método alcalino denominado AFEX (GAO et al., 2011), que não extrai ou separa a hemicelulose da biomassa. Desta forma, a adição de xilanases em biomassa pré-tratada, contendo hemicelulose, promove maiores rendimentos de hidrólise, pois ao degradar o xilano o acesso à celulose fica facilitado, e também ocorre a anulação da inibição que os xilooligômeros promovem na atividade hidrolítica das celulasas (HU et al., 2011; MAEDA et al., 2011; QING, YANG, WYMAN, 2010).

Neste contexto, os genes que codificam para xilanases tem sido clonados em hospedeiros heterólogos para a produção de proteínas visando a super produção das mesmas e possibilitando estudos que visam alterar suas propriedades para atender as demandas das aplicações comerciais (AHMED, RIAZ, JAMIL, 2009).

Entre os trabalhos que envolvem a expressão heteróloga da Xyn2 de *Trichoderma*, encontra-se o de (LA GRANGE, PRETORIUS, VAN ZYL, 1996) que reportou a expressão da Xyn2 de *T. reesei* QM6a em *Saccharomyces cerevisiae* Y294. Estes autores reportaram as maiores atividades da Xyn2 recombinante em pH 6,0 e

temperatura de 60 °C, e os ensaios de termo-estabilidade mostraram retenção de mais de 90 % de atividade depois de 30 minutos a 50 °C. Este trabalho aponta ainda a importância do sistema desenvolvido para a indústria de papel, uma vez que foi obtida, no sobrenadante de cultura de *Saccharomyces cerevisiae* Y294, a Xyn2 recombinante livre de atividade celulolítica.

Mais recentemente, (HE et al., 2009) reportaram a expressão da Xyn2 de *T. reesei* Rut-C30 em *Pichia pastoris* X33. Vale citar que estes autores utilizaram o mesmo plamídeo pPIC-Z- $\alpha$ -A, usado por nós para a expressão da Xyn2 de *T. harzianum* IOC-3844 em *Pichia pastoris* KM71H muts. De modo muito semelhante ao trabalho de (LA GRANGE, PRETORIUS, VAN ZYL, 1996) com o *T. reesei* QM6a, o trabalho de (HE et al., 2009) mostrou para a Xyn2 recombinante de *T. reesei* Rut-C30, as maiores atividades em pH 6,0 e temperatura de 60 °C, e os ensaios de termo-estabilidade mostraram retenção de mais de 94 % de atividade depois de 30 minutos a 50 °C. Interessantemente, os ensaios de (HE et al., 2009) foram realizados diretamente a partir do sobrenadante de cultura da *P. pastoris* e os autores reportaram a Xyn2 como a principal proteína presente no sobrenadante, assim como encontrado por nós para a Xyn2 de *T. harzianum* IOC-3844 expressa em *P. pastoris* KM71H muts (Figura 10). Por fim, além de reportarem atividade xilanásica praticamente livre de atividade celulolítica da Xyn2 recombinante de *T. reesei* Rut-C30, (HE et al., 2009) reportaram também altos níveis de atividade desta enzima, como sugerido aqui para Xyn2 recombinante de *T. harzianum* IOC-3844 (Figura 12).

Diferentemente da EGIII de *T. harzianum* que ainda não tinha relatos de expressão heteróloga na literatura, a enzima Xyn2 de *T. harzianum* C4 já foi expressa de modo recombinante em *Saccharomyces cerevisiae* SEY2102 por (LEE et al., 2009). Depois de ensaiar 30 colônias transformantes em placas contendo Xilano coradas com Vermelho Congo, estes autores selecionaram a colônia contendo a Xyn2 recombinante mais ativa para utilizar os sobrenadantes de cultura para os ensaios de atividade subsequentes. Com isso foram reportadas as maiores atividades da Xyn2 recombinante em pH 5,5 e temperatura de 50 °C. Neste trabalho também foram realizados ensaios adicionais com a Xyn2 purificada diretamente do sobrenadante de cultura do *T. harzianum* C4 através de cromatografia. Com isso, as maiores atividades encontradas foram em pH 5,1 e temperatura de 50 °C. Os autores ainda reportaram a produção e purificação de Xyn2 do sobrenadante de *T. harzianum* C4 e a produção da enzima com uso do *Saccharomyces cerevisiae* SEY2102 transformado, como livres de atividade

celulolítica, apresentando assim importância para a indústria de papel e celulose, apesar de não terem apresentado altos rendimentos (LEE et al., 2009).

Os ensaios mostrados aqui envolvendo de expressão heteróloga da Xyn2 de *T. harzianum* IOC-3844 foram bem sucedidos (Figura 10) e a enzima se mostrou ativa em ensaios qualitativos em placas contendo Xilano e coradas com Vermelho Congo (Figura 12). A expressão recombinante de Xyn2 de *T. harzianum* IOC-3844 realizada neste trabalho permitiu a realização de ensaios de purificação e caracterização enzimática que estão sendo conduzidos no Laboratório de Biologia Molecular DGE-UFSCar por outro estudante. A produção recombinante da Xyn2 também possibilita futuros testes de combinação desta enzima, com misturas de celulasas para avaliar as possíveis influências via interações sinérgicas, no rendimento de hidrólise durante a degradação de biomass vegetal contendo hemicelulose. Além disso, podem ser realizados futuros ensaios de termo-estabilidade e de evolução dirigida, objetivando a aplicabilidade desta enzima na indústria de papel.

## 6 CONCLUSÕES

- Foi construída uma biblioteca de cDNA, do fungo *Trichoderma harzianum* IOC-3844, induzida para produção de transcritos envolvidos na degradação de biomass vegetal.
- Um total de 2016 clones da biblioteca induzida do *Trichoderma harzianum* IOC-3844 foi sequenciado, totalizando 1544 seqüências após a exclusão de seqüências de baixa qualidade e de plasmídeo.
- O sequenciamento da biblioteca de induzida de *Trichoderma harzianum* IOC-3844 permitiu a identificação de vinte transcritos envolvidos com a degradação de biomass vegetal, indicando que este fungo possui um pequeno repertório destes transcritos, assim como reportado para *Trichoderma reesei*.
- Sete transcritos representativos do transcriptoma induzido do *Trichoderma harzianum* IOC-3844 foram estudados por PCR em tempo real e tiveram os níveis de expressão comparados: com transcriptoma do fungo não induzido, com as primeiras horas de indução e com a expressão do transcrito de *actina* no transcriptoma induzido.
- Os dados de expressão do *T. harzianum* IOC-3844, analisados por contagem de ESTs e por PCR em tempo real, indicaram níveis de expressão significativamente aumentados neste fungo, em comparação ao *Trichoderma reesei*, para os transcritos relacionados com a desconstrução da biomass vegetal.
- Foram desenvolvidos protocolos de expressão heteróloga para as enzimas EGIH e Xyn2 de *Trichoderma harzianum* IOC-3844.
- Os protocolos de expressão heteróloga, das enzimas de *Trichoderma harzianum* IOC-3844, abriram caminho para estudos que envolvem: caracterização enzimática, determinação de estruturas, engenharia de proteínas, evolução dirigida, interações sinérgicas e otimização de preparados enzimáticos.
- Os estudos desenvolvidos neste trabalho indicaram o fungo *Trichoderma harzianum* IOC-3844 como uma linhagem promissora para a produção de enzimas que degradam biomass vegetal, apresentando potencial aplicabilidade para a produção de etanol de segunda geração.
- Os resultados também indicaram este fungo como uma fonte potencial de enzimas para a indústria de papel.

## 7 PERSPECTIVAS FUTURAS

- Completar o seqüenciamento das ORFs para as proteínas de *T. harzianum* IOC-3844 envolvidas com degradação de biomassa vegetal para coparação com as ORFs correspondentes de *T. reesei*;
- Anotar a biblioteca induzida de *T. harzianum* IOC-3844;
- Buscar esclarecimento da função de transcritos sem função associada com alta contagem de ESTs da biblioteca de *T. harzianum* IOC-3844;
- Comparar a biblioteca induzida (deste trabalho) com as não induzidas (disponíveis publicamente) de *T. harzianum*;
- Comparar a biblioteca induzida de *T. reesei* com as não induzidas (disponíveis publicamente);
- Avaliar a expressão por PCR em tempo real das celulasas de *Trichoderma reesei* para comparar com as celulasas de *Trichoderma harzianum* IOC-3844;
- Sequenciar e comparar os transcriptomas induzidos de *T. reesei* e de *T. harzianum* em plataforma de nova geração;
- Sequenciar o genoma do *T. harzianum* IOC-3844 em plataforma de nova geração;
- Continuar os ensaios de expressão heteróloga e caracterização cinética das enzimas de *T. harzianum* IOC-3844 envolvidas com a degradação de biomassa;
- Comparar atividades das enzimas de *T. reesei* e *T. harzianum* produzidas no mesmo hospedeiro;
- Anular o fator de transcrição *cre1* do *T. harzianum* IOC-3844 para avaliar o possível aumento de produtividade de celulasas e hemicelulasas;
- Melhorar as enzimas do *T. harzianum* IOC-3844 com evolução dirigida e engenharia de proteínas;
- Avaliar combinações de enzimas recombinantes do *T. harzianum* IOC-3844 com outras enzimas e preparados enzimáticos na degradação de biomassa vegetal;
- Analisar as sequencias promotoras dos genes envolvidos com degradação de biomassa de *T. harzianum* IOC-3844 e comparar com as seqüências de *T. reesei* para mapear diferenças na regulação transcricionl frente a substratos diversos.



## 8 REFERÊNCIA

ABRAHÃO-NETO, J. et al. Mitochondrial functions mediate cellulase gene expression in *Trichoderma reesei*. **Biochemistry**, v. 34, p.10456-10462, 1995.

AHAMED, A.; VERMETTE, P. Enhanced enzyme production from mixed cultures of *Trichoderma reesei* RUT-C30 and *Aspergillus niger* LMA grown as fed batch in a stirred tank bioreactor. **Biochemical Engineering Journal**, v. 42, p. 41-46, 2008.

AHMED, S.; RIAZ, S.; JAMIL, A. Molecular cloning of fungal xylanases: an overview. **Applied Microbiology and Biotechnology**, v. 84, p. 19-35, 2009.

ALBERSHEIM, P. et al. Principles of cell wall architecture and assembly. In: **Plant cell walls: from chemistry to biology**. New York: Garland Science, 2010, p. 234-271.

ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, p. 403-410, 1990.

ALTSCHUL, S.F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, p. 3389-3402, 1997.

ARANTES, V.; SADDLER, J. N. Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. **Biotechnology for Biofuels**, v. 3, n. 4, p. 1-11, 2010.

ARO, N. et al. ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression. **Applied and Environmental Microbiology**, v. 69, n.1, p. 56-65, 2003.

ARO, N.; PAKULA, T.; PENTTILA, M. Transcriptional regulation of plant cell wall degradation by filamentous fungi. **FEMS Microbiology Reviews**, v. 29, p. 719-739, 2005.

ARO, N. et. al. ACEII, a novel transcriptional activator involved in regulation of cellulase and xylanase genes of *Trichoderma reesei*. **The Journal of Biological Chemistry**, v. 276, n. 26, p. 24309-24314, 2001.

BAJPAI, P.; ANAND, A.; BAJPAI, P. K. Bleaching with lignin-oxidizing enzymes. **Biotechnology Annual Review**, v. 12, p. 349-378, 2006.

BALAT, M.; BALAT, H. Recent trends in global production and utilization of bio-ethanol fuel. **Applied Energy**, v. 86, p. 2273-2282, 2009.

BANERJEE, G. et al. Synthetic enzyme mixtures for biomass deconstruction: production and optimization of a core set. **Biotechnology and Bioengineering**, v. 106, n. 5, p. 707-720, 2010a.

BANERJEE, G. et al. Rapid optimization of enzyme mixtures for deconstruction of diverse pretreatment/biomass feedstock combinations. **Biotechnology for Biofuels**, v. 3, n. 22, p. 1-15, 2010b.

BANERJEE, G.; SCOTT-CRAIG, J. S.; WALTON, J. D.: Improving enzymes for biomass conversion: a basic research perspective. **Bioenergy Research**, v. 3, p. 82-92, 2010.

BENÍTEZ, T. et al. Biocontrol mechanisms of *Trichoderma* strains. **International Microbiology**, v. 7, p. 249-260, 2004.

BENSON, D. A. et al GenBank. **Nucleic Acids Research**, v. 39, p. D32-D37, 2011.

BERMAN, H. M. et al. The protein data bank. **Acta Crystallographica Section D: Biological Crystallography**, v. 58, n. 1, p. 899-907, 2002.

BRODEUR, G. et al. Chemical and physicochemical pretreatment of lignocellulosic biomass: a review. **Enzyme Research**, n. 2011, p. 1-17, 2011.

BUCKHOLZ; R. G.; GLEESON, M. A. G. Yeast systems for the commercial production of heterologous proteins. **Nature Biotechnology**, v. 9, p. 1067-1072, 1991.

CAMPBELL, M. M.; SEDEROFF, R. R. Variation in lignin content and composition: mechanisms of control and implications for the genetic improvement of plants. **Plant Physiol**, v. 110, p. 3-13, 1996.

CAPELARI, M.; TOMÁS-PEJÓ, E. Lignin degradation and in vitro digestibility of wheat straw treated with Brazilian tropical species of white-rot fungi. **Folia Microbiologica**, v. 42, p. 481-487, 1997.

CARLE-URIOSTE; J. C. et al. Cellulase induction in *Trichoderma reesei* by cellulose requires its own basal expression. **The Journal of Biological Chemistry**, v. 272, n. 15, p. 10169-10174, 1997.

CASTRO, A. M. et al. High-yield endoglucanase production by *Trichoderma harzianum* IOC-3844 cultivated in pretreated sugarcane mill byproduct. **Enzyme Research**, v. 2010, p. 1-8, 2010a.

CASTRO, A. M. et al. *Trichoderma harzianum* IOC-4038: a promising strain for the production of a cellulolytic complex with significant  $\beta$ -glucosidase activity from sugarcane bagasse cellulignin. **Applied Biochemistry and Biotechnology**, v. 62, n. 7, p. 2111-2122, 2010b.

CHAVERRI, P. et al. Multilocus phylogenetic structure within the *Trichoderma harzianum/Hypocrea lixii* complex. **Molecular Phylogenetics and Evolution**, v. 27, p. 302-313, 2003.

DASHTBAN, M.; SCHRAFT, H.; QIN, W. Fungal bioconversion of lignocellulosic residues; opportunities & perspectives. **International Journal of Biological Sciences**, v. 5, n. 6, p. 578-595, 2009.

DELABONA, P. S. et al. Use of a new *Trichoderma harzianum* strain isolated from the Amazon rainforest with pretreated sugar cane bagasse for on-site cellulase production. **Bioresource Technology**, v. 107, p. 517-521, 2012.

DIENER, S. E. et al. Characterization of the protein processing and secretion pathways in a comprehensive set of expressed sequence tags from *Trichoderma reesei*. **FEMS Microbiology Letters**, v. 230, n. 2, p. 275-282, 2004.

DILLON, A. J. P. et al. Use of 2-deoxyglucose in liquid media for the selection of mutant strains of *Penicillium echinulatum* production increased cellulase and  $\beta$ -glucosidase activities. **Applied Microbiology and Biotechnology**, v. 70, p. 740-746, 2006.

DIVNE, C. et al. High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*. **Journal of Molecular Biology**, v. 275, p. 309-325, 1998.

EK, M.; GELLERSTEDT, G.; HENRIKSSON, G. Chemistry of chemical pulping In: EK, M.; GELLERSTEDT, G.; HENRIKSSON, G. **Pulp and paper chemistry and**

**technology:** pulping chemistry and technology. Stockholm Sweden: Walter de Gruyter & Co, 2009. v. 2, p. 91-120.

EL-GOGARY, S. et al. Mechanism by which cellulose triggers cellobiohydrolase I gene expression in *Trichoderma reesei*. **Proceedings of the National Academy of Sciences**, v. 86, p. 6138-6141, 1989.

ERIKSSON, O.; GORING, DAI.; LINDGREN, B. O. Structural studies on the chemical bonds between lignins and carbohydrates in spruce wood. **Wood science and Technology**, v. 14, n. 4, p. 267-279, 1980.

EWING, B.; GREEN, P. Base-calling of automated sequencer traces using phred. II: Error probabilities. **Genome Research**, v. 8, n. 3, p. 186-194, 1998.

EWING, B. et al. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Research**, v. 8, n. 3, p. 175-185, 1998.

FOREMAN, P. K. Et al. Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. **The Journal of Biological Chemistry**, v. 278, n. 34, p. 31988-31997, 2003.

GAO, D. et al. Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. **Biotechnology for Biofuels**, v. 4, n. 5, p. 1-11, 2011.

GENE RUNNER. Versão 3.01. Disponível em: <<http://www.generunner.net>>. Acesso em: 16/ago/2010.

GENEROSO, W. C. et al: Recombinant expression and characterization of an endoglucanase III (cel12a) from *Trichoderma harzianum* (Hypocreaceae) in the yeast *Pichia pastoris*. **Genetics and Molecular Research**, v. 11, n. 2, p. 1544-1557, 2012.

GHISALBERTI, E. L.; SIVASITHAMPARAM, K. Anti fungal antibiotics produced by *Trichoderma* spp. **Soil Biology & Biochemistry**, v. 23, p. 1011-1020, 1991.

GOLDMAN, G. H.; HAYES, C.; HARMAN, G. E. Molecular and cellular biology of biocontrol by *Trichoderma* spp. **Trends in Biotechnology**, v. 12, p. 478-482, 1994.

GRITZALI, M.; BROWN JUNIOR, R. D. The cellulose system of *Trichoderma*: the relationship between purified extracellular enzymes from induced or cellulose grown cells. In: BROWN JUNIOR, R.D.; JURASEK, L. (Ed.). **Hydrolysis of cellulose: mechanisms of enzymatic and acid catalysis**. Washington DC: American Chemical Society; 1979. v. 181. p. 237-260.

GUILFOILE, P. et al. Cloning and sequencing of a cellobiohydrolase gene from *Trichoderma harzianum* FP108. **Journal of the Minnesota Academy of Science**, v. 64, n. 1, p. 18-22, 1999.

GUILLÉN, F. et al. Oxygen activation during oxidation of methoxyhydroquinones by laccase from *Plerotus eryngii*. **Applied and Environmental Microbiology**, v. 66, n. 1, p. 170-175, 2000.

GUO, Y. et al. dCAS: a desktop application for cDNA sequence annotation. **Bioinformatics**, v. 25, n. 9, p. 1195-1196, 2009.

GUSAKOV, A. V. Alternatives to *Trichoderma reesei* in biofuel production. **Trends in Biotechnology**, v. 29, n. 9, p. 419-425, 2011.

HARMAN, G. E. et al. *Trichoderma* species- opportunistic, avirulent plant symbionts. **Nature Reviews Microbiology**, v. 2, p. 43-56, 2004.

HARRIS, J. F. Acid hydrolysis and dehydration reactions for utilizing plant carbohydrates. **Applied Polymer Symposium**, v. 28, p. 131-144, 1975.

HE, J. et al. Expression of endo-1, 4-beta-xylanase from *Trichoderma reesei* in *Pichia pastoris* and functional characterization of the produced enzyme. **BMC Biotechnology**, v. 9, n. 56, p. 1-10, 2009.

HENDRIKS A. T. W. M.; ZEEMAN G. Pretreatments to enhance the digestibility of lignocellulosic biomass. **Bioresource Technology**, v. 100, p. 10-18, 2009.

HENRIKSSON, G. et al. Endoglucanase 28 (Cel12A), a new *Phanerochaete chrysosporium* cellulose. **European Journal of Biochemistry**, v. 259, p. 88-95, 1999.

HERPOËL-GIMBERT, I. et al. Comparative secretome analyses of two *Trichoderma reesei* RUT-C30 and CL847 hypersecretory strains. **Biotechnology for Biofuels**, v. 1, n.18, p. 1-12, 2008.

HILL, J. et al. Climate change and health costs of air emissions from biofuels and gasoline. **Proceedings of the National Academy of Sciences**, v. 106, n. 6, p. 2077-2082, 2009.

HIMMEL, M. E. et al. Biomass recalcitrance: engineering plants and enzymes for biofuels. **Science**, v. 315, n. 5813, p. 804-807, 2007.

HU, J.; ARANTES, V.; SADDLER, J. N. The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: is it an additive or synergistic effect?. **Biotechnology for Biofuels**, v. 4, n. 36, p. 1-13, 2011.

HUANG, X.; MADAN, A. CAP3: A DNA sequence assembly program. **Genome Research**, v. 9, n. 9, p. 868-877, 1999.

ILMÉN, M. et al. Regulation of cellulase gene expression in the filamentous fungus *Trichoderma reesei*. **Applied and Environmental Microbiology**, v. 63, n. 4, p. 1298-1306, 1997.

ILMÉN, M.; THRANE, C.; PENTTILA, M. The glucose repressor gene creI of *Trichoderma*: isolation and expression of a full-length and a truncated mutant form. **Molecular and General Genetics**, v. 251, p. 451-460, 1996.

JÄGER, G. et al. How recombinant swollenin from *Kluyveromyces lactis* affects cellulosic substrates and accelerates their hydrolysis. **Biotechnology for Biofuels**, v. 4, p. 33, 2011.

JUN, H.; KIESELBACH, T.; JÖNSSON, L. J. Enzyme production by filamentous fungi: analysis of the secretome of *Trichoderma reesei* grown on unconventional carbon source. **Microbial Cell Factories**, v. 10, n. 68, p. 1-10, 2011.

KABEL, M. A. et al. Standard assays do not predict the efficiency of commercial cellulase preparations towards plant materials. **Biotechnology and Bioengineering**, v. 93, n. 1, p. 56-63, 2006.

KÁDÁR, Z.; SZENGYEL, Z.; RÉCZE, K. Simultaneous saccharification and fermentation (SSF) of industrial wastes for the production of ethanol. **Industrial Crops and Products**, v. 20, p. 103-110, 2004.

KARKEHABADI S. et al. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. **Journal of Molecular Biology**, v. 383, p. 144-154, 2008.

KARLSSON, J. et al. Homologous expression and characterization of Cel61A (EGIV) of *Trichoderma reesei*. **European Journal of Biochemistry**, v. 268, p. 6498-6507, 2001.

KARLSSON, J. et al. Enzymatic properties of the low molecular mass endoglucanases Cel12A (EG III) and Cel45A (EG V) of *Trichoderma reesei*. **Journal of Biotechnology**, v. 99, p. 63-7, 2002.

KERSTEN, P.; CULLEN, D. Extracellular oxidative systems of the lignin-degrading Basidiomycete *Phanerochaete chrysosporium*. **Fungal Genetics and Biology**, v. 44, p. 77-87, 2007.

KNAUF, M.; MONIRUZZAMAN, M. Lignocellulosic biomass processing: a perspective. **International Sugar Journal**, v. 106, n. 1263, p. 147-150, 2004.

KREDICS, L. et al. Clinical importance of the genus *Trichoderma*. A review. **Acta Microbiologica et Immunologica Hungarica**, v. 50, p. 105-117, 2003.

KUBICEK, C. P. et al. The *Trichoderma cellulase* regulatory puzzle: from the interior life of a secretory fungus. **Enzyme and Microbial Technology**, v. 15, n. 2, p. 90-99, 1993.

KUBICEK, C. P. et al. Metabolic engineering strategies for the improvement of cellulase production by *Hypocrea jecorina*. **Biotechnology for Biofuels**, v. 2, p. 19, 2009.

KUHAD, R. C.; GUPTA, R.; SINGH, A. Microbial cellulases and their industrial applications. **Enzyme Research**, v. 2011, p. 1-10, 2011.

KUHLS, K. et al. Molecular evidence that the asexual industrial fungus *Trichoderma reesei* is a clonal derivative of the ascomycete *Hypocrea jecorina*. **Proceedings of the National Academy of Sciences**, v. 93, p. 7755-7760, 1996.

LA GRANGE, D. C.; PRETORIUS, I. S.; VAN ZYL, W. H. Expression of a *Trichoderma reesei* b-xylanase gene (Xyn2) in *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 62, n. 3, p. 1036-1044, 1996.

LEE, H. L. et al. Construction and characterization of different fusion proteins between cellulases and  $\beta$ -glucosidase to improve glucose production and thermostability. **Bioresource Technology**, v. 102, n. 4, p. 3973-3976, 2011.

LEE, J. M. et al. Molecular cloning and expression of the *Trichoderma harzianum* C4 endo- $\beta$ -1,4-xylanase gene in *Saccharomyces cerevisiae*. **Journal of Microbiology and Biotechnology**, v. 19, n. 8, p. 823-828, 2009.

LIBERATO, M. V. et al. Crystallization and preliminary X-ray diffraction analysis of endoglucanase III from *Trichoderma harzianum*. **Acta Crystallographica - Section F**, n. 68, p. 306-309, 2012.

LIN, H. et al. Advances in the study of directed evolution for cellulases. **Frontiers of Environmental Science & Engineering in China**, v. 5, n. 4, p. 519-525, 2011.

LINDER, M.; TEERI, T. T. The cellulose-binding domain of the major cellobiohydrolase of *Trichoderma reesei* exhibits true reversibility and a high exchange rate on crystalline cellulose. **Proceedings of the National Academy of Sciences**, v. 93, p. 12251-12255, 1996.

LIU, P.; YANG, Q. Identification of genes with a biocontrol function in *Trichoderma harzianum* mycelium using the expressed sequence tag approach. **Research in Microbiology**, v. 156, p. 416-423, 2005.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_T}$  method. **Methods**, v. 25, n. 4, p. 402-408, 2001.

LIYAMA, K.; LAM, T. B. T.; STONE, B. A. Covalent cross-links in the cell wall. **Plant Physiology**, v. 104, p. 315-320, 1994.

LÓPEZ-MONDÉJAR, R.; ROS, M.; PASCUAL, J. A. Mycoparasitism-related genes expression of *Trichoderma harzianum* isolates to evaluate their efficacy as biological control agent. **Biological Control**, v. 56, p. 59-66, 2011.



LYND, L. R. et al. Microbial cellulose utilization: fundamentals and biotechnology. **Microbiology and Molecular Biology Reviews**, v. 66, n. 3, p. 506-577, 2002.

MACH-AIGNER, A. R. et al. Transcriptional regulation of *xyl1*, encoding the main regulator of the xylanolytic and cellulolytic enzyme system in *Hypocrea jecorina*. **Applied and Environmental Microbiology**, v. 74, n. 21, p. 6554-6562, 2008.

MACLEAN, H. L.; LAVE, L. B. Evaluating automobile fuel/propulsion system technologies. **Progress in Energy and Combustion Science**, v. 29, p. 1-69, 2003.

MAEDA, R. N. Et al. Enzymatic hydrolysis of pretreated sugar cane bagasse using *Penicillium funiculosum* and *Trichoderma harzianum* cellulases. **Process Biochemistry**, v. 46, p. 1196-1201, 2011.

MALAGÓ-JR, W.; SOARES-COSTA, A.; HENRIQUE-SILVA, F. DNA as genetic material: revisiting classic experiments through a simple, practical class. **Biochemistry and Molecular Biology Education**, v. 37, n. 5, p. 290-295, 2009.

MANDELS, M.; PARRISH, F. W.; REESE, E. T. Sophorose as an inducer of cellulase in *Trichoderma viride*. **Journal of Bacteriology**, v. 83, n. 2, p. 400-408, 1962.

MANDELS, M.; REESE, E. T. Induction of cellulase in fungi by cellobiose. **Journal of Bacteriology**, v. 79, n. 6, p. 816-826, 1960.

MANDELS, M.; REESE, E. T. Induction of cellulase in *Trichoderma viride* as influenced by carbon sources and metals. **Journal of Bacteriology**, v. 73, p. 269-278, 1957.

MARCHESSAULT, R. H.; HOWSMON, J. A. Experimental evaluation of the lateral order distribution in cellulose. **Textile Research Journal**, v. 27, p. 30-41, 1957.

MARGOLLES-CLARK, E.; ILMÉN, M.; PENTTILA, M. Expression patterns of ten hemicellulase genes of the filamentous fungus *Trichoderma reesei* on various carbon sources. **Journal of Biotechnology**, v. 57, p. 167-179, 1997.

MARGOLLES-CLARK, E. et al. The  $\alpha$ -glucuronidase-encoding gene of *Trichoderma reesei*. **Gene**, v. 172, n. 1, p. 171-172, 1996d.

MARGOLLES-CLARK, E. et al. Three  $\alpha$ -galactosidase genes of *Trichoderma reesei* cloned by expression in yeast. **European Journal of Biochemistry**, v. 240, p. 104-111, 1996c.

MARGOLLES-CLARK, E. et al. Cloning of genes encoding  $\alpha$ -L-Arabinofuranosidase and  $\alpha$ -Xylosidase from *Trichoderma reesei* by expression in *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 62, n. 10, p. 3840-3846, 1996b.

MARGOLLES-CLARK, E. et al. Acetyl xylan esterase from *Trichoderma reesei* contains an active-site serine residue and a cellulose-binding domain. **European Journal of Biochemistry**, v. 237, p. 553-560, 1996a.

MARTINEZ, D. et al. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. **Nature Biotechnologies**, v. 22, p. 695-700, 2004.

MARTINEZ, D. et al. Genome sequencing of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). **Nature Biotechnology**, v. 26, n. 5, p. 553-560, 2008.

MERINO, S. T.; CHERRY, J. Progress and challenges in enzyme development for biomass utilization. **Advances in Biochemical Engineering / Biotechnology**, v. 108, p. 95-120, 2007.

MEYER, A. S.; ROSGAARD, L.; SØRENSEN, H. R. The minimal enzyme cocktail concept for biomass processing. **Journal of Cereal Science**, v. 50, p. 337-344, 2009.

MIN, L. J. et al. Molecular cloning and expression of the *Trichoderma harzianum* C4 Endo- $\beta$ -1,4-xylanase gene in *Saccharomyces cerevisiae*. **Journal of Microbiology and Biotechnology**, v. 19, n. 8, p. 823-828, 2009.

MOSIER, N. et al. Features of promising technologies for pretreatment of lignocellulosic biomass. **Bioresource Technology**, v. 96, p. 673-686, 2005.

NAKARI-SETALA, T. et al. Genetic modification of carbon catabolite repression in *Trichoderma reesei* for improved protein production. **Applied and Environmental Microbiology**, v. 75, n. 14, p. 4853-4860, 2009.

NAKAZAWA, H. et al. Directed evolution of endoglucanase III (Cel12A) from *Trichoderma reesei*. **Applied Microbiology and Biotechnology**, v. 83, p. 649-657, 2009.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. Gene. Disponível em: <<http://www.ncbi.nlm.nih.gov/gene>>. Acesso em: 13/out/2009.

NEHRING, R. Traversing the mountaintop: world fossil fuel production to 2050. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 364, p. 3067-3079, 2009.

OGASAWARA, W. et al. Cloning, functional expression and promoter analysis of xylanase III gene from *Trichoderma reesei*. **Applied Microbiology and Biotechnology**, v. 72, n. 5, p. 995-1003, 2006.

OKADA, H. et al. Molecular characterization and heterologous expression of the gene encoding a low-molecular-mass endoglucanase from *Trichoderma reesei* QM9414. **Applied and Environmental Microbiology**, v. 64, n. 2, p. 555-563, 1998.

OLOFSSON, K.; BERTILSSON, M.; LIDÉN, G. A short review on SSF - an interesting process option for ethanol production from lignocellulosic feedstocks. **Biotechnology for Biofuels**, v. 1, p. 7-14, 2008.

PALLAPOLU, V. R. et al. Effects of enzyme loading and b-glucosidase supplementation on enzymatic hydrolysis of switchgrass processed by leading pretreatment technologies. **Bioresource Technology**, v. 102, p. 11115-11120, 2011.

PAN, X. et al. Biorefining of softwoods using ethanol organosolv pulping: preliminary evaluation of process streams for manufacture of fuel-grade ethanol and co-products. **Biotechnology and Bioengineering**, v. 90, n. 4, p. 473-481, 2005.

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

PEREIRA, N. J.; COUTO, M. A. P. G.; ANNA, L. M. M. S. **Biomass of lignocellulosic composition for fuel ethanol production within the context of biorefinery**. Rio de Janeiro: Biblioteca Nacional, 2008. 45 p. (Series on Biotechnology, v. 2)

PETERSON, R.; NEVALAINEN, H. *Trichoderma reesei* RUT-C30 - thirty years of strain improvement. **Microbiology**, v. 158, n. 1, p. 58-68, 2012.

PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acid Research**, v. 29, n. 900, p. 2002-2007, 2001.

POLIZELI, M. L. T. M. et al. Xylanases from fungi properties and industrial applications. **Applied Microbiology and Biotechnology**, v. 67, p. 577-591, 2005.

PORTNOY, T. et al. The CRE1 carbon catabolite repressor of the fungus *Trichoderma reesei*: a master regulator of carbon assimilation. **BMC Genomics**, v. 12, n. 269, p. 1-12, 2011.

Primer3. Disponível em: <<http://frodo.wi.mit.edu/primer3>>. 04/mar/2011.

QING, Q.; YANG, B.; WYMAN, C. E. Xylooligomers are strong inhibitors of cellulose hydrolysis by enzymes. **Bioresource Technology**, v. 101, n. 24, p. 9624-9630, 2010.

QUIROZ-CASTAÑEDA, R. E. et al. Loosenin, a novel protein with cellulose-disrupting activity from *Bjerkandera adusta*. **Microbial Cell Factories**, v. 10, n. 8, p. 1-9, 2011.

RAUSCHER, R. et al. Transcriptional regulation of *xynI*, encoding xylanase I, in *Hypocrea jecorina*. **Eukaryotic Cell**, v. 5, p. 447-456, 2006.

REST. Software, 2009. Disponível em: <<http://www.qiagen.com/REST>>. Acesso em dia 18/jul/2011.

ROMANOS, M. A.; SCORER, C. A., CLARE, J. J. Foreign gene expression in yeast: a review. **Yeast**, v. 8, p. 423-488, 1992.

RULLER, R. et al. Thermostable variants of the recombinant xylanase A from *Bacillus subtilis* produced by directed evolution show reduced heat capacity changes. **Proteins**, v. 70, n. 4, p. 1280-1293, 2008.

SADDLER, J. N.; HOGAN, C. M. C.; LOUIS-SEIZE, G. A comparison between the cellulase systems of *Trichoderma harzianum* E58 and *Trichoderma reesei* C30. **Applied Microbiology and Biotechnology**, v. 22, p. 139-145, 1985.

SAHA, B. C. Hemicellulose bioconversion. **Journal of Industrial Microbiology and Biotechnology**, v. 30, p. 279-291, 2003.

SALOHEIMO, A. et al. Isolation of the *ace1* gene encoding a Cys2-His2 transcription factor involved in regulation of activity of the cellulase promoter *cbh1* of *Trichoderma reesei*. **The Journal of Biological Chemistry**, v. 275, p. 5817-5825, 2000.

SALOHEIMO, M., et al. cDNA cloning of a *Trichoderma reesei* cellulase and demonstration of endoglucanase activity by expression in yeast. **European Journal of Biochemistry**, v. 249, p. 584-591, 1997.

SALOHEIMO, M. et al. Swollenin, a *Trichoderma reesei* protein with sequence similarity to the plant expansins, exhibits disruption activity on cellulosic materials. **European Journal of Biochemistry**, v. 269, p. 4202-4211, 2002.

SALVACHÚA, D. et al. Fungal pretreatment: An alternative in second-generation ethanol from wheat straw. **Bioresource Technology**, v. 102, p. 7500-7506, 2011.

SAMBROOK, J.; RUSSELL, D. W. **Molecular cloning**: a laboratory manual. 3. ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001.

SAMIRIT, D.; JAGDEEP, K. Microbial mannanases: An overview of production and applications. **Critical Reviews in Biotechnology**, v. 27, n. 4, p. 197-216, 2008.

SANDGREN, M.; STAHLBERG, J.; MITCHINSON, C. Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. **Progress in Biophysics and Molecular Biology**, v. 89, p. 246-291, 2005.

SCHUSTER, A.; SCHMOLL, M. Biology and biotechnology of *Trichoderma*. **Applied Microbiology and Biotechnology**, v. 87, p. 87-799, 2010.

SERVICE, R. F. Is there a road ahead for cellulosic ethanol? **Science**, v. 329, p. 784-785, 2010.

SHALLOM, D.; SHOHAM, Y. Microbial hemicellulases. **Current Opinion in Microbiology**, v. 6, p. 219-228, 2003.

SHARMA, P. et al. Biocontrol genes from *Trichoderma* species: a review. **African Journal of Biotechnology**, v. 10, n. 86, p. 19898-19907, 2011.

SHI, J.; CHINN, M. S.; SHARMA-SHIVAPPA, R. R. Microbial pretreatment of cotton stalks by solid state cultivation of *Phanerochaete chrysosporium*. **Bioresource Technology**, v. 99, p. 6556-6564, 2008.

SIGNALP 4.0 SERVER. Versão 4.0. Disponível em: <<http://www.cbs.dtu.dk/services/SignalP>>. Acesso em 21/jan/2010.

SLADE, R.; BAUEN, A.; SHAH, N. The greenhouse gas emissions performance of cellulosic ethanol supply chains in Europe. **Biotechnology for Biofuels**, v. 2, n. 15, p. 1-19, 2009.

SMITH, P. K. et al. Measurement of protein using bicinchoninic acid. **Analytical Biochemistry**, v. 1, p. 76-85, 1985.

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

SPREY, B.; BOCHEM, H. P. Effect of endoglucanase and cellobiohydrolase from *Trichoderma reesei* on cellulose microfibril structure. **FEMS Microbiology Letters**, v. 97, p. 113-118, 1992.

STALBRAND, H. et al. Purification and characterization of two  $\beta$ -mannanases from *Trichoderma reesei*. **Journal of Biotechnology**, v. 29, p. 229-242, 1993.

STEINDORFF, A. S. et al. *Trichoderma harzianum* expressed sequence tags for identification of genes with putative roles in mycoparasitism against *Fusarium solani*. **Biological Control**, v. 61, n. 2, p. 134-140, 2012.

STRICKER, A. R. et al. Xyr1 (xylanase regulator 1) regulates both the hydrolytic enzyme system and D-xylose metabolism in *Hypocrea jecorina*. **Eukaryotic Cell**, v. 5, p. 2128-2137, 2006.

SUKUMARAN, R. K. et al. Cellulase production using biomass feed stock and its application in lignocellulose saccharification for bio-ethanol production. **Renewable Energy**, v. 34, n. 2, p. 421-424, 2009.

SUKUMARAN, R. K.; SINGHANIA, R. R.; PANDEY, A. Microbial cellulases – Production, applications and challenges. **Journal of Scientific and Industrial Research**, v. 64, p. 832-844, 2005.

TAN, L. U. L. et al. Purification and characterization of two d-xylanases from *Trichoderma harzianum*. **Enzyme and Microbial Technology**, v. 7, n. 9, p. 425-430, 1985.

TEERI, T. T. et al. *Trichoderma reesei* cellobiohydrolases: why so efficient on crystalline cellulose? **Biochemical Society Transactions**, v. 26, p. 173-178, 1998.

TENKANEN, M. et al. Action of *Trichoderma reesei* mannanase on galactoglucomannan in pine kraft pulp. **Journal of Biotechnology**, v. 57, p. 191-204, 1997.

TENKANEN, M.; PULS, J.; POUTANEN, K. Two major xylanases of *Trichoderma reesei*. **Enzyme and Microbial Technology**, v. 14, p. 566-574, 1992.

TIEN, M. Properties of ligninase from *Phanerochaete chrysosporium* and their possible applications. **Critical Reviews in Microbiology**, v. 15, n. 2, p. 141-168, 1987.

TORRIE, J. P.; SENIOR, D. J.; SADDLER, J. N. Production of  $\beta$ -mannanases by *Trichoderma harzianum* E58. **Applied Microbiology and Biotechnology**, v. 34, p. 303-307, 1990.

TURNER, P.; MAMO, G.; KARLSSON, E. N. Potential and utilization of thermophiles and thermostable enzymes in biorefining. **Microbial Cell Factories**, v. 6, n. 9, p. 1-23, 2007.

ULKER, A.; SPREY, B. Characterization of an unglycosylated low molecular weight 1,4-b-glucan-glucohydrolase of *Trichoderma reesei*. **FEMS Microbiology Letters**, v. 69, p. 215-219, 1990.

U.S. DEPARTMENT OF ENERGY. **Breaking the Biological Barriers to Cellulosic Ethanol**: a Joint Research Agenda, DOE/SC-0095. Washington, DC: US-DOE, 2006. p. 181-184.

VAAJE-KOLSTAD, G. et al. The non-catalytic chitin-binding protein CBP21 from *Serratia marcescens* is essential for chitin degradation. **The Journal of Biological Chemistry**, v. 280, p. 28492–28497, 2005.

VAN VLEET, J. H.; JEFFRIES, T. W. Yeast metabolic engineering for hemicellulosic ethanol production. **Current Opinion in Biotechnology**, v. 20, p. 300-306, 2009.

VERBEKE, J. et al. Transcriptional profiling of cellulase and expansin-related genes in a hypercellulolytic *Trichoderma reesei*. **Biotechnology Letters**, v. 31, p. 1399-1405, 2009.

VETTORE, A. L. et al. Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. **Genome Research**, v. 13, n. 12, p. 2725-2735, 2003.

VIKARI, L. et al. Xylanases in bleaching: from an idea to the industry. **FEMS Microbiology Letters**, v. 13, p. 335-355, 1994.

VIKARI, L. et al. Thermostable enzymes in lignocellulose hydrolysis. In: SCHEPER, T. et al (Ed.). **Advances in Biochemical Engineering/Biotechnology**, v.108, p.121-145, 2007.

VIZCAÍNO, J. Á. et al. Generation, annotation and analysis of ESTs from *Trichoderma harzianum* CECT 2413. **BMC Genomics**, v. 7, n. 193, p. 1-14, 2006.

VIZCAÍNO, J. Á. et al. Generation, annotation, and analysis of ESTs from four different *Trichoderma* strains grown under conditions related to biocontrol. **Applied Microbiology and Biotechnology**, v. 75, p. 853-862, 2007.

VOET, D.; VOET, J. G. **Biochemistry**. New York: John Wiley & Sons; 1995.

VOUTILAINEN, S. P. et al. Expression of *Talaromyces emersonii* cellobiohydrolase Cel7A in *Saccharomyces cerevisiae* and rational mutagenesis to improve its thermostability and activity. **Protein Engineering Design and Selection**, v. 23, n. 2, p. 69-79, 2010.

WALTON, J.; BANERJEE, G.; CAR, S. GENPLAT: an automated platform for biomass enzyme discovery and cocktail optimization. **Journal of Visualized Experiment**, v. 56, p. 1-5, 2011.



WAN, C.; LI, Y. Microbial pretreatment of corn stover with *Ceriporiopsis subvermispora* for enzymatic hydrolysis and ethanol production. **Bioresource Technology**, n. 16, p. 6398-6403, 2010.

WANG, D. et al Maximum saccharification of cellulose complex by an enzyme cocktail supplemented with cellulase from newly isolated *Aspergillus fumigatus* ECU0811. **Applied Microbiology and Biotechnology**, v. 166, p. 176-186, 2012.

WANG, T. et al. Directed evolution for engineering pH profile of endoglucanase III from *Trichoderma reesei*. **Biomolecular Engineering**, v. 22, p. 89-94, 2005.

WONG, K. K. Y. et al. Purification of a third distinct xylanase from the xylanolytic system of *Trichoderma harzianum*. **Canadian Journal of Microbiology**, v. 32, n. 7, p. 570-576, 1986.

WONG, K. K. Y.; TAN, L. U. L.; SADDLER, J. N. Multiplicity of  $\beta$ -1,4-Xylanase in microorganisms: Functions and applications. **Micricrobiological Reviews** 1988, v. 52, n. 3, p. 305-317, 1988.

WU, C. H. et al. The Universal Protein Resource (UniProt): an expanding universe of protein information. **Nucleic Acids Research**, v. 1, n. 34, p. 187-191, 2006.

WYMAN, C. E. What is (and is not) vital to advancing cellulosic ethanol. **Trends in Biotechnology**, v. 25, n. 4, p. 153-157, 2007.

XIONG, H. et al. Improved xylanase production by *Trichoderma reesei* grown on L-arabinose and lactose or D-glucose mixtures. **Applied Microbiology and Biotechnology**, v. 64, p. 353-358, 2004.

XU, J. et al Regulation of xyn3 gene expression in *Trichoderma reesei* PC-3-7. **Applied Microbiology and Biotechnology**, v. 54, n. 3, p. 370-375, 2000.

YEDIDIA, I.; BENHAMOU, N.; CHET, I. Induction of defense responses in cucumber plants (*Cucumis sativus* L.) by the biocontrol agent *Trichoderma harzianum*. **Applied and Environmental Microbiology**, v. 65, n. 3, p. 1061-1070, 1999.

YUAN, S.; WU, Y.; COSGROVE, D. J. A Fungal endoglucanase with plant cell wall extension activity. **Plant Physiology**, v. 127, p. 324-333, 2001.

ZEILINGER, S. et al. Different inducibility of expression of the two xylanase genes *xyn1* and *xyn2* in *Trichoderma reesei*. **The Journal of Biological Chemistry**, v. 271, n. 41, p. 25624-25629, 1996.

ZHAO, X.; CHENG, K.; LIU, D. Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis. **Applied Microbiology and Biotechnology**, v. 82, n. 5, p. 815-827, 2009.

ZIMBARDI, F.; RICCI, E.; BRACCIO, G. Technoeconomic study on steam explosion application in biomass processing. **Applied Biochemistry and Biotechnology**, v. 98-100, p. 89-99, 2002.

## 9 APÊNDICES

---

Apêndice A: Artigo redigido e submetido para publicação

---

Apêndice B: Artigo publicado

---

Apêndice C: Artigo publicado

---

Apêndice D: Artigo publicado

---

Apêndice E: Objetivos e cronograma propostos

---

Apêndice F: Panorama dos estudos de clonagem e expressão heteróloga

---

Apêndice G: Alinhamentos entre as enzimas seqüenciadas de *Trichoderma harzianum* IOC-3844 e as correspondentes homólogas de *Trichoderma reesei*

---

Apêndice H: Valores de expressão calculados para as figuras 5, 6 e 7

---

**Apêndice A: Artigo redigido e submetido para publicação.**

**Transcriptional profile of biomass degrading genes from *Trichoderma harzianum* IOC-3844: a promising source of cellulases and hemicellulases for biomass deconstruction and second generation ethanol conversion.**

Wilson Malagó-Jr <sup>1</sup>, Ludier Kesser Santos-Silva <sup>1</sup>, Nei Pereira Jr <sup>2</sup>, Flávio Henrique-Silva <sup>1,\*</sup>

<sup>1</sup> Department of Genetics and Evolution, Federal University of São Carlos, Rodovia Washington Luis, Km 235, São Carlos, 13565-905, SP, Brazil.

<sup>2</sup> Biochemical Engineering Department, School of Chemistry, Federal University of Rio de Janeiro, P.O. Box 68542, Rio de Janeiro, 21941-598, RJ, Brazil

\*Corresponding author

Email addresses:

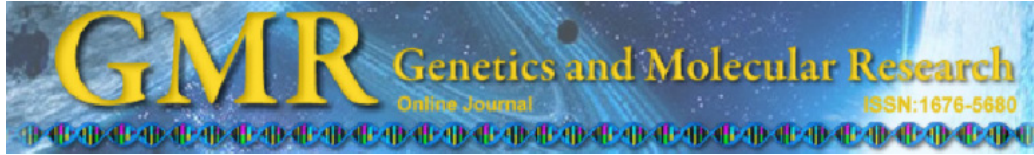
WMJ: malago.jr@gmail.com

LKSS: ludierludy@yahoo.com.br

NPJ: nei@eq.ufrj.br

FHS: dfhs@ufscar.br

**Apêndice B: Artigo publicado.**



# Recombinant expression and characterization of an endoglucanase III (cel12a) from *Trichoderma harzianum* (Hypocreaceae) in the yeast *Pichia pastoris*

W.C. Generoso<sup>1</sup>, W. Malagó-Jr.<sup>1</sup>, N. Pereira Jr.<sup>2</sup> and F. Henrique-Silva<sup>1</sup>

<sup>1</sup>Laboratório de Biologia Molecular, Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos, SP, Brasil

<sup>2</sup>Laboratórios de Desenvolvimento de Bioprocessos, Departamento de Engenharia Bioquímica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil

Corresponding author: F. Henrique-Silva  
E-mail: [dfhs@ufscar.br](mailto:dfhs@ufscar.br)

Genet. Mol. Res. 11 (2): 1544-1557 (2012)

Received February 7, 2012

Accepted April 25, 2012

Published May 21, 2012

DOI <http://dx.doi.org/10.4238/2012.May.21.11>

**ABSTRACT.** Filamentous fungi from the genus *Trichoderma* have been widely investigated due to their considerable production of important biotechnological enzymes. Previous studies have demonstrated that the *T. harzianum* strain IOC-3844 has a high degree of cellulolytic activity. After excluding the native signal peptide, the open reading frame of the *T. harzianum* endoglucanase III enzyme was cloned in the expression vector pPICZ $\alpha$ A, enabling protein secretion to the culture medium. The recombinant plasmid was used to transform *Pichia pastoris*. Recombinant expression in the selected clone yielded 300 mg pure enzyme per liter of induced medium. The recombinant enzyme proved to be active in a qualitative analysis using Congo red. A quantitative assay, using dinitrosalicylic acid, revealed a high degree of activity at pH 5.5 and around 48°C. This information contributes to our understanding of the cellulolytic repertory of *T. harzianum* and the

determination of a set of enzymes that can be incorporated into mixes for second-generation ethanol production.

**Key words:** *Trichoderma harzianum*; Second-generation ethanol; *Pichia pastoris*

## INTRODUCTION

Plant biomass has emerged as a feasible source for biofuel production due to its relatively low cost and large-scale feedstock procurement (Lynd et al., 2002, 2008). As Brazil is the main producer of sugarcane worldwide, sugarcane bagasse is the country's most abundant source of plant residual biomass and is mainly destined for the production of electricity in boilers (Pessoa-Jr et al., 2005). This bagasse contains about 50% cellulose, 25% hemicellulose and 25% lignin, which can be used as feedstock for different organisms, resulting in value-added products or essentially more sugar for ethanol fermentation, which is important to the sustainability of energy consumption, especially when considering the limited land resources and noncompetitive land use (Pandey et al., 2000; Lynd et al., 2008).

The depolymerization of polysaccharides is the key step in ethanol production from lignocellulosic biomass. Known since 1819, acid treatment is the precursor method of biomass hydrolysis and involves either concentrated acids at low temperatures or diluted acids at high temperatures (Galbe and Zacchi, 2002). The main disadvantage of acid hydrolysis is the generation of fermentation inhibitors (Galbe and Zacchi, 2002). The use of microbial enzymes is an attractive alternative for cellulose depolymerization due to the specificity and efficiency of the enzymes as well as the non-inhibitor feature and rapid hydrolysis (Arantes and Saddler, 2010; Banerjee et al., 2010).

The cellulolytic complex basically involves  $\beta$ -1,4-endoglucanases, which internally cleave the cellulose molecule, cellobiohydrolases, which break down cellulose into cellobiose, and  $\beta$ -glycosidases, which hydrolyze cellobiose into glucose (Foreman et al., 2003). A number of different microorganisms produce cellulases, among which the most often employed are the filamentous fungi of the genus *Trichoderma* (Lynd et al., 2002; Schuster and Schmoll, 2010). *Trichoderma reesei* (*Hypocrea jecorina*) is known to be an effective cellulase producer and is currently the main source of enzymatic cocktails for the degradation of biomass (Schuster and Schmoll, 2010). On the other hand, strains of *Trichoderma harzianum* (*Hypocrea lixii*) are seldom reported as cellulase producers and are more frequently known as control agents against fungal pathogens (Arantes and Saddler, 2010; Banerjee et al., 2010). Nevertheless, recent studies have revealed the potential of this filamentous fungus for large-scale cellulase production and industrial applications (Ahmed et al., 2009; de Castro et al., 2010a,b). In studying the *T. harzianum* IOC-3844 strain, de Castro and collaborators (2010a) found a well-balanced enzymatic complex comparable to commercial cocktails and with greater endoglucanase content, making this strain a promising producer of cellulolytic complex for the hydrolysis of sugarcane bagasse.

The present paper describes the cloning of an endoglucanase from *T. harzianum* IOC-3844 and its heterologous production in the yeast *Pichia pastoris*. This enzyme exhibits a high degree of similarity with endoglucanase III (cel12a) from *T. reesei*. Endoglucanase III (EGIII) enzymes have a low molecular mass and lack the cellulose binding domain (CBD) (Okada

et al., 1998; Henriksson et al., 1999; Nakazawa et al., 2009). These enzymes are also able to degrade amorphous cellulose and have a compact  $\beta$ -jellyroll fold (Okada et al., 1998; Henriksson et al., 1999; Sandgren et al., 2003, 2005; Nakazawa et al., 2009).

The aim of the present study was to determine the proprieties of EGIII from *T. harzianum* IOC-3844 expressed in *P. pastoris* with regard to biochemical and enzyme parameters.

## MATERIAL AND METHODS

### Strains and media

*Escherichia coli* DH5 $\alpha$  (Invitrogen, USA) was used for the general propagation of plasmids. Bacterial cells were grown with agitation at 37°C in a low-salt Luria-Bertani medium containing Zeocin™ for a final concentration of 25  $\mu$ g/mL (EasySelect Pichia Expression kit, 2001). *P. pastoris* KM71H<sup>(Mut<sup>S</sup>)</sup> (Invitrogen) was used as the expression host. Yeast cultures were maintained on a yeast extract-peptone-dextrose (YEPD) medium. The media for growth and induction were BMGY (buffered glycerol-complex medium) and with BMMY (buffered methanol-complex medium), respectively, both at pH 6.5 (EasySelect Pichia Expression kit, 2001).

### Cloning of endoglucanase ORF and transformation in *P. pastoris*

EGIII cDNA was obtained from a clone of a cDNA library constructed from *T. harzianum* IOC-3844 induced by microcrystalline cellulose (Malagó-Jr. W, Santos-Silva LK, Pereira N Jr. and Henrique-Silva F, unpublished results). This clone was found by comparison using the BLASTx (Altschul et al., 1990) local alignment program and had its open reading frame (ORF) entirely sequenced.

The EGIII ORF was cloned into pPICZ $\alpha$ A, excluding the putative signal peptide. The mature ORF was amplified with the forward primer 5'-CGGAATTCCAGACCAGCTGCGAA C-3' and reverse primer 5'-CGCGTTCGACGTTGATAGATGCGGTCC-3' (underlined bases indicate introduced restriction sites).

Both the amplicon and pPICZ $\alpha$ A plasmid were digested with *Eco*RI and *Sal*I and ligated, creating a recombinant plasmid containing the mature EGIII ORF flanked by the secretion signal peptide ( $\alpha$ -factor) at the N-terminal and 6xHisTag at the C-terminal. The pPICZ vectors also contain an inducible promoter from the alcohol oxidase gene AOX1, which is induced by methanol.

Before *P. pastoris* transformation, the recombinant plasmid was linearized with *Pme*I endonuclease and then introduced into the yeast by electroporation (1.5 kV, 25  $\mu$ F, 200  $\Omega$ ), following the method described by Creeg (2007). Transformants were cultivated on solid YEPD with 1 M sorbitol and 100  $\mu$ g/mL Zeocin. Recombinants were identified by polymerase chain reaction using the  $\alpha$ -factor and AOX 3'-primers.

### Expression, purification and identification of recombinant enzyme

The transformant yeasts were screened for protein induction in 24-well plates, as described by Boettner et al. (2002). A recombinant yeast was selected for enzyme production and purification.



Expression induction for purification was carried out following the manual of the *P. pastoris* expression kit (EasySelect Pichia Expression kit, 2001). A single colony of the recombinant yeast was grown overnight in 5 mL BMGY at 30°C with shaking at 250 rpm. The culture was used to inoculate 500 mL fresh BMGY in a baffled shake flask and was grown at 30°C and at 250 rpm, for 24 h ( $OD_{600\text{nm}} = 6$ ). The content was centrifuged at 1500 g, for 5 min. The supernatant was discarded and the cells were resuspended in 100 mL BMMY containing 1% methanol. The culture was maintained at 30°C and at 250 rpm, with the daily addition of 0.75% methanol to maintain induction until the best production time. The supernatants were recovered and analyzed using SDS-PAGE (Sambrook and Russell, 2001) and as well as through the qualitative identification of enzyme activity using the Congo red method (Teather and Wood, 1982), with modifications. For such, 20  $\mu\text{L}$  of each supernatant was dispensed on plates containing 1% carboxymethylcellulose (CMC) and 2% agar and incubated for 20 min at 30°C. The plate was then stained with 1% Congo red solution for about 30 min and destained with 1 M sodium chloride until the appearance of degradation halos.

Recombinant enzyme purification from the supernatant was performed by affinity chromatography using a 5-mL silica resin column containing nickel (Ni-NTA Superflow resin, Qiagen, USA). The binding buffer employed consisted of 10 mM Tris base, 50 mM sodium phosphate and 100 mM sodium chloride at pH 8.0. The elution buffer was the same as the binding buffer, except for the addition of imidazole ranging from 5 to 250 mM. The polyhistidine-tagged protein was eluted with a 2-fold resin volume of elution buffer at each imidazole concentration used. The fractions containing purified protein were pooled and dialyzed in membranes of 14,000 MW (Pierce, USA) three times for 4 h at 4°C in 2 L 50 mM Tris-HCl buffer, pH 7.0. The enzyme solution was sterilized by filtration through a 0.22- $\mu\text{m}$  membrane, quantified using the BCA Protein Assay kit (Thermo Scientific, USA) and stored at -20°C.

For the identification of EGIII from *T. harzianum* IOC-3844, the purified protein and supernatant of the fungus were separated on 12% acrylamide gels and were used for Western blotting. *T. harzianum* was induced with 1% microcrystalline cellulose for 36 h, following the method described by Mandels et al. (1962). For Western blotting, SDS-PAGE-separated proteins were blotted onto a polyvinylidene fluoride (PVDF) membrane (Pierce) using the Mini-Trans-Blot Module (Mini-Protean Tetra Cell, Bio-Rad, USA) for 2 h in transfer buffer [20 mM Tris base, 55 mM glycine and 16.5% (v/v) methanol]. The separated proteins were detected with the primary anti-ThEGIII antibody and the secondary anti-mouse IgG-alkaline phosphatase-conjugated antibody (Sigma, Germany), both diluted to 1:10,000 in Tris-buffered saline (TBS, pH 7.4). The BCIP/NBT-Blue Liquid Substrate System (Sigma) was used for detecting alkaline phosphatase.

Anti-Th-EGIII polyclonal antibodies were produced in 20-day-old Swiss albino mice (*Mus musculus*) subjected to two injections of 100  $\mu\text{g}$  purified recombinant *T. harzianum* EGIII. The first immunization was performed using 100  $\mu\text{L}$  protein solution combined with 100  $\mu\text{L}$  Freund's complete adjuvant (Sigma). The second inoculation was performed after 45 days, with 100  $\mu\text{L}$  protein conjugated with 100  $\mu\text{L}$  incomplete Freund's adjuvant (Sigma). The mice were sacrificed after 10 days; blood was collected and centrifuged at 15,700 g and at 4°C, for 5 min to obtain the serum containing the antibodies.

### CMCase activity assays

CMCase activity was determined using the 3,5-dinitrosalicylic acid method (Miller,

1959) based on the quantification of the reducing sugar content, using glucose as the standard for the calibration curves. All assays were performed in triplicate with 2% low-viscosity CMC (50 to 200 cps, Sigma).

Briefly, 45  $\mu$ L 2% CMC in 100 mM McIlvaine's buffer, pH 5.5, was incubated in a Peltier bath at 42°C for 15 min; 5  $\mu$ L enzyme solution was then added to the reaction. The mixture was incubated for the enzyme reaction and the reaction was stopped after 12 min by adding 100  $\mu$ L dinitrosalicylic acid solution. The mixture was boiled for 5 min, and the reducing sugars were measured at 540 nm with a spectrophotometer (Victor<sup>3</sup>, Perkin-Elmer, USA).

The optimal temperature for EGIII was determined by altering the reaction temperature within a range of 20° to 50°C in 100 mM McIlvaine's buffer, pH 5.5. The optimal temperature was then employed to determine the optimal pH, altering the reaction buffer within a range from 2.5 to 7.0. For such, the following buffers were used: 100 mM sodium-citrate buffer (pH 2.5 to 4.0), 100 mM McIlvaine's buffer (pH 4.5 to 6.0), 100 mM sodium phosphate buffer (pH 6.5 to 7.0) and 100 mM Tris-HCl buffer (pH 7.5 to 8.0).

Thermostability was analyzed through the measurement of CMCase activity following pre-incubation of the purified enzyme for 1 h in the absence of substrate at temperatures ranging from 0° to 65°C. The thermal resistance of the enzyme was determined by monitoring the residual activity every 24 h of pre-incubation at the last stable temperature (40°C). Residual activity was measured at optimal pH and temperature.

$K_{cat}$  and  $K_m$  values were derived from a non-linear Michaelis-Menten curve of enzyme activity, measured by varying CMC concentrations from 3 to 32.4 g/L at the optimal pH and temperature.

All experimental data and graphs were generated with the Prism 5 statistical program (GraphPad, USA). Values are reported as the mean and standard deviation.

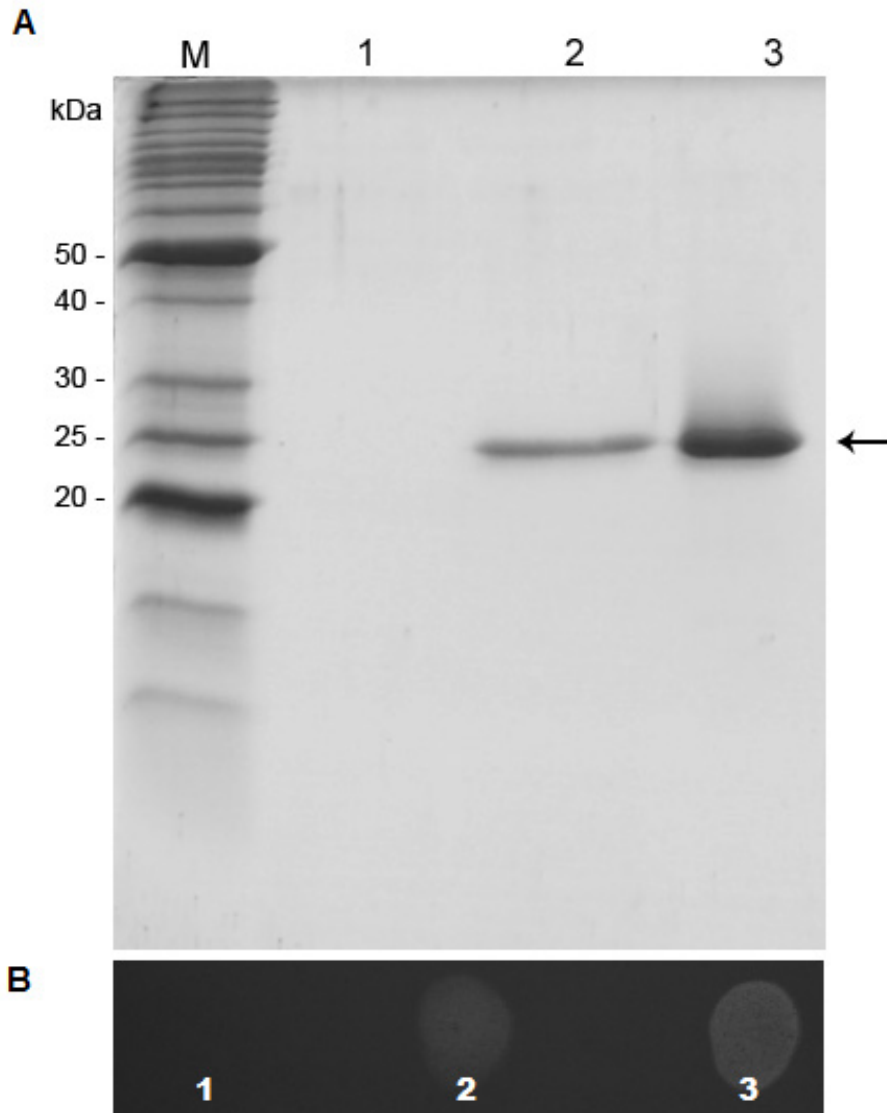
## RESULTS

### Sequence analysis of EGIII in *P. pastoris*

EGIII cDNA is deposited in the GenBank database (accession No. JQ923478) and contains an ORF of 705 bp, which encodes a protein with 235 amino acids. *T. harzianum* EGIII has 81% similarity with the peptide sequence of EGIII from *T. reesei*, 57% similarity with that from *Aspergillus fumigatus* and 47% similarity with that from *Humicola grisea* (for alignments, see Supplementary material). The SignalP program (Bendtsen et al., 2004) predicts a putative signal peptide of 16 amino acids, likely cleaved in the AVA\*QT sequence. The predicted peptide sequence has 56.25% similarity with EGIII from *T. reesei*. No N-glycosylation site was predicted using the NetNGlyc program (<http://www.cbs.dtu.dk/services/NetNGlyc/>).

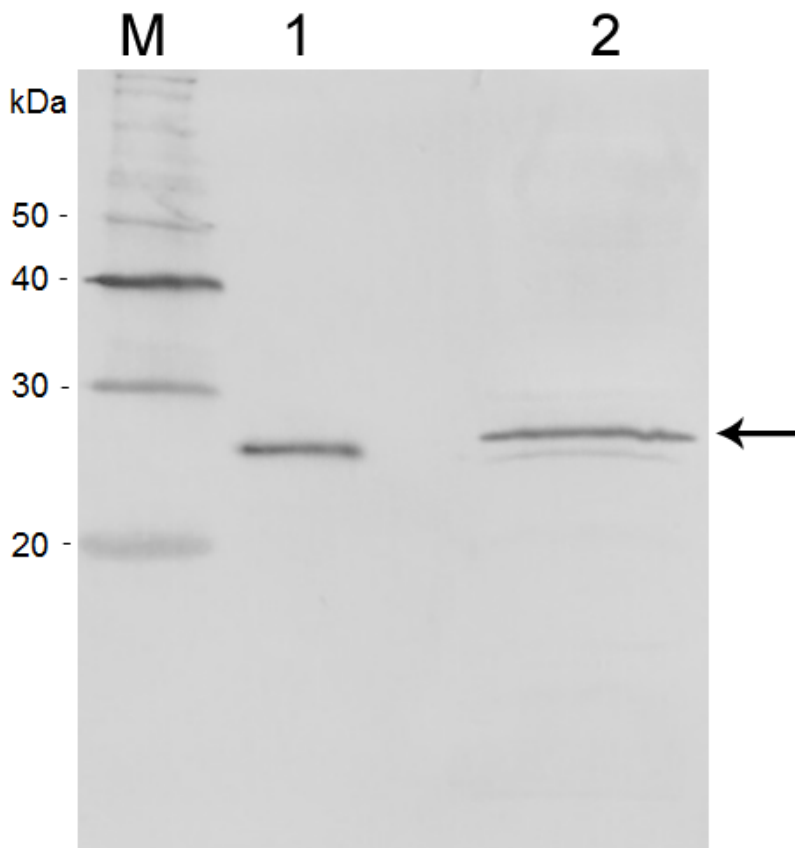
### Expression, purification and identification of EGIII

Several recombinants were analyzed in a small-scale expression assay regarding the production of EGIII (data not shown). The recombinant with the best expression was employed for the production and purification of the recombinant enzyme (Figure 1A). The recombinant enzyme, called rThEGIII, was secreted in an active form by the yeast (Figure 1B). The purification yielded 300 mg enzyme per liter of induced medium.



**Figure 1.** SDS-PAGE analysis of rThEGIII (A) and rThEGIII Congo red assay with carboxymethylcellulose (B). Lane M = BenchMark molecular mass ladder (Invitrogen); lane 1 = supernatant of recombinant *Pichia pastoris* before induction with methanol; lane 2 = supernatant of recombinant *P. pastoris* induced by 1% methanol for 24 h; lane 3 = rThEGIII purified from supernatant. The arrow indicates the recombinant protein.

The purified recombinant enzyme was used to produce polyclonal antibodies, which were used against the supernatant of *T. harzianum* IOC-3844 induced with 1% microcrystalline cellulose. The molecular mass of the enzyme detected in the fungal supernatant was very similar to that of the recombinant enzyme produced in *P. pastoris* (Figure 2).

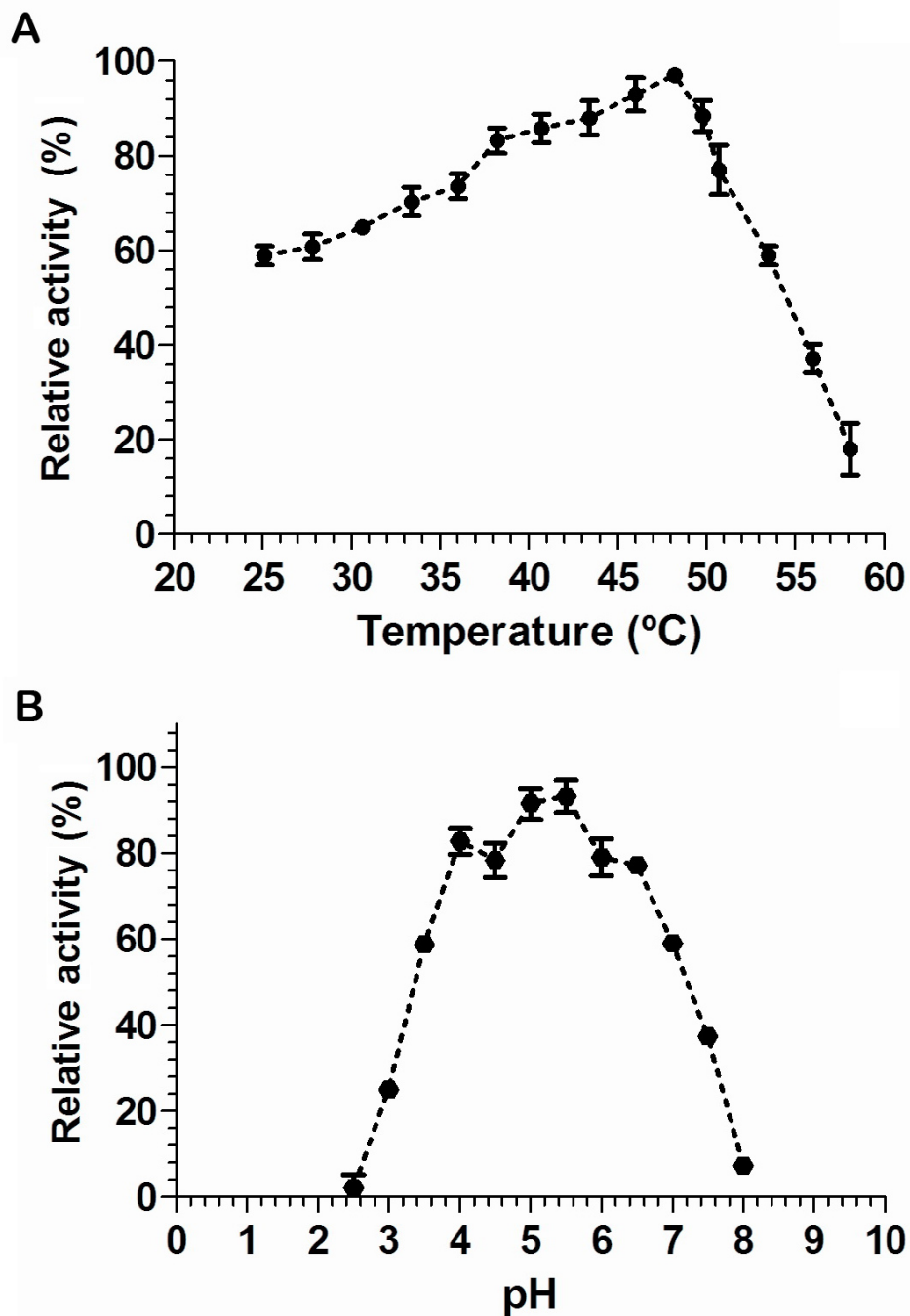


**Figure 2.** Western blotting analysis of EGIII from *Trichoderma harzianum* IOC-3844. Lane M = MagicMark XP Western molecular mass ladder (Invitrogen); lane 1 = purified rThEGIII; lane 2 = supernatant of *T. harzianum* induced with 1% microcrystalline cellulose for 36 h. The arrow indicates the recombinant protein.

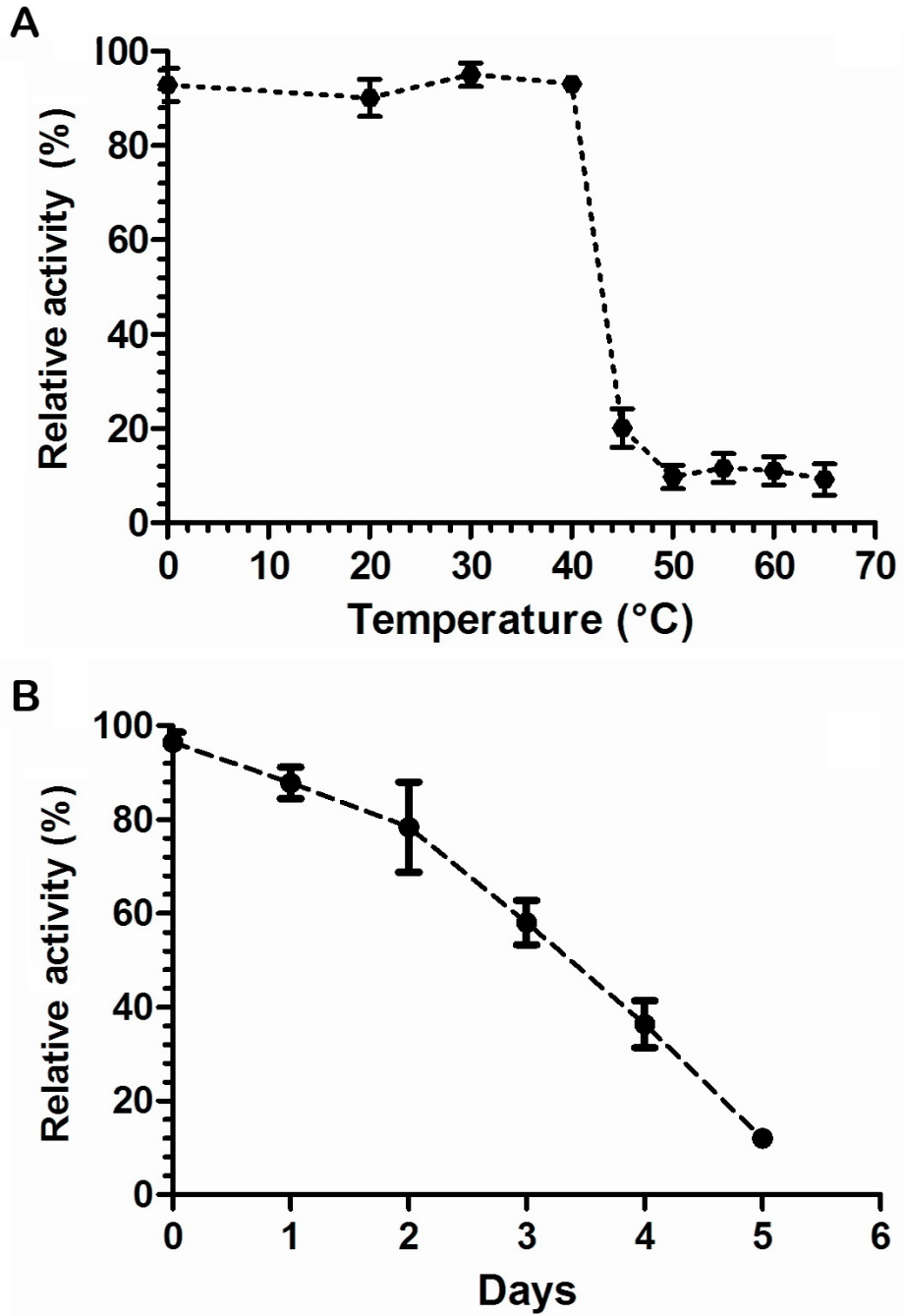
### Characterization of biochemical parameters

For the hydrolytic assays, the enzyme concentration of 0.4  $\mu\text{M}$  was established from previously obtained curves of different enzyme dilutions and reaction times (data not shown). The optimal temperature for rThEGIII was estimated by varying the temperature assays from 21° to 58°C. The best consumption of CMC was achieved at 48.2°C (Figure 3A). Optimal pH was determined within a range of 2.5 to 8.0. The previously determined optimal temperature (48.2°C) was used for the optimal pH tests. The greatest rThEGIII activity was seen at pH 5.5 (Figure 3B).

A thermostability curve was obtained after 1 h of incubation at temperatures ranging from 0° to 65°C. The enzyme was stable up at temperatures up to 40°C (Figure 4A). Considering the decrease in activity over 40°C, enzymatic thermal resistance was evaluated at this temperature for different time intervals. The rThEGIII enzyme exhibited at least 50% activity up to the third day of pre-incubation (Figure 4B).



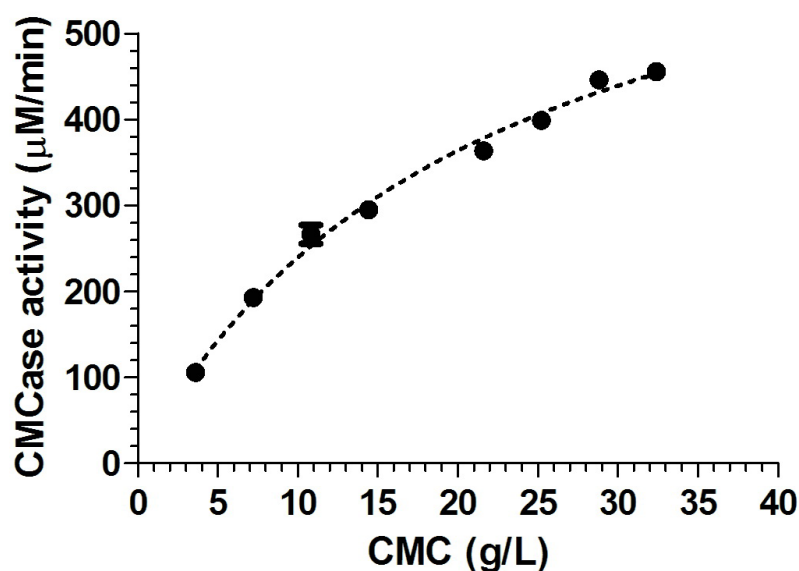
**Figure 3.** Effect of temperature (A) and pH (B) on activity of purified rThEGIII (see Methods for reactions and conditions).



**Figure 4.** Thermostability of rThEGIII after 1 h of incubation (A) and thermal resistance at 40°C (B) (see Methods for reactions and conditions).

### Characterization of enzyme parameters

For the kinetic characterization of rThEGIII, the CMC concentration ranged from 0.3 to 32.4 g/L. The Michaelis-Menten curve fit revealed a  $K_m$  of 21.35 g/L and a turnover number of 1880  $\text{min}^{-1}$ . These data are displayed in Figure 5 and Table 1.



**Figure 5.** Michaelis-Menten curve fit of rate of reduction sugar generated ( $\mu\text{M}/\text{min}$ ) versus carboxymethylcellulose (CMC) concentration (g/L).

**Table 1.** Kinetic enzymatic parameters of rThEGIII.

	$K_m$ (g/L)	$K_{cat}$ ( $\text{min}^{-1}$ )	$K_{cat}/K_m$ ( $\text{L}\cdot\text{g}^{-1}\cdot\text{min}^{-1}$ )
Value	21.35	1880	88.06
Standard deviation	$\pm 1.87$	$\pm 82$	

### DISCUSSION

There is a growing body of research on the isolation and enhancement of enzymes related to the degradation of plant biomass for the development of feasible methods regarding the production of second-generation ethanol. In the present study, a  $\beta$ -1,4-endoglucanase cDNA from the *T. harzianum* strain IOC-3844 was isolated and heterologously expressed in the yeast *P. pastoris*. The recombinant enzyme was characterized considering its possible application in the process of cellulose depolymerization.

The  $\beta$ -1,4-endoglucanase studied has a high degree of similarity with EGIII from *T.*

*reesei*. EGIII from *T. harzianum* belongs to glycosyl hydrolase family 12 and this enzyme can therefore also be described as cel12a (Foreman et al., 2003). Glycosyl hydrolase family 12 comprises hydrolytic enzymes that cleave glycosidic bonds between two carbohydrates or a carbohydrate from a non-carbohydrate moiety. The cellulases from this family have no CBD and do not easily access crystalline cellulose (Macarron et al., 1993; Okada et al., 1998). However, these enzymes contribute to the breakdown of biomass by hydrolyzing amorphous cellulose and acting synergistically with expansin-like proteins (Arantes and Saddler, 2010).

In the present study, the *P. pastoris* system was employed for the heterologous expression of the identified *T. harzianum* EGIII. The use of this yeast host is recommended primarily due to its simplicity and proximal enzymatic machinery in comparison with *T. harzianum*. Thus, the recombinant enzyme could have similar characteristics as the native one. Another advantage in using *P. pastoris* is the low amount of native-secreted proteins (Cereghino and Cregg, 2000), as illustrated in Figure 1A. For the industrial use of rThEGIII, the purification step could be unnecessary. However, due to the low degree of production of native cellulolytic enzymes of *P. pastoris*, the purification step was necessary for the characterization of the recombinant enzyme.

A large amount of rThEGIII was produced after 24 h of methanol induction. In approximately 48 h, about 300 mg purified rThEGIII can be obtained from 1 L of induced medium. This high yield, which can be further increased in industrial fermenters, is important for the competitiveness of biomass ethanol.

The estimated molecular mass for rThEGIII in SDS-PAGE was about 25 kDa, which is very similar to the predicted mass of 24.6 kDa (Figure 1A). This low molecular mass is also reported for cel12a cellulases and could be related to the absence of CBD and N-glycosylation sites, which are also characteristics of cel12a, unlike other endoglucanases with a high molecular mass (Kwon et al., 1999; Sandgren et al., 2005).

The optimal pH and temperature for rThEGIII were 5.5 and 48.2°C, respectively. This optimal pH is quite similar to that reported for EGIII from some filamentous fungi, such as *T. reesei*, *T. koningii*, *Hypocrea schweinitzii* and *H. grisea* (Karlsson et al., 2002; Sandgren et al., 2005; Wang et al., 2005). On the other hand, the optimal temperature for rThEGIII is not similar to that for EGIII from other filamentous fungi. The optimal temperature for EGIII from *H. schweinitzii* and *Gliocladium roseum* is about 47°C, whereas the optimal temperature for *T. reesei*, *T. koningii* and *Fusarium javanicum* is about 55°C, and finally nearly 70°C is reported for *H. grisea* (Karlsson et al., 2002; Lynd et al., 2002; Sandgren et al., 2005; Nakazawa et al., 2009).

Analyzing the crude extract of *T. harzianum* IOC-3844, de Castro et al. (2010a) found a high degree of endoglucanase activity at temperatures around 52°C and pH 5.1. However, at about 48°C (the optimal temperature of rThEGIII), high endoglucanase activity was also detected in the fungal extract (de Castro et al., 2010a). This difference between the crude extract and the purified rThEGIII could be explained by the existence of other endoglucanases and the synergy among them within the fungal culture supernatant. Analyzing an endoglucanase from *T. harzianum* ETS 323 with molecular mass similar to that of rThEGIII, Liu et al. (2010) found an optimal temperature of 50°C, which is very similar to that of rThEGIII.

These optimal features are interesting, since, under suitable conditions of simultaneous saccharification and fermentation, an enzyme with an optimal temperature closer to 37°C and pH 5.5 seems to be more appropriate (Vasquez et al., 2007). Even with a lower degree of



thermostability than the EGIII from *T. reesei*, which displays a high degree of activity until 50°C (Nakazawa et al., 2009), rThEGIII demonstrated stability at temperatures close to the optimal, lasting several days with fairly good activity. Thus, EGIII may be a promising enzyme for the purpose of simultaneous saccharification and fermentation, since this process requires hydrolysis under mild conditions.

Compared to the kinetic parameters reported by Macarron et al. (1993) and Nakazawa et al. (2009) for CMCase activity, the  $K_{cat}$  value obtained here is higher in comparison to EGIII from *T. reesei*, with the  $K_m$  being twice as high. The  $K_m$  is similar to that found for endoglucanase activity in the crude extract of *T. harzianum* IOC-3844 (19 g/L) and the endoglucanase from *T. harzianum* ETS 323 (23 g/L) (de Castro et al., 2010a; Liu et al., 2010). This lower specificity in comparison to *T. reesei* does not preclude the use of this enzyme in saccharification, since a high cellulose concentration is used in this process.

In conclusion, the heterologous production of the  $\beta$ -1,4-endoglucanase III (cel12a) from *T. harzianum* was successfully performed in *P. pastoris*, yielding large amounts of enzyme. The use of yeasts for the production of recombinant enzymes for enzymatic cocktails can be of great value. The primary reason is that there is a low expression of native extracellular proteins by *P. pastoris*, but also, the industrial system for ethanol production is already well suited for the use of yeasts. These results indicate that this enzyme may be a feasible component of commercial enzyme cocktails for both separate and simultaneous hydrolysis and fermentation, mainly due to a high turnover rate. However, further investigations are required to measure the synergy between this enzyme and others from the cellulolytic complex and to evaluate its use on an industrial scale.

## ACKNOWLEDGMENTS

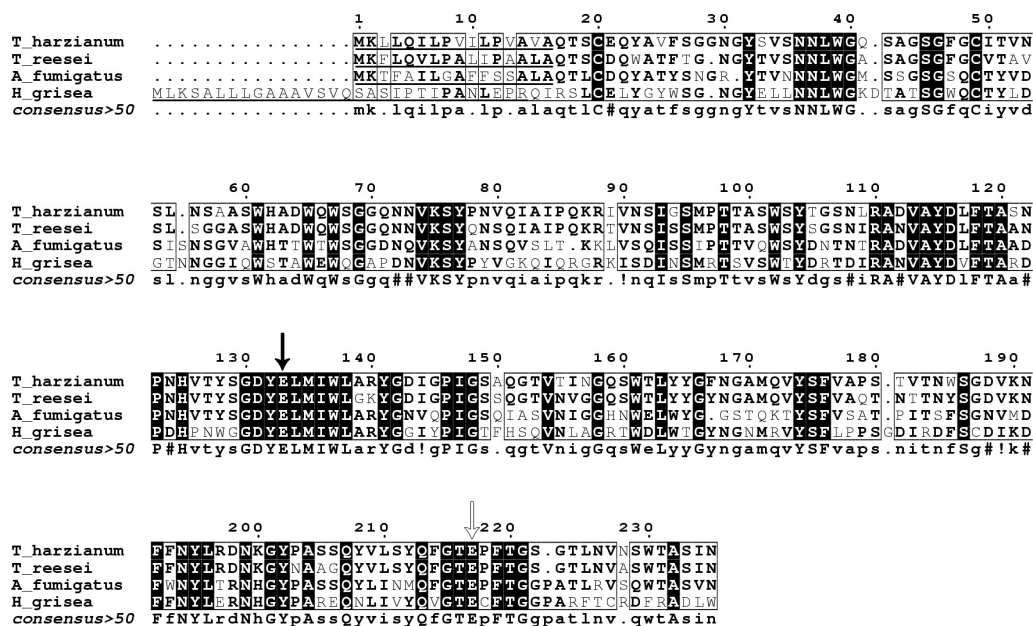
Research supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

## REFERENCES

- Ahmed S, Bashir A, Saleem H, Saadia M, et al. (2009). Production and purification of cellulose-degrading enzymes from a filamentous fungus *Trichoderma harzianum*. *Pak. J. Bot.* 41: 1411-1419.
- Altschul SF, Gish W, Miller W, Myers EW, et al. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
- Arantes V and Saddler JN (2010). Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol. Biofuels* 3: 4.
- Banerjee G, Scott-Craig JS and Walton JD (2010). Improving enzymes for biomass conversion: a basic research perspective. *Bioenerg. Res.* 3: 82-92.
- Bendtsen JD, Nielsen H, von Heijne G and Brunak S (2004). Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340: 783-795.
- Boettner M, Prinz B, Holz C, Stahl U, et al. (2002). High-throughput screening for expression of heterologous proteins in the yeast *Pichia pastoris*. *J. Biotechnol.* 99: 51-62.
- Cereghino JL and Cregg JM (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol. Rev.* 24: 45-66.
- Corpet F (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16: 10881-10890.
- Cregg JM (2007). *Pichia* Protocols: Methods in Molecular Biology. Humana Press, Totowa.
- de Castro AM, Ferreira MC, da Cruz JC, Pedro KC, et al. (2010a). High-yield endoglucanase production by *Trichoderma harzianum* IOC-3844 cultivated in pretreated sugarcane mill byproduct. *Enzyme Res.* 2010: 854526.
- de Castro AM, Pedro KC, da Cruz JC, Ferreira MC, et al. (2010b). *Trichoderma harzianum* IOC-4038: A promising strain for the production of a cellulolytic complex with significant beta-glucosidase activity from sugarcane bagasse

- cellulignin. *Appl. Biochem. Biotechnol.* 162: 2111-2122.
- EasySelect™ Pichia Expression Kit (2001). A Manual of Methods for Expression of Recombinant Proteins Using pPICZ and pPICZα in *Pichia pastoris*. Version G, Invitrogen, Carlsbad.
- Foreman PK, Brown D, Dankmeyer L, Dean R, et al. (2003). Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. *J. Biol. Chem.* 278: 31988-31997.
- Galbe M and Zacchi G (2002). A review of the production of ethanol from softwood. *Appl. Microbiol. Biotechnol.* 59: 618-628.
- Henriksson G, Nutt A, Henriksson H, Pettersson B, et al. (1999). Endoglucanase 28 (Cel12A), a new *Phanerochaete chrysosporium* cellulase. *Eur. J. Biochem.* 259: 88-95.
- Karlsson J, Siika-aho M, Tenkanen M and Tjerneld F (2002). Enzymatic properties of the low molecular mass endoglucanases Cel12A (EG III) and Cel45A (EG V) of *Trichoderma reesei*. *J. Biotechnol.* 99: 63-78.
- Kwon I, Ekino K, Goto M and Furukawa K (1999). Heterologous expression and characterization of endoglucanase I (EGI) from *Trichoderma viride* HK-75. *Biosci. Biotechnol. Biochem.* 63: 1714-1720.
- Liu SY, Shibu MA, Jhan HJ, Lo CT, et al. (2010). Purification and characterization of novel glucanases from *Trichoderma harzianum* ETS 323. *J. Agric. Food Chem.* 58: 10309-10314.
- Lynd LR, Weimer PJ, van Zyl WH and Pretorius IS (2002). Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* 66: 506-577.
- Lynd LR, Laser MS, Bransby D, Dale BE, et al. (2008). How biotech can transform biofuels. *Nat. Biotechnol.* 26: 169-172.
- Macarron R, Acebal C, Castillon MP, Dominguez JM, et al. (1993). Mode of action of endoglucanase III from *Trichoderma reesei*. *Biochem J.* 289: 867-873.
- Mandels M, Parrish FW and Reese ET (1962). Sophorose as an inducer of cellulase in *Trichoderma viride*. *J. Bacteriol.* 83: 400-408.
- Miller GL (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31: 426-428.
- Nakazawa H, Okada K, Onodera T, Ogasawara W, et al. (2009). Directed evolution of endoglucanase III (Cel12A) from *Trichoderma reesei*. *Appl. Microbiol. Biotechnol.* 83: 649-657.
- Okada H, Tada K, Sekiya T, Yokoyama K, et al. (1998). Molecular characterization and heterologous expression of the gene encoding a low-molecular-mass endoglucanase from *Trichoderma reesei* QM9414. *Appl. Environ. Microbiol.* 64: 555-563.
- Pandey A, Soccol CR, Nigam P and Soccol VT (2000). Biotechnological potential of agro-industrial residues. I: sugarcane bagasse. *Bioresour. Technol.* 74: 69-80.
- Pessoa-Jr A, Roberto IC, Menossi M, dos Santos RR, et al. (2005). Perspectives on bioenergy and biotechnology in Brazil. *Appl. Biochem. Biotechnol.* 121-124: 59-70.
- Sambrook J and Russell DW (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
- Sandgren M, Gualfetti PJ, Paech C, Paech S, et al. (2003). The *Humicola grisea* Cel12A enzyme structure at 1.2 Å resolution and the impact of its free cysteine residues on thermal stability. *Protein Sci.* 12: 2782-2793.
- Sandgren M, Stahlberg J and Mitchinson C (2005). Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. *Prog. Biophys. Mol. Biol.* 89: 246-291.
- Schuster A and Schmoll M (2010). Biology and biotechnology of *Trichoderma*. *Appl. Microbiol. Biotechnol.* 87: 787-799.
- Teather RM and Wood PJ (1982). Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Appl. Environ. Microbiol.* 43: 777-780.
- Vasquez MP, da Silva JN, de Souza MJB and Pereira N Jr (2007). Enzymatic hydrolysis optimization to ethanol production by simultaneous saccharification and fermentation. *Appl. Biochem. Biotechnol.* 137-140: 141-153.
- Wang T, Liu X, Yu Q, Zhang X, et al. (2005). Directed evolution for engineering pH profile of endoglucanase III from *Trichoderma reesei*. *Biomol. Eng.* 22: 89-94.

**SUPPLEMENTARY MATERIAL**



**Supplementary Figure 1.** Sequence alignment of glycosyl hydrolase family 12 amino acid sequences from filamentous fungi. Underlined sequences indicate putative signal peptide. Black narrow indicates nucleophile residue and open narrow the acid-base residue. The protein sequence GenBank accession codes were: *Trichoderma harzianum* IOC-3844 (this study), *Trichoderma reesei* QM9414 (BAA20140), *Aspergillus fumigatus* Af293 (XP\_748895), and *Humicola grisea* (AAM77714). The sequences were obtained from GenBank, and the alignment conducted with the MultAlin software (Corpet, 1988).

**Apêndice C: Artigo publicado.**

Marcelo Vizoná Liberato,<sup>a</sup>  
Wesley Cardoso Generoso,<sup>b</sup>  
Wilson Malagó Jr,<sup>b</sup> Flávio  
Henrique-Silva<sup>b</sup> and Igor  
Polikarpov<sup>a\*</sup>

<sup>a</sup>Instituto de Física de São Carlos, Universidade de São Paulo, Avenida Trabalhador São-carlense 400, 13566-590 São Carlos-SP, Brazil, and

<sup>b</sup>Departamento de Genética e Evolução, Universidade Federal de São Carlos, Rodovia Washington Luís, km. 235, 13565-905 São Carlos-SP, Brazil

Correspondence e-mail:  
ipolikarpov@if.sc.usp.br

Received 20 September 2011  
Accepted 9 January 2012

## Crystallization and preliminary X-ray diffraction analysis of endoglucanase III from *Trichoderma harzianum*

Endoglucanases are enzymes that hydrolyze cellulose and are important components of the cellulolytic complex. In contrast to other members of the complex, they cleave internal  $\beta$ -1,4-glycosidic bonds in the cellulose polymer, allowing cellulose to be used as an energy source. Since biomass is an important renewable source of energy, the structural and functional characterization of these enzymes is of interest. In this study, endoglucanase III from *Trichoderma harzianum* was produced in *Pichia pastoris* and purified. Crystals belonging to the orthorhombic space group  $P2_12_12_1$ , with unit-cell parameters  $a = 47.54$ ,  $b = 55.57$ ,  $c = 157.3$  Å, were obtained by the sitting-drop vapour-diffusion method and an X-ray diffraction data set was collected to 2.07 Å resolution.

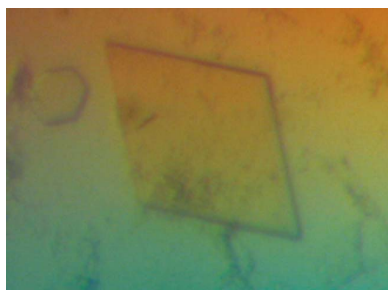
### 1. Introduction

To date, humans have mostly relied on energy provided by fossil fuels, but the increasing energy needs of the growing human population and the recognition of the problems of fossil fuels (pollution, global warming and scarcity, to mention three) have provided a driver for looking at alternative sustainable sources of energy (Singh *et al.*, 2011).

The method of producing energy through the hydrolysis of biomass has been known for a long time (Antoni *et al.*, 2007). Biomass is a renewable resource and the final product of its transformation *via* the fermentation of depolymerized sugars, cellulosic ethanol, is less toxic, more readily biodegradable and produces fewer airborne pollutants than petroleum fuel (John *et al.*, 2011). Although it is not yet possible to fulfil the world's energy demands using renewable fuels, they are increasingly becoming a feasible option.

Cellulose forms tightly interwoven crystalline fibres which are generally very resistant to degradation. However, a number of microorganisms are able to efficiently hydrolyze cellulose. For example, *Hypocrea* fungi in general and *H. jecorina* in particular are able to secrete large amounts of cellulases to hydrolyse cellulosic biomass. Another member of the same group, *H. lixii*, also known as *Trichoderma harzianum*, has been utilized in phytopathogen control (Monteiro *et al.*, 2010) and more recently has been described as a competent cellulolytic fungus which produces a balanced mixture of cellulases (de Castro, Ferreira *et al.*, 2010). It has been demonstrated that *T. harzianum* secretes large quantities of endoglucanases when cultivated on pretreated sugarcane bagasse (de Castro, Pedro *et al.*, 2010).

The cellulolytic complex produced by most fungal species is formed of at least three major components: (i) endoglucanases (EGs), which hydrolyze the internal  $\beta$ -1,4-glycosidic bonds in the cellulose polymer, thus introducing two new internal chain termini; (ii) cellobiohydrolases (CBHs), which act on the reduced and nonreduced termini of cellulose polymer chains, thus degrading cellulose into cellobiose; and (iii)  $\beta$ -glycosidases, which hydrolyze cellobiose to glucose (Medve *et al.*, 1998). A large number of endoglucanases have been identified and they have been classified into glycoside hydrolase families based on their amino-acid sequence similarity (Henrissat & Davies, 2000). Despite the differences in their three-dimensional structures, the catalytic mechanism attributed to all endoglucanases is



similar, *i.e.* they hydrolyze the  $\beta$ -1,4-glycosidic bond in cellulose via a double-displacement reaction and a glycosyl-enzyme intermediate that results in retention of the anomeric configuration in the product (Birsan *et al.*, 1998).

The present paper describes the crystallization of *T. harzianum* endoglucanase III and its preliminary X-ray diffraction studies. The enzyme belongs to the GH12 family, a group of low-molecular-weight endoglucanases that contain a single catalytic domain and lack a cellulose-binding module (Macarrón *et al.*, 1993; Nakazawa *et al.*, 2009). The members of this group have a compact  $\beta$ -sandwich structure that is curved to create an extensive cellulose-binding site on the concave face of the  $\beta$ -sheet (Sulzenbacher *et al.*, 1999). This architecture is based on a compact  $\beta$ -jelly-roll fold which contains one  $\alpha$ -helix and two  $\beta$ -sheets consisting of a total of 15  $\beta$ -strands (Sandgren *et al.*, 2001).

Several crystal structures of endoglucanase III from various organisms have previously been solved (Crennell *et al.*, 2006; Khademi *et al.*, 2002; Sulzenbacher *et al.*, 1997), including one from *T. reesei* (Sandgren *et al.*, 2001, 2004). Despite their strong structural similarity, the endoglucanases reveal considerable differences in their enzymatic activities, biochemical properties and thermal stabilities (Sandgren *et al.*, 2003). Little structural information is currently available for the *T. harzianum* glycosyl hydrolases. To at least partially fill this gap, we expressed, purified and crystallized endoglucanase III (EG3) from *T. harzianum* and collected X-ray diffraction data from the crystals.

## 2. Experimental methods

### 2.1. Protein production

The EG3 cDNA clone was obtained from a cDNA library constructed from *T. harzianum* IOC-3844 induced by microcrystalline cellulose (unpublished data). The EG3 ORF was cloned into pPICZ $\alpha$ A (Invitrogen), excluding the putative signal peptide. The mature ORF was amplified with the forward primer 5'-CGGAA-TTCCAGACCAGCTGCGAAC-3' and the reverse primer 5'-CGC-GTCGACGTTGATAGATGCGGTCC-3' (the bases shown in bold indicate introduced restriction sites).

The amplicon and the pPICZ $\alpha$ A plasmid were digested with *Eco*RI and *Sal*I and ligated, creating a recombinant plasmid containing the mature EG3 ORF flanked by the secretion signal peptide ( $\alpha$ -factor) at the N-terminus and a 6 $\times$ His tag at the C-terminus. The pPICZ vectors also contained an inducible promoter from the alcohol oxidase gene AOX1, which is induced by methanol.

Before transformation, the recombinant plasmid was linearized with *Pme*I endonuclease and then introduced into *Pichia pastoris* (MutS) (Invitrogen) by electroporation (1.5 kV, 25  $\mu$ F, 200  $\Omega$ ). Transformants were cultivated in solid yeast extract peptone dextrose (YEPD) with 1 M sorbitol and 100  $\mu$ g ml<sup>-1</sup> Zeocin. Recombinants were identified by polymerase chain reaction using the  $\alpha$ -factor and AOX 3' primers. For large-scale expression, a recombinant colony was inoculated in 10 ml buffered glycerol-complex medium (BMGY) and incubated at 303 K and 250 rev min<sup>-1</sup> for 24 h. The 10 ml culture was added to 500 ml BMGY and incubated at 303 K and 250 rev min<sup>-1</sup> for 24 h (OD<sub>600nm</sub> = 6). The cells were collected by centrifugation (1500g for 5 min) and the supernatant was discarded. The cells were resuspended in 100 ml buffered methanol-complex medium (BMMY) with 1% methanol and incubated at 303 K and 250 rev min<sup>-1</sup> for 24 h. The culture containing the recombinant protein was centrifuged and the supernatant was vacuum-filtered through a 0.44  $\mu$ m membrane.

### 2.2. Purification

Purification of the recombinant enzyme from the supernatant was performed by affinity chromatography using a 5 ml silica-resin column containing Ni-NTA Superflow resin (Qiagen). The resin was equilibrated with 25 ml binding buffer (10 mM Tris-HCl, 50 mM sodium phosphate, 100 mM sodium chloride pH 8.0). 100 ml of the supernatant was applied to 5 ml resin; after washing the resin with 15 ml binding buffer, the polyhistidine-tagged protein was eluted with 5 ml aliquots of elution buffer containing an increasing imidazole concentration (5–250 mM). The elution fractions were analyzed by 15% SDS-PAGE (Sambrook & Russell, 2001). The fractions containing purified protein were pooled and dialyzed against three changes of 50 mM Tris-HCl buffer pH 7.5 (2 l) at 277 K for 4 h. The protein concentration was determined using the BCA Protein Assay Kit (Thermo Scientific) and the samples were stored at 253 K.

### 2.3. Analysis of the hydrolysis of different substrates

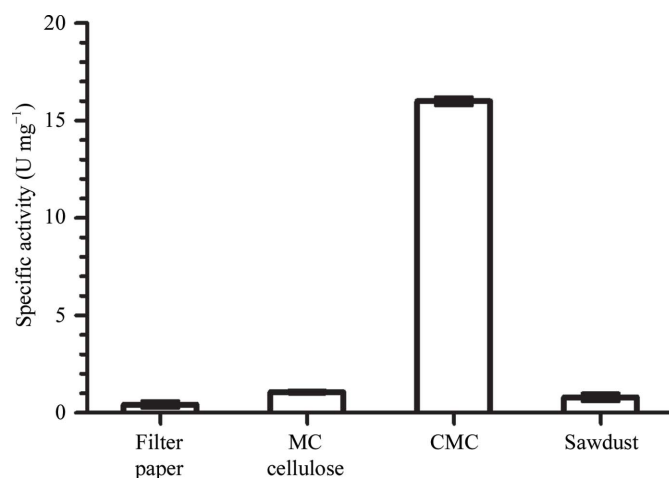
The hydrolytic activity of EG3 was determined using the 3,5-dinitrosalicylic acid method based on quantification of the reducing-sugar content (Miller, 1959). The relative activity of the purified enzyme was checked using filter paper (Whatman, USA), carboxymethylcellulose (50–200 cps; Sigma, Germany), microcrystalline cellulose (Sigma Germany) and sawdust as substrates.

Briefly, 400  $\mu$ l of 2% substrate in 100 mM potassium phosphate buffer pH 6.5 was mixed with 100  $\mu$ l enzyme solution (0.4  $\mu$ M final concentration) and incubated at 315 K. After 15 min, the reaction was stopped by the addition of 1 ml dinitrosalicylic acid solution. The mixture was boiled for 5 min, diluted to 10 ml with water and the liberated reducing sugars were measured at a wavelength of 540 nm using a spectrophotometer (U-5100; Hitachi, Japan).

All assays were performed in triplicate and a reaction stopped just after addition of the enzyme was used as a blank. All experimental data and graphics were generated with the *Prism* 5 statistical program (GraphPad, USA).

### 2.4. Crystallization

Purified EG3 (23 781 Da; amino acids 18–236) at a concentration of 10 mg ml<sup>-1</sup> was submitted to crystallization screening using the sitting-drop vapour-diffusion technique. Drops of 2  $\mu$ l final volume (1:1 ratio of protein and screen solutions) were set up automatically

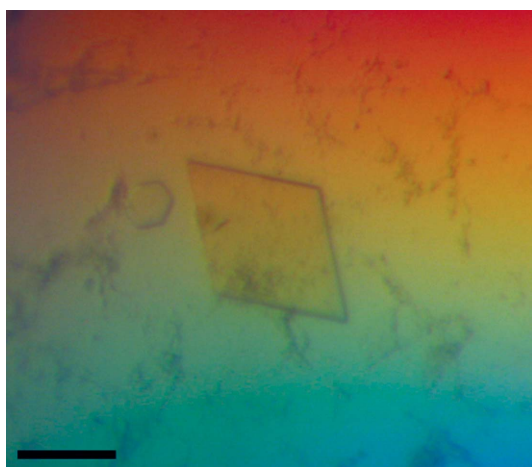


**Figure 1** Substrate specificity of the purified recombinant EG3 from *T. harzianum*. Filter paper, microcrystalline cellulose (MC cellulose), carboxymethylcellulose (CMC) and sawdust were used as substrates.

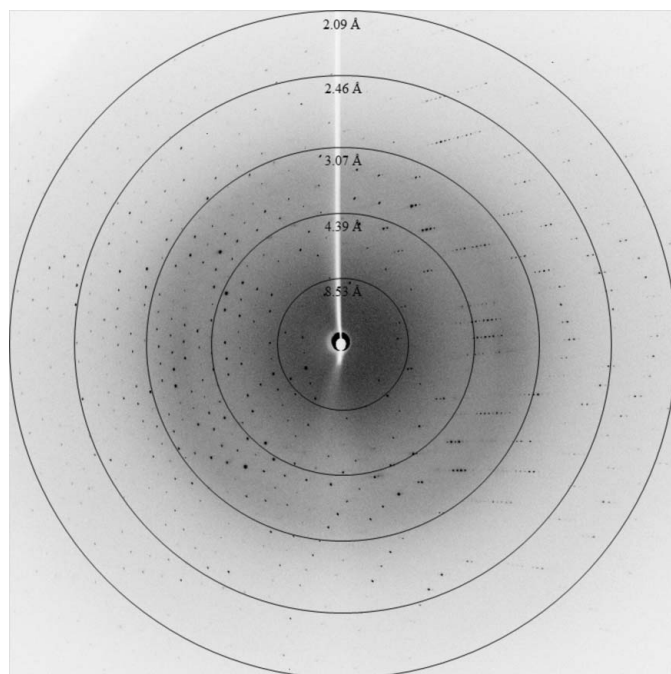
using a Honeybee 931 crystallization robot (Genomic Solutions Inc.) with a variety of commercially available screens (PEGs I and II Suite, Classics Suite and MbClassics Suite from Qiagen) and maintained at a temperature of 291 K. 96-well crystallization plates from Greiner Bio-One were used.

## 2.5. Data collection and processing

A single EG3 crystal was transferred into a cryosolution consisting of 12% PEG 6000, 2 M sodium chloride, 20% ethylene glycol, mounted in a cryoloop and directly flash-cooled in a cold nitrogen stream prior to X-ray data analysis. Diffraction data were collected on beamline MX2 at the Brazilian Synchrotron Light Source (LNLS), Campinas, Brazil using a MAR charge-coupled device detector (Guimarães *et al.*, 2009). Data covering 225° were collected using a 0.5° oscillation step to a maximum resolution of beyond 2.0 Å using a



**Figure 2**  
Crystals of EG3 grown in 25% (w/v) PEG 4000, 0.2 M ammonium sulfate, 0.1 M sodium acetate pH 4.6. The black bar corresponds to 200 µm.



**Figure 3**  
Typical diffraction pattern of an EG3 crystal.

**Table 1**

Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Beamline	MX2, LNLS
Wavelength (Å)	1.46
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 47.54, b = 55.57, c = 157.3$
Resolution range (Å)	50–2.07 (2.14–2.07)
Unique reflections	26039
Total reflections	171836
Multiplicity	6.6 (6.3)
Completeness (%)	99.1 (97.5)
$R_{\text{merge}}^{\dagger}$ (%)	12.1 (21.6)
$\langle I/\sigma(I) \rangle$	14.2 (8.1)

$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$ , where  $I_i(hkl)$  is the intensity of the  $i$ th observation of reflection  $hkl$  and  $\langle I(hkl) \rangle$  is the average intensity of reflection  $hkl$ .

MAR Mosaic225 detector. Data integration was carried out using *DENZO* and data were scaled with *SCALEPACK* in the *HKL-2000* package (Otwinowski, 1993).

## 3. Results and discussion

The purified EG3 was tested for enzymatic activity and proved to be active. The specific activity of the enzyme towards carboxymethyl cellulose (CMC) was 16 U mg<sup>-1</sup>. This activity is comparable with that of *T. reesei* EG3 (Nakazawa *et al.*, 2009). Carboxymethyl cellulose was also preferentially consumed by EG3 compared with filter paper, microcrystalline cellulose and sawdust (Fig. 1).

Crystals of EG3 appeared after 5 d at 291 K in the original sitting-drop screen plates using the PEGs II Suite. A cluster of needles were obtained using a reservoir solution consisting of 25% (w/v) PEG 4000, 0.2 M ammonium sulfate, 0.1 M sodium acetate pH 4.6 (condition No. 65). A plate-shaped crystal (Fig. 2) was subsequently obtained in an optimization screen from a crystallization solution consisting of 12% (w/v) PEG 6000, 2.0 M sodium chloride (condition No. 77).

The plate-shaped crystal was tested for X-ray diffraction quality (Fig. 3) and was used to collect a complete diffraction data set to 2.07 Å resolution. Based on the data-set processing statistics, the EG3 crystal belonged to the primitive orthorhombic system ( $P2_12_12_1$ ), with unit-cell parameters  $a = 47.54, b = 55.57, c = 157.3$  Å. The calculated Matthews coefficient of 2.16 Å<sup>3</sup> Da<sup>-1</sup> (Matthews, 1968) indicated that there were two molecules in the asymmetric unit; the solvent content of the crystal was 43.2%. Data-collection and processing statistics are presented in Table 1.

The structure of EG3 was solved by the molecular-replacement technique with *Phaser* v.2.1.4 (McCoy *et al.*, 2007) using the atomic coordinates of the *T. reesei* EG3 crystal structure (PDB entry 1h8v; Sandgren *et al.*, 2001) as a search model, giving a solution LLG (log-likelihood gain) of 2016.85. The amino-acid sequence identity between the EG3s from *T. reesei* and *T. harzianum* is 83%. Initial structure-refinement cycles using *REFMAC5* (Murshudov *et al.*, 2011) resulted in a model with an  $R_{\text{work}}$  of 26.85% and an  $R_{\text{free}}$  of 30.96%. Further model building and refinement is in progress.

We are grateful to the Brazilian Synchrotron Light Source (LNLS) and to the staff members of the MX2 beamline. This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) via research grants 2009/14333-6, 2008/56255-9 and 2009/52840-7 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) via INCT do Bioetanol and grants 471834/2009-2 and 490022/2009-0.

## References

- Antoni, D., Zverlov, V. V. & Schwarz, W. H. (2007). *Appl. Microbiol. Biotechnol.* **77**, 23–35.
- Birsan, C., Johnson, P., Joshi, M., MacLeod, A., McIntosh, L., Monem, V., Nitz, M., Rose, D. R., Tull, D., Wakarchuck, W. W., Wang, Q., Warren, R. A., White, A. & Withers, S. G. (1998). *Biochem. Soc. Trans.* **26**, 156–160.
- Castro, A. M. de, Ferreira, M. C., da Cruz, J. C., Pedro, K. C., Carvalho, D. F., Leite, S. G. & Pereira, N. (2010). *Enzyme Res.* **2010**, 854526.
- Castro, A. M. de, Pedro, K. C., da Cruz, J. C., Ferreira, M. C., Leite, S. G. & Pereira, N. Jr (2010). *Appl. Biochem. Biotechnol.* **162**, 2111–2122.
- Crennell, S. J., Cook, D., Minns, A., Svergun, D., Andersen, R. L. & Nordberg Karlsson, E. (2006). *J. Mol. Biol.* **356**, 57–71.
- Guimarães, B. G., Sanfelici, L., Neuenschwander, R. T., Rodrigues, F., Grizolli, W. C., Raulik, M. A., Piton, J. R., Meyer, B. C., Nascimento, A. S. & Polikarpov, I. (2009). *J. Synchrotron Rad.* **16**, 69–75.
- Henrissat, B. & Davies, G. J. (2000). *Plant Physiol.* **124**, 1515–1519.
- John, R. P., Anisha, G. S., Nampoothiri, K. M. & Pandey, A. (2011). *Bioresour. Technol.* **102**, 186–193.
- Khademi, S., Zhang, D., Swanson, S. M., Wartenberg, A., Witte, K. & Meyer, E. F. (2002). *Acta Cryst.* **D58**, 660–667.
- Macarrón, R., Acebal, C., Castellón, M. P., Domínguez, J. M., de la Mata, I., Pettersson, G., Tomme, P. & Claeysens, M. (1993). *Biochem. J.* **289**, 867–873.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Medve, J., Karlsson, J., Lee, D. & Tjerneld, F. (1998). *Biotechnol. Bioeng.* **59**, 621–634.
- Miller, G. L. (1959). *Anal. Chem.* **31**, 426–428.
- Monteiro, V. N., do Nascimento Silva, R., Steindorff, A. S., Costa, F. T., Noronha, E. F., Ricart, C. A., de Sousa, M. V., Vainstein, M. H. & Ulhoa, C. J. (2010). *Curr. Microbiol.* **61**, 298–305.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Nakazawa, H., Okada, K., Onodera, T., Ogasawara, W., Okada, H. & Morikawa, Y. (2009). *Appl. Microbiol. Biotechnol.* **83**, 649–657.
- Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 55–62. Warrington: Daresbury Laboratory.
- Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning. A Laboratory Manual*, 3rd ed. New York: Cold Spring Harbor Laboratory Press.
- Sandgren, M., Berglund, G. I., Shaw, A., Ståhlberg, J., Kenne, L., Desmet, T. & Mitchinson, C. (2004). *J. Mol. Biol.* **342**, 1505–1517.
- Sandgren, M., Gualfetti, P. J., Shaw, A., Gross, L. S., Saldajeno, M., Day, A. G., Jones, T. A. & Mitchinson, C. (2003). *Protein Sci.* **12**, 848–860.
- Sandgren, M., Shaw, A., Ropp, T. H., Wu, S., Bott, R., Cameron, A. D., Ståhlberg, J., Mitchinson, C. & Jones, T. A. (2001). *J. Mol. Biol.* **308**, 295–310.
- Singh, A., Nigam, P. S. & Murphy, J. D. (2011). *Bioresour. Technol.* **102**, 10–16.
- Sulzenbacher, G., Mackenzie, L. F., Wilson, K. S., Withers, S. G., Dupont, C. & Davies, G. J. (1999). *Biochemistry*, **38**, 4826–4833.
- Sulzenbacher, G., Shareck, F., Morosoli, R., Dupont, C. & Davies, G. J. (1997). *Biochemistry*, **36**, 16032–16039.



**Apêndice D: Artigo publicado.**

## Laboratory Exercises

# DNA as Genetic Material: Revisiting Classic Experiments Through a Simple, Practical Class<sup>§</sup>

Received for publication, February 20, 2009, and in revised form, June 17, 2009

Wilson Malagó Jr., Andrea Soares-Costa, and Flávio Henrique-Silva‡

From the Department of Genetics and Evolution, Federal University of São Carlos, Rodovia Washington Luis, Km 235, São Carlos, Sao Paulo, Brazil

In 1928, Frederick Griffith demonstrated a transmission process of genetic information by transforming *Pneumococcus*. In 1944, Avery *et al.* demonstrated that Griffith's transforming principle was DNA. We revisited these classic experiments in a practical class for undergraduate students. Both experiments were reproduced in simple, adapted forms. Griffith's experiment was reproduced by mixing heat-killed, ampicillin-resistant *E. coli* with live ampicillin-susceptible *E. coli*, followed by plating samples in the presence or absence of the antibiotic. Cells were also plated separately as controls. Avery's work was reproduced by treating a purified plasmid harboring the ampicillin resistance gene with DNase I. Treated and untreated plasmids were then used to transform *E. coli* cells, which were plated in culture media containing ampicillin. The students received a class guide for understanding and performing the experiments. The original articles by Griffith and Avery *et al.* were also provided, along with a list of questions to encourage a discussion on the experimental approach and results. The expected results were obtained and the students successfully revisited the classic experiments, which revealed that DNA is genetic material. The class was very well accepted, as indicated by students' evaluations. Thus, we presented a quick, inexpensive class involving important concepts, which can be easily reproduced in any laboratory with minor resources.

**Keywords:** Genetic material, Transforming principle, Frederick Griffith, Oswald Avery.

At the beginning of the 20th century, with the rediscovery of Mendel's works, the blending theory was abandoned, as inheritance patterns were attributed to hereditary particles. As reviewed by Martins in 1999 [1], the evidence published by Walter Sutton and Theodor Boveri between 1902 and 1903 contributed to the belief that the behavior of Mendel's particles parallels the movements of chromosomes in meiosis, suggesting that genes are physically located on the chromosomes. These studies originated the chromosome theory of inheritance, which was supported in 1905 by Brush and Stevens [2] and Wilson [3], with their studies involving the behavior of sex chromosomes. Additional evidence appeared in the second decade of 20th century, with the linkage studies carried out by Thomas Hunt Morgan and his students. As reported by Neatherall [4], in 1908 Archibald Garrod speculated that some diseases were the consequence of a lack of enzyme function resulting from defective Mendelian factors (i.e.

genes). This hypothesis was later confirmed in 1941 by Beadle and Tatum [5], who worked with mutant strains of *Neurospora crassa* lacking different enzyme activities involved in the arginine biosynthesis pathway. These authors postulated that genes mediate biochemical reactions through the production of specific enzymes. Thus, Garrod's hypothesis was confirmed and stated as "one gene/one enzyme." However, among the evidence indicating the gene location and mechanisms of gene action, an intriguing question remained to be answered in the scientific context of the time: What is the molecular nature of genes?

In 1928, primarily interested in the cause and control of human pneumonia, Frederick Griffith demonstrated the transmission of genetic information [6]. When studying the occurrence of different types of *Pneumococci* found in pneumonia patients, Griffith isolated and produced *in vitro* a naturally occurring strain of a nonvirulent *Pneumococcus*. These nonvirulent cells had a rough appearance (R cells) due to the lack of the outer polysaccharide capsule. Contrarily, virulent cells had a smooth appearance (S cells) due to the presence of the outer polysaccharide capsule. Griffith observed that stable strains of R cells could be regularly transformed into S cells when mixing heat-killed S cells with live R cells before injecting them

<sup>§</sup>Additional Supporting Information may be found in the online version of this article.

‡To whom correspondence should be addressed. Tel: 55-16-3351-8378; Fax: 55-16-3351-8377. E-mail: dfhs@power.ufscar.br.

into mice. When the mice died from pneumonia, the S cells could be recovered from them. Griffith believed that some S substance, likely protein, was passing from S to R cells during the transformation.

Over the following years, the chemical nature of “Griffith’s transforming principle” was pursued by a research team headed by Oswald Avery. Their belief was that understanding how capsule production was activated by the transforming principle would possibly lead to pneumonia therapy based on the inhibition of this process. In 1931, Dawson and Sia [7] brought Griffith’s transformation out of the mouse and into the test tube. In 1932, Alloway [8] eliminated the need for intact heat-killed S cells by preparing cell-free active filtered extracts. In subsequent years, Colin Munro MacLeod removed protein from the extracts using chloroform, with no loss of transforming activity. MacLeod also found the presence of RNA and DNA in the active extract. Maclyn McCarty used specific enzymes and demonstrated that polysaccharide and RNA did not need to be present in the principle for transformation to occur. The team also showed that the active substance of the extract was not digested by trypsin or chymotrypsin. Moreover, lipids were removed from extracts without impairing the transforming activity. At the time, it was well known that chromosomes were mainly composed of DNA and protein and the researchers gradually considered DNA to be genetic material [9]. The final proof came when Avery’s team showed that enzyme preparations enriched with DNase destroyed the transformation activity. Finally, in 1944, Avery et al. reported their evidence indicating that DNA was the transforming principle and responsible for inducing enzymatic polysaccharide syntheses in *Pneumococci* S cells [10]. Despite Avery’s evidence, many would only accept DNA as genetic material after experiments published in 1952 by Hershey and Chase, who recovered labeled phage DNA from the bacterial cell fraction following virus infection and demonstrated that DNA was required for new phages to reproduce inside the bacterial cell [11]. Nevertheless, according to McCarty, some scientists promptly accepted the evidences presented in the Avery’s article published in 1944. In fact, important findings were the consequence of studies started before the Hershey and Chase’s publication in 1952. For example, Erwing Chargaff turned all of his efforts to the study of nucleic acids, after reading the Avery’s 1944 article. Hence, when searching for the biological specificity of DNA, Chargaff discovered the phenomena of base complementarity. Furthermore, the information in the Avery’s article served as the basis for James Watson and Francis Crick to begin their studies on DNA structure in 1951 [9].

The educational approach proposed here directs students to the important historical context of Molecular Biology involving the identification of the genetic material. We have adapted simple experiments that parallel the works of Griffith and Avery *et al.*, revealing DNA as genetic material. The students’ thoughts are turned to that era by dealing with equivalent evidence through adapted, up-to-date experiments. Moreover, the students deal with the phenomenon of bacterial transformation, which is central to understanding the usefulness of DNA manipulation methods.

TABLE I  
*Parallelism between Griffith’s work and our adapted experiment*

<i>Pneumococcus</i> cell	Phenotype	<i>E. coli</i> cell	Phenotype
S live	Mouse dies	AR live	Grow in Amp
R live	Mouse lives	AS live	Do not grow in Amp
S heat-killed	Mouse lives	AR boiled	Do not grow anyway
R live + S heat-killed	Mouse dies	AS live + AR boiled	Grow in Amp

In the week before the class, we scheduled a half hour meeting in which the original articles by Griffith and Avery *et al.* were given to the students, along with a brief explanation of the main historical points, emphasizing central concepts and asking them to try to understand the context and relevance of the studies. We recommended that they perform a selective reading of the main points of the articles that may be useful for the equivalent experiments to be carried out the following week.

#### METHODS

The class reported here is the first of a series of four classes on “Topics of Molecular Genetics.” This series is part of an elective, essentially practical discipline denominated “Practices of Genetics,” offered to the students of Biological Sciences undergraduate course. The team that administered the series of four classes was composed of two graduate students—researchers at the Molecular Biology Laboratory—and the head of the Laboratory, who was the professor in charge. The applicants had to be familiar with laboratory practices. Therefore, they would have at least completed the related disciplines of Cell Biology and Microbiology, which have significant number of practical activities. Among the theoretical disciplines, they had to have at least completed Classical Genetics.

At the beginning of the class, the main concepts involving the experiments were explained and important details regarding the protocols were discussed. The students received a class guide, which is available on the web [12], containing brief theoretical explanations and detailed protocols. Class activities were accomplished in 4 hours, from 2 to 6 p.m. Fifteen students were equally divided into three groups. All were continually encouraged to participate actively in the manipulation.

Griffith’s experiment was simulated by transforming ampicillin-susceptible DH5 $\alpha$  *E. coli* cells (AS<sup>1</sup> cells) into ampicillin-resistant cells (AR cells) through the transmission of the plasmid pTZ57R (Fermentas, Glen Burnie, MD, K1214), which harbors the ampicillin resistance gene. Transformation was achieved by mixing boiled heat-killed AR cells with live AS cells before plating. A parallel between Griffith’s main results and our adaptation is summarized in Table I.

Avery’s DNA isolation can be compared with *E. coli* standard plasmid preparations, as both eliminate unwanted cellular components, thereby yielding purified DNA. Thus, we gave a brief explanation of both preparations to the students and simulated Avery’s procedures by beginning from the key point in which the DNA was submitted to DNase digestion. We prepared CaCl<sub>2</sub>-competent DH5 $\alpha$  *E. coli* cells during the class to transform with the plasmids treated and untreated with DNase. Regarding the features of the plasmid used, we only explained

<sup>1</sup>The abbreviations used are: AS cells, Ampicillin Susceptible cells; AR cells, Ampicillin Resistant cells; +Amp, With Ampicillin; –Amp, Without Ampicillin; LB, Luria Broth culture media; RPM, Revolutions per Minute; CFU, Colony forming units.

TABLE II  
List of questions given

Question	Score
1) What is the nature of the transforming principle and what are the experimental evidences which indicates the nature of it?	8.0
2) Why did we also plate AR live, AS live, and AR boiled cells individually in agar culture media +Amp and –Amp?	9.0
3) Why did we plate the AS live + AR boiled cell mixture in agar media –Amp?	7.0
4) Why did we plate transforming cells with DNA not digested with DNase I?	9.0
5) What is the importance of the transformation phenomenon to the Molecular Biology?	8.0
6) What could explain the presence of some colonies in the plates “AR boiled cells” –Amp and +Amp of group two?	10.0
7) What could explain the absence of colonies in the plate “DNase I untreated plasmid” of group one?	9.0

The average scores are ranging from 1 to 10.

that it carried a gene capable of changing the growth phenotype of *E. coli*.

The following week, we scheduled a 2-hour class to analyze the results. A discussion was stimulated and the students answered an individual questionnaire (Table II). The students' evaluations were determined based on the questionnaire and individual participation in class. The students also filled out an additional question form with scores ranging from 1 to 5 providing an evaluation on the overall quality of the class (Table III).

#### MATERIALS

Apart from the laboratory equipment, all the materials used were previously prepared and/or allocated in triplicate to facilitate the work of the three groups (Table IV). The concentration and amount of reagents used are mentioned in “Protocols.”

#### EXPERIMENTAL ASPECTS AND PROTOCOLS

Some useful information to help with the protocol can be accessed on the web from “<http://www.ncbe.reading.ac.uk/ncbe/protocols/transformer.html>.” A general guide for the appropriate use of micropipettes can be downloaded at “<http://www.gilson.com/ServiceTraining/pipeUsersGuides.asp>.” However, some protocol innovations are only now being made public in this article.

Although the DH5 $\alpha$  *E. coli* cells used are not potentially pathogenic, the students were advised to take precautions during handling as done in previous practical microbiological disciplines. After class, the culture media containing *E. coli* cells and all contaminated materials were sterilized before disposal.

The protocols were all previously tested under the same conditions of the classroom experiments and produced the desired results. On the day before class, AR cells (DH5 $\alpha$  *E. coli* containing the pTZ57R plasmid – Fermentas, Glen Burnie, MD, K1214) and AS cells (DH5 $\alpha$  *E. coli* without plasmid) were preinoculated in 10 mL of LB Medium in the presence and absence of ampicillin, respectively. The cells grew overnight at 37 °C. The next morning, the AR +Amp and AS –Amp preinocula were diluted at 1:50 in 15 mL of LB with and without Amp, respectively. The samples were incubated for approximately 3 hours until Optical

Density achieved 0.4 at 600 nm. Each type of cell was divided into three samples of 3 mL for the different groups and kept in ice until the class began. We also separated two aliquots of 1 mL of AS cells for each group for the competent cell preparation. The LB agar plates were also prepared before the class. There were 6 +Amp and 4 –Amp plates per group. Each group began the adapted Griffith procedures by boiling 200  $\mu$ L of AR cells for 10 min. The boiled AR cells were placed in ice for 2 min and 100  $\mu$ L were mixed with 100  $\mu$ L of live AS cells. The mixture was kept in ice for 5 min and then incubated for 50 min at 37 °C and 300 RPM in Eppendorf, Westbury, NY, Thermomixer. Aliquots of 200  $\mu$ L of live AR live AS and boiled AR cells were incubated separately for 50 min at 37 °C and 300 RPM. After incubation, the samples were then kept in ice until plating. Meanwhile, the students began to prepare DH5 $\alpha$  *E. coli* competent cells. Each group centrifuged two aliquots of 1 mL of AS cells for 7 min at 1,500  $\times$  g in a microcentrifuge. The supernatant was removed and cells were resuspended in 350  $\mu$ L of iced 0.1 M CaCl<sub>2</sub>. The samples were centrifuged for 7 min at 1,150  $\times$  g and the supernatant was removed. The cells were resuspended in 100  $\mu$ L of iced 0.1 M CaCl<sub>2</sub> and kept in ice until use. Each group received two 1.5 tubes, containing 200 ng of pTZ57R (Fermentas, Glen Burnie, MD, K1214) in 2  $\mu$ L of water. Thus, the adapted Avery experiments were begun by adding to

TABLE III  
Evaluation of the class

Topics	Average
Evaluation of the class	
1) Relevance of the topics in molecular biology	4.9
2) Up-to-date nature of the class content	4.3
3) The class stimulated the student's commitment to do the experiments	4.2
4) Quality of the knowledge acquired by carrying out the experiments	4.5
5) Motivation to check out the results	4.4
6) Pleasure in carrying out the experiments	4.3
7) Were your expectations fulfilled?	4.6
8) Proximity of the experimental steps to the reality of a research laboratory	4.4
9) Quality of the class	4.7
10) Organization in terms of the theoretical approach	3.8
11) Organization in terms of the practical approach	3.7
12) Overall evaluation	4.5
Evaluation of the teachers	
1) Organization and development of the theoretical content	4.0
2) Organization and development of the practical content	4.1
3) Technical explanations about the experimental steps	4.1
4) Availability/receptivity to explanations/clarifications during experimental steps	4.8
5) Stimulation of student participation	4.3
6) Quality/clarity of the explanations	4.6
7) Preparation of the experimental material	4.1
8) Conduct when faced with unexpected incidents	4.4
9) General orientation	4.5
10) Overall evaluation of the teachers	4.6
Self-evaluation	
1) Participation in the experiments	4.3
2) Theoretical knowledge acquired	4.2
3) Practical knowledge acquired	4.3
4) Overall self-evaluation	4.3

The questionnaire was answered by all of the students and the numbers indicate the average score attributed to each answer, on a scale of 1 to 5.

TABLE IV  
List of required material

Reagents and molecular device
Culture medium (LB/LB agar)
Ampicillin for plasmid selection
<i>E. coli</i> DH5 $\alpha$ cells
Crushed ice
DNase I
Plasmid with ampicillin resistance gene
EDTA
Distilled water
Sterile solution of CaCl <sub>2</sub> 0.1 M
100% and 70% Ethanol
Equipments and laboratory stuff
Microcentrifuge
Timer
Test tube rack
Permanent marker
Beaker for waste/used tips
Autoclave
1.5 mL microcentrifuge tubes
Water boiling apparatus
Thermomixer block or Orbital shaker 37 °C
Water bath 42 °C
Micropipettes and tips
Laminar flow hood (optional)
50 ml polypropylene tubes
Petri plates
Bunsen burner
Spectrophotometer
Cell spreader
Incubator 37 °C

one of the tubes: 6.5 of deionized water, 1  $\mu$ L of DNase I buffer, and 0.5  $\mu$ L of DNase I (Invitrogen, Grand Island, NY, #18068-015). The samples were incubated for 15 min at 37 °C and the reaction was inactivated by adding 1  $\mu$ L of 25 mM EDTA and incubating for 10 min at 65 °C. The tubes were kept in ice until the transformation. Amounts of 200 ng of plasmid DNAs treated and untreated with DNase I were used to transform the freshly prepared *E. coli* cells. The DNAs were added to the competent cell aliquots, gently mixed and kept in ice for 10 min. Heat shock was performed by placing the tubes for 90 seconds in a water bath at 42 °C. The tubes were placed in ice again for 1 min and 900  $\mu$ L of LB was then added to the samples, which were incubated for 45 min at 37 °C and 300 RPM. Meanwhile, the samples from the adapted Griffith protocol were arranged and plated in LB agar media in the presence and absence of ampicillin as follows: 50  $\mu$ L of live AR cells, 50  $\mu$ L of live AS cells, 50  $\mu$ L of boiled AR cells, and 50  $\mu$ L of the mixture of boiled AR cells plus live AS cells. Finally, 50  $\mu$ L of transformed cells with plasmid DNA treated and untreated with DNase I were plated in LB agar media containing ampicillin.

## RESULTS

### Experimental Results

Bacterial growth 1 day before class was as expected: AR cells grew in the presence and absence of ampicillin and AS cells grew only in the absence of ampicillin. This excluded any possibility of contamination or misuse of bacterial cells with undesired phenotypes. Groups 1 and 3 obtained all the expected Griffith parallel results. Figure 1 shows the plates from Group 3. The results include the following: live AR cell growth in the presence and absence of ampicillin; live AS cell growth only in absence of ampicillin; lack growth of boiled AR cells in the presence and absence of ampicillin; and growth of boiled AR mixed with live AS cells in the absence and presence of ampicillin. Growth of the mixture of boiled AR and live

AS cells in the presence of ampicillin demonstrated that the growth phenotype in the presence of the antibiotic was passed from AR to AS cells. Relatively lesser growth in plates +Amp indicated that only a small proportion of cells were transformed (Fig. 1*h*). Among the plates from Group 2 there were unexpected results: a colony (which seemed to be yeast) grew on the boiled AR –Amp plate and two *E. coli* colonies grew on the boiled AR +Amp plate (data not shown). Before the class, we also plated boiled AS cells, which did not grow. However, we did

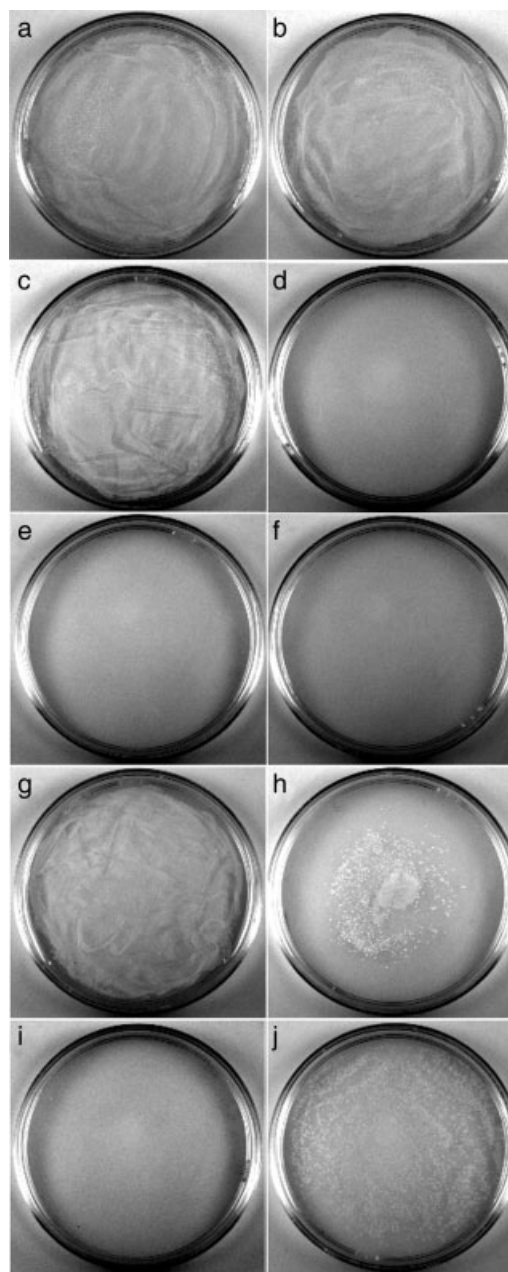


FIG. 1. Plating results of Griffith's (a–h) and Avery's (i, j) adapted experiments. (a) AR live cells (–Amp); (b) AR live cells (+Amp); (c) AS live cells (–Amp); (d) AS live cells (+Amp); (e) AR boiled cells (–Amp); (f) AR boiled cells (+Amp); (g) mixture of AR boiled cells plus AS live cells (–Amp); (h) mixture of AR boiled cells plus AS live cells (+Amp); (i) DNase I treated plasmid transformation (+Amp); (j) DNase I untreated plasmid transformation (+Amp).

not find it necessary to perform this test in class, as AR boiled plating results are enough to control the effectiveness of inactivation by boiling.

The simulated Avery DNase I digestion experiments from Groups 2 and 3 achieved the expected results, as can be seen in the plates from Group 3 (Fig. 1). DNase-treated plasmid did not transform the cells, but untreated plasmid did (Figs. 1*i* and 1*j*). This indicates that DNA is responsible for the phenotype of ampicillin resistance. Group 1 plates revealed unexpected results: two *E. coli* colonies grew on the plate treated with DNase I and no colonies grew on the untreated plasmid plates (data not shown). The students counted the plates from Groups 2 and 3 to calculate the competence of the DH5 $\alpha$  *E. coli* cells, which was  $2.5 \times 10^5$  CFU per microgram of pTZ57R (Fermentas, K1214) plasmid (Fig. 1*j*).

### Learning Results

The questionnaire answers showed that the students' understanding was generally good. The average scores are listed in Table II and the comments for each question are as follows: 1) Nearly all correctly referred to DNase I assay results as the main evidence of the nature of the transforming principle. One student answered: "Without DNA, there is no transforming principle". Interesting, two students also referred to evidence generated from the parallel Griffith experiment and claimed that nucleic acid resistance to boiling temperatures also indicates DNA as the transforming principle. Two others incorrectly answered that the evidence came from the digestion of heat-killed AR cells with DNase I. These answers may reflect a lack of attention. 2) Despite a few incomplete citations of detailed aspects involving the required plating controls, the students generally understood the results and the usefulness of the additional plating. Interestingly, one student answered "No growth was expected among heat-killed AR cells plates. However, the transforming principle was there, but unviable, due to its dependence on the cellular machinery inactivated by heat." 3) Most of students referred to the mixture of killed AR and live AS cells plated in  $-$ Amp plates as an important control for determining the viability of live AS cells after the mixing. However, only five stated that the low percentage transformed was due to the low efficiency of live AS cells. Therefore, in a future class, it would be useful to include a question such as "What may explain the difference between the growth in plates G and F?" to raise this issue for discussion. 4) The majority of students answered that undigested plasmid plating was an important control to determine the viability of competent cells, which, in turn, ensured the result of plasmid exposure to DNase I. The same two students who wrongly answered Question 1 stated that plating was useful in monitoring DNase I activity. 5) The majority understood the examples discussed after the experiment and stated that the acquirement of new characters by target organisms is a phenomenon of great importance to fundamental and applied research. 6) All of students mentioned the possibility of contamination to explain the unexpected

results in Group 2. However, three students also stated the possibility of some sort of boiling inactivation failure in the case of the presence of *E. coli* (which is certainly less probable). 7) All members in Group 1 stated that vigorous handling could have been responsible for the damage of competent cells. Other students stated that forceful handling combined with a high temperature of the cell spreader could culminate in the death of competent cells.

### Discipline Achievement

We evaluated class participation in a comparative manner, based on the observation of student's feedback whereas stimulated by us to handling the samples and to ask about the protocols and concepts involved in the experiments. The scores given ranged from 1 to 10. Some students did not participate equally in the experiments. In fact, two of them less participated and stated that this was because they were unfamiliar with the equipment. We believe that this represented a lack of interest and they receive a participation score of 6. By the other hand, nine students were fully motivated. They were always ready to handling the samples and were also continually asking about the protocol steps and theoretical concepts. This lead us to give a score 10 for them. The remaining four students had an intermediate participation and received a score of 8.

The overall evaluation of the class itself (score ranging from 1 to 5) was determined using a questionnaire administered to the students which included an assessment of the class, teachers and a self-evaluation. All students filled out the form, clearly indicating their approval of the class. Among the averages calculated, only two scores were below 4.0 (Table III).

### DISCUSSION

The results of the parallel Avery experiments with DNase I were undoubtedly definitive toward reinforcing the evidence that DNA is genetic material (Fig. 1*i*). However, the students were astonished with the transformation results obtained in the Griffith simulation (Fig. 1*h*). Some said: "Oh! It transformed!" They suspected that the transforming principle might be a plasmid harboring the ampicillin resistance gene. We confirmed their suspicions, informing them that it was the same plasmid used in the Avery simulation experiment. It would be perhaps useful to divide the class and perform the Avery simulation only after the observation of the results of Griffith simulation, thereby avoiding presenting the plasmid and allowing the students to consider the possibilities.

During the discussion of the results, the students could realize that the phenomena of transformation can occur in a very simple way, also being able to occur in natural environment, as it depends roughly on the contact of the host cell with some exposed DNA. This learning did not arise in previous classes in which the students simply transformed *E. coli* cells following the established methods and also did not simulate both Griffith's and Avery's experiments. In that cases the students had the idea that

the transformation only could occur in acutely controlled laboratory conditions, with appropriate prepared DNA and cells.

In previous experiments, we also achieved a transformation by mixing the purified plasmid with live AS cells (not CaCl<sub>2</sub>-competent cells). This was not performed in class due to time restraints, but could be employed in future classes. Moreover, we attempted to digest heat-killed AR cells with DNase I, but it did not work, possibly due to the presence of some DNase I-inhibition activity within the cellular debris.

The competent cell preparation was included in the class because DNA transfer is one of the most important tools in molecular biology [13]. Furthermore, among the existing methods, *E. coli* transformation was the first to be performed and is of central importance. During the postexperiment discussion, questions arose regarding the use of competent cells. We emphasized that, with the intention of setting up constant transformation assays, Avery's team was the first to optimize and prepare what we currently call "competent cells" [9]. We briefly also explained that cDNA libraries require a high number of clones to be representative and, thus, highly efficient competent cells are needed.

Despite the fact that AS cells did not grow in +Amp media during the preparation for the class, it would be useful to include a preparation of an extra tube of competent cells in order to plate them directly (without transformation) in +Amp plates, thereby demonstrating their inability to grow in the presence of ampicillin.

Some students asked about the applications of the DNase I enzyme. After emphasizing its usefulness in Avery's final proof, we briefly explained its use in RT-PCR and DNase fingerprint assays.

Interestingly, during the experiment and subsequent discussion, some students' thoughts, beliefs, and expressions had to be redressed. Regarding contamination, one student stated that, "Some DNA from the air could transform the cells." Another said: "The plating controls were to monitor potentially strange bodies." There was also: "The principle was passed to the AS cells, transforming them into virulent-resistant cells."

The overall evaluation survey demonstrated the effectiveness of the class (Table III). One student commented: "This was a simple and very interesting class." Only two elements had scores below 4.0. The theoretical approach was criticized. This may be because some students may be not accustomed to reading scientific articles, which may have led to an inappropriate reading of the original weighty articles provided before the class. Moreover, some students appeared to be in a rush to perform the experiments and possibly did not read the theoretical

explanations in the class guide with full attention. Nonetheless, during the discussion of the results, they were able to understand the concepts fully. The practical approach was also criticized. This may be due to the optimized program of the experiment tasks, in which the experiments were overlapped and organized in such a sequence as to adapt to the available time (i.e. Griffith's simulation, competent cells, DNase I digestion, DNase plasmid transformation, adapted Griffith plating, and adapted Avery plating). This may have puzzled some students, but also could have been solved with a detailed reading of the class guide.

#### CONCLUDING REMARKS

We proposed a cheap, quick, and innovative practical class, revisiting Griffith's and Avery's classic experiments, which demonstrated that DNA is genetic material. Central concepts and important tools of molecular biology were employed. The students were placed in contact with important historical issues and laboratory training. Furthermore, the class can be easily reproduced in a regular laboratory with minor resources.

#### REFERENCES

- [1] L. A.-C. P. Martins (1999) Did Sutton and Boveri propose the so-called Sutton-Boveri chromosome hypothesis? *Genet. Mol. Biol.* **22**, 261–272.
- [2] S. G. Brush (1978) Nettie M. Stevens and the discovery of sex determination by chromosomes, *Isis*, **69**, 163–172.
- [3] E. B. Wilson (1905) The chromosomes in relation to the determination of sex in insects, *Science* **22**, 500–502.
- [4] D. J. Weatherall (2008) The centenary of Garrod's Croonian lectures, *Clin. Med.* **8**, 309–311.
- [5] G. W. Beadle, E. L. Tatum (1941) Genetic control of biochemical reactions in neurospora, *Proc. Natl. Acad. Sci. USA* **27**, 499–506.
- [6] F. Griffith (1928) The significance of pneumococcal types, *J. Hyg.* **27**, 113–159.
- [7] M. H. Dawson, R. H. P. Sia (1931) In vitro transformations of pneumococcal types I. A technique for inducing transformation of pneumococcal types in vitro, *J. Exp. Med.* **54**, 681–700.
- [8] J. L. Alloway (1932) The transformation in vitro of *R. pneumococci* into S forms of different specific types by the use of filtered pneumococcus extracts, *J. Exp. Med.* **55**, 91–99.
- [9] M. McCarty (1985) *The Transforming Principle—Discovering that Genes are Made of DNA*, W.W. Norton & Company, New York.
- [10] O. T. Avery, C. M. MacLeod, M. McCarty (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III, *J. Exp. Med.* **79**, 137–158.
- [11] A. D. Hershey, M. Chase (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage, *J. Gen. Physiol.* **36**, 39–56.
- [12] Federal University of Sao Carlos, Molecular Biology Laboratory, Available at [http://www.ufscar.br/~dge/Class\\_Guide\\_BAMBED\\_2009.pdf](http://www.ufscar.br/~dge/Class_Guide_BAMBED_2009.pdf).
- [13] J. D. Watson, A. A. Caudy, R. M. Myers, J. A. Witkowski (2007) *Recombinant DNA Genes and Genomes—A Short Course*, 3rd ed., W.H. Freeman and Company, New York.

## Apêndice E: Objetivos e cronograma propostos.

Cronograma inicialmente proposto no projeto:

	Sem 2/2008	Sem 1/2009	Sem 2/2009	Sem 1/2010	Sem 2/2010	Sem 1/2011	Sem 2/2011	Sem 1/2012
<b>Objetivo específico 1</b>	X	X	X					
<b>Objetivo específico 2</b>		X	X	X				
<b>Objetivo específico 3</b>			X	X	X	X		
<b>Objetivo específico 4</b>				X	X	X	X	
<b>Objetivo específico 5</b>			X	X	X	X		
<b>Redação de Tese</b>								X

Cumprimento do cronograma :

Objetivos específicos propostos no projeto	Objetivo alcançado (%)				
1-Construir uma biblioteca de cDNA de <i>Trichoderma harzianum</i> induzido para a produção de celulases e seqüenciar os clones.	20 %	40 %	60 %	80 %	100 %
2- Identificar e caracterizar os genes que codificam para celulases por meio de comparação com celulases conhecidas de outros organismos.	20 %	40 %	60 %	80 %	100 %
3-Isolar e clonar os genes que codificam para as celulases de <i>Trichoderma harzianum</i> .	20 %	40 %	60 %	80 %	100 %
4-Desenvolver sistemas de expressão e purificação das celulases do <i>Trichoderma harzianum</i> em <i>Pichia pastoris</i> .	20 %	40 %	60 %	80 %	100 %
5-Estudar o transcriptoma de <i>Trichoderma harzianum</i> induzido para a produção de celulases comparando com o transcriptoma do fungo não induzido.	20 %	40 %	60 %	80 %	100 %

O Objetivo específico 5 foi considerado 60 % executado pois, apesar de terem sido feitos estudos de PCR em tempo real com alguns genes alvo (*cbh1*, *cbh2*, *egl1*, *egl2*, *egl3*, *egl7* e *swol*) representativos das celulases, todo o transcriptoma induzido do *T. harzianum* IOC-3844 poderia ter sido analisado em comparação com transcriptomas não induzidos deste fungo disponíveis publicamente. Este trabalho foi iniciado mas as ferramentas e a metodologia envolvendo bioinformática apresentaram problemas que comprometeram a execução do em tempo hábil.



## Apêndice F: Panorama dos estudos de clonagem e expressão heteróloga.

Estão resumidos abaixo os ensaios realizados no Laboratório de Biologia Molecular do DGE-UFSCar, envolvendo as ORFs de *Trichoderma harzianum* IOC-3844 relacionadas a desconstrução de biomassa.

Proteína	ORF Completa em pDONR	ORF Seqüenciada	Clonada em Plasmídeo	Expressão	Expressou	Ativa
CBHI	OK	OK	OK		Não	-
CBHII	OK	OK	-		-	-
EGI	OK	OK	OK		OK	OK
EGII	OK	-	-		-	-
EGIII	OK	OK	OK		OK	OK
EGIV	OK	OK	OK		OK	OK
EGV	OK	OK	-		-	-
EGVII	OK	OK	OK		Não	-
BGL1	OK	OK	OK		-	-
BGL2	OK	OK	OK		Não	-
BGL5	OK	-	-		-	-
Xyn2	OK	OK	OK		OK	OK
Xyn3	OK	OK	OK		OK	OK
Bxl1	OK	-	-		-	-
Man1	OK	-	-		-	-
SWOI	OK	OK	-		-	-
Cip1	14° aa*	-	-		-	-
Cip2	OK	-	-		-	-
Xyr1	OK	-	-		-	-
ACEI	6° aa*	-	-		-	-

\*Amino-ácido a partir do qual o alinhamento entre o cDNA e o melhor “hit” do blastx se inicia.

**Apêndice G: Alinhamentos entre as enzimas seqüenciadas de *Trichoderma harzianum* IOC-3844 e as correspondentes homólogas de *Trichoderma reesei*.**

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/ BLAST/ blastp suite-2sequences/ Formatting Results - TUMAGYGN113](#)[Formatting options](#)[Download](#)

## Blast 2 sequences

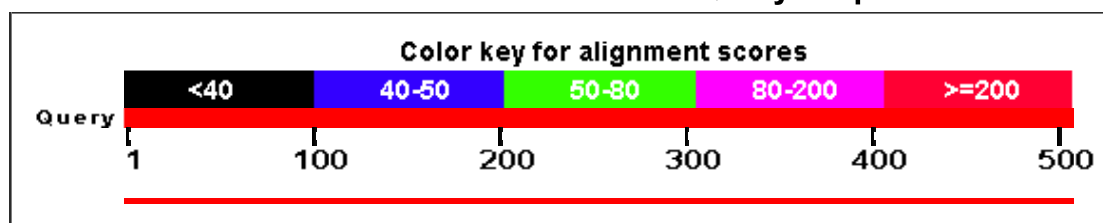
## CBHI T. harzianum IOC-3844 (505 letters)

**Query ID** |c|26373  
**Description** CBHI T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 505

**Subject ID** 26375  
**Description** CBHI T. reesei AAQ76092.1  
**Molecule type** amino acid  
**Subject Length** 514  
**Program** BLASTP 2.2.26+

[Graphic Summary](#)

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
26375	CBHI T. reesei AAQ76092.1	<u>822</u>	822	100%	0.0	80%	

## Alignments

>lcl|26375 CBHI T. reesei AAQ76092.1  
Length=514

Score = 822 bits (2123), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 412/514 (80%), Positives = 452/514 (88%), Gaps = 9/514 (2%)

Query	1	MYQKLA AISAF LAAARAQQVCTQQAETHPPLTWQKCTASG-CTPQQG SVVLDANWRWTHD	59
Sbjct	1	MY+KLA ISAF LA ARAQ CT Q+ETHPPLTWQK++ G CT Q GSVV+DANWRWTH	60
Query	60	TKSTTNCYDGNTWSSTLCPDDATCAKNCCLDGANYSGTYGVTTSGDALTLQFVTAS---N	116
Sbjct	61	TNSSTNCYDGNTWSSTLCPDNETCAKNCCLDGAAYASTYGVTTSGNLSLSIGFVTQSAQKN	120
Query	117	VGSRLYLMANDSTYQEF T LSGNEFSFDVDVSQLPCGLNGALYFVSM DADGGQSKYPGNAA	176
Sbjct	121	VGARLYLMA SDT TYQEF T L LSGNEFSFDVDVSQLPCGLNGALYFVSM DADGGVSKYPTNTA	180
Query	177	GAKYGTGYCDSQCPRDLKFINGQANVEGWEPSSNNANTGVGGHGSCCSEMDIWEANSISE	236
Sbjct	181	GAKYGTGYCDSQCPRDLKFINGQANVEGWEPSSNNANTG+GGHGSCCSEMDIWEANSISE	240
Query	237	ALTPHPCETV GQ T MCGSDSCGGTYSNDRYGGTCDPDGCDWNPYRLGNTSFYGP GSSSFALD	296
Sbjct	241	ALTPHPCTTVGQEICEGDGCGGTYSNDRYGGTCDPDGCDWDPYRLGNTSFYGP GSSSF TLD	300
Query	297	TTKKLTVVTQFATDGSISRYYVQNGVKFQQPN AQVGSYSGNTINTDYCAA EQTAFGGT SF	356
Sbjct	301	TTKKLTVVTQFETSGAINRYYVQNGVTFQQPNAELGSYSGNGLNDDYCTAEAEFGGSSF	360
Query	357	TDKGGLAQINKAFQGGMVLVMSLWDDYAVNMLWLDSTYPTNATASTPGAKRGSCSTSSGV	416
Sbjct	361	SDKGGLTQFKKATSGGMVLVMSLWDDYYANMLWLDSTYPTNETSSTPGAVRGSCSTSSGV	420
Query	417	PAQVEAQSPNSKVIYSNIRFGPIGSTG-----GNTGSNPPGTSTTRAPPSSTGSSPTATQ	471
Sbjct	421	PAQVESQSPNAKVTF SNIKFGPIGSTGDPSSGGNPPGGNPPGTTTTTRRPATTTGSSPGPTQ	480
Query	472	THYGQCGGTGWTGPTRCASGYTCQVLNPFYSQCL	505
Sbjct	481	SHYGQCGGIGYSGPTVCASGTTTCQVLNPFYSQCL	514

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/BLAST/blastp\\_suite-2sequences/Formatting Results - TUR543VG112](#)[Formatting options](#)[Download](#)

## Blast 2 sequences

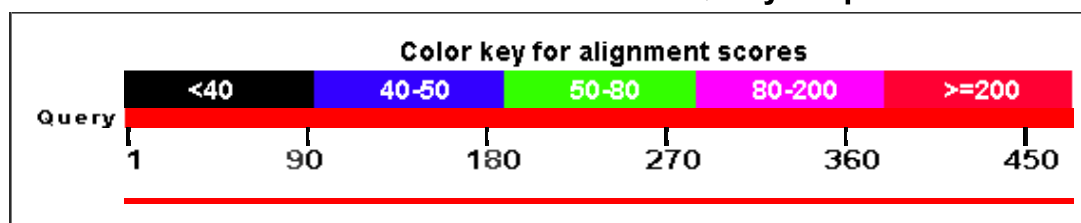
## CBHII T. harzianum IOC-3844 (472 letters)

**Query ID** lc|48055  
**Description** CBHII T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 472

**Subject ID** 48057  
**Description** CBHII T. reesei AAA34210.1  
**Molecule type** amino acid  
**Subject Length** 471  
**Program** BLASTP 2.2.26+

[Graphic Summary](#)

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
48057	CBHII T. reesei AAA34210.1	<a href="#">805</a>	805	100%	0.0	90%	

## Alignments

>lcl|48057 CBHII T. reesei AAA34210.1  
Length=471

Score = 805 bits (2079), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 424/472 (90%), Positives = 446/472 (94%), Gaps = 1/472 (0%)

Query	1	MVVGILATLATLATLAASVPLEERQSCSSVWVGQCGGQNWAGPFCCASGSTCVYSNDYYSQ	60
Sbjct	1	M+VGIL TLATLATLAASVPLEERQ+CSSVWVGQCGGQNW+GP CCASGSTCVYSNDYYSQ	60
Query	61	CLPGAASSSSSTRASSTTSRVSSATSTRSSASTPPPASSTTPAPPVGSATYSGNPFAG	120
Sbjct	61	CLPGAASSSSSTRAASTTSRVSPPTS-RSSSATPPPGSTTTTRVPPVGSATYSGNPFVFG	119
Query	121	VTPWANSFYASEVSTLAIPLSLTGAMATAAAAVAKVPSFMWLDTLDKTPLMSSTLSDIRAA	180
Sbjct	120	VTPWAN++YASEVS+LAIPLSLTGAMATAAAAVAKVPSFMWLDTLDKTPLM TL+DIR A	179
Query	181	NKAGGNYAGQFVVYDLPDRDCAAAAASNGEYSIADGGVDKYKNYIDTIRGIVTTFS DVRIL	240
Sbjct	180	NKNGGNYAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTL	239
Query	241	LVIEPDSLANLVTNLATPKCSNAQSAYLECIYAITQLNLPNVAMYLDAGHAGWLGWPAN	300
Sbjct	240	LVIEPDSLANLVTNL TPKC+NAQSAYLECIYAVTQLNLPNVAMYLDAGHAGWLGWPAN	299
Query	301	QDPAAQLFANVYKNASSPRAVRGLATNVANYNAWNIITPPSYTQGNNAVYNEKLYIHAIGP	360
Sbjct	300	QDPAAQLFANVYKNASSPRALRGLATNVANYNGWNIITPPSYTQGNNAVYNEKLYIHAIGP	359
Query	361	LLANHGWSNAFFITDQGRSGKOPTGQLEWGNWCNAVGTGFGIRPSANTGDSLLDSFVWVK	420
Sbjct	360	LLANHGWSNAFFITDQGRSGKOPTGQ +WG+WCN +GTGFGIRPSANTGDSLLDSFVWVK	419
Query	421	PGGECDGTSNSSAPRFDYHCASADALQPAPQAGSWFQAYFVQLLTNANPSFL	472
Sbjct	420	PGGECDGTS+SSAPRFD HCA DALQPAPQAG+WFQAYFVQLLTNANPSFL	471

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/BLAST/blastp\\_suite-2sequences/Formatting Results - TURU929W113](#)

[Formatting options](#)

[Download](#)

## Blast 2 sequences

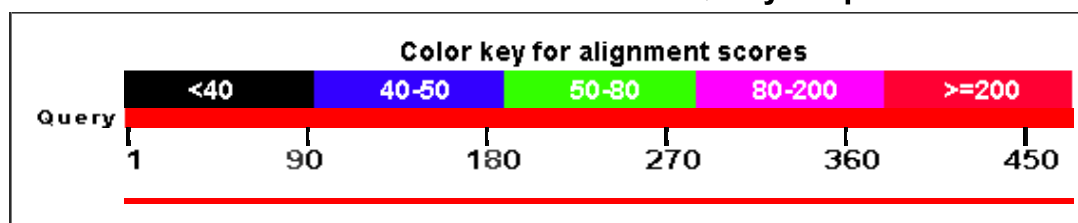
## EGI T. harzianum IOC-3844

**Query ID** |c|56343  
**Description** EGI T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 471

**Subject ID** 56345  
**Description** EGI T. reesei AAA34212.1  
**Molecule type** amino acid  
**Subject Length** 459  
**Program** BLASTP 2.2.26+

## Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence



# Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
56345	EGI T. reesei AAA34212.1	<a href="#">719</a>	719	100%	0.0	81%	

## Alignments

>lcl|56345 EGI T. reesei AAA34212.1  
Length=459

Score = 719 bits (1857), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 380/471 (81%), Positives = 412/471 (87%), Gaps = 12/471 (3%)

Query	1	MALSGPFPPLATAALIAFIAQIVAGQQPGTSTPEVHPKLT	60
		MA S PL TA L IA++VA QQP	
Sbjct	1	MAPSVTLPLTTAILA--IARLVAAQQPGTSTPEVHPKLT	58
Query	61	NYRWMHDKNFNSCTVNGGVNTTLC	120
		NYRWMHD N+NSCTVNGGVNTTLC	
Sbjct	59	NYRWMHDANYNNSCTVNGGVNTTLC	118
Query	121	SSSGGYSSVSPRLYLLGSDGDYELLQ	180
		SSSGGYSSVSPRLYLL SDG+Y +L+LNGQELS	
Sbjct	119	SSSGGYSSVSPRLYLLSDGGEYVMLK	178
Query	181	NQYNTAGANYGSGYCD	240
		NQYNTAGANYGSGYCD	
Sbjct	179	NQYNTAGANYGSGYCD	238
Query	241	TACDASGCGFN	300
		TACD++GCGFN	
Sbjct	239	TACDSAGCGFN	298
Query	301	GVDVPSAQSGGDTISSCPSASAYGGL	360
		GVD+PSAQ GGD	
Sbjct	299	GVDIPSAQPGGDTISSCPSASAYGGL	358
Query	361	PCSSTEGNPSTIVANNPGTHVIFSNIRWGDIGST	420
		PCSSTEGNPS I+ANNP THV+FSNIRWGDIGST	
Sbjct	359	PCSSTEGNPSNILANNPNTHVVFSNIRWGDIGST	408
Query	421	TSTSSRTTSSAAGCTQTHWGQCGNGYTGCKICTAGTTCQYSNDYYSQCL	471
		++T +TTSS+ CTQTHWGQCGG GY+GCK CT+GTTCQYSNDYYSQCL	
Sbjct	409	STTRSSTTSSSPSCTQTHWGQCGGIGYSGCKTCTSGTTCQYSNDYYSQCL	459



## BLAST®

## Basic Local Alignment Search Tool

[NCBI/ BLAST/ blastp suite-2sequences/ Formatting Results - TUS1S5K111R](#)

[Formatting options](#)

[Download](#)

## Blast 2 sequences

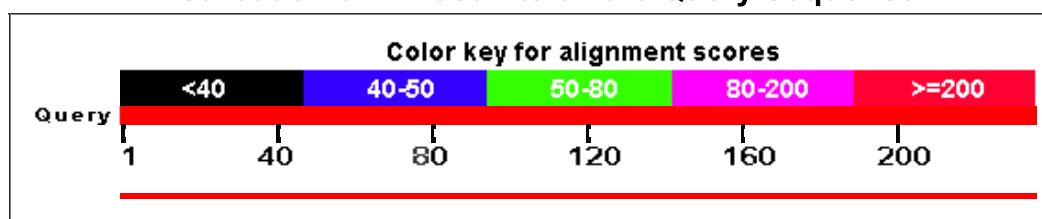
## EGIII T. harzianum IOC-3844

**Query ID** |cl|51303  
**Description** EGIII T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 235

**Subject ID** 51305  
**Description** EGIII T. reesei BAA20140.1  
**Molecule type** amino acid  
**Subject Length** 234  
**Program** BLASTP 2.2.26+

Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
51305	EGIII T. reesei BAA20140.1	<a href="#">395</a>	395	100%	1e-144	81%	

## Alignments

>lcl|51305 EGIII T. reesei BAA20140.1  
Length=234

Score = 395 bits (1015), Expect = 1e-144, Method: Compositional matrix adjust.  
Identities = 191/235 (81%), Positives = 212/235 (90%), Gaps = 1/235 (0%)

Query	1	MKLLQILPVILPVAVAQTSCEQYAVFSGGNGYSVSNLWQSQAGSGFGCITVNSLNSAAS	60
Sbjct	1	MK LQ+LP ++P A+AQTSC+Q+A F+G NGY+VSNNLWG SAGSGFGC+T SL+ AS	59
Query	61	WHADWQWSGGQNNVKSYPNVQIAIPQKRIVNSIGSMPTTASWSYTGSNLRADVAYDLFTA	120
Sbjct	60	WHADWQWSGGQNNVKSYPNVQIAIPQKR VNSI SMPTTASWSY+GSN+RA+VAYDLFTA	119
Query	121	SNPNHVTYSGDYELMIWLARYGDIGPIGSAQGTVTINGQSWTLYYGFNGAMQVYSFVAPS	180
Sbjct	120	ANPNHVTYSGDYELMIWL +YGDIGPIGS+QGTV + GQSWTLYYG+NGAMQVYSFVA +	179
Query	181	TVTNSGDKVKNFFNYLRDNKGYPASSQYVLSYQFGTEPFTGSGTLNVNSWTASIN	235
Sbjct	180	TN+SGDKVKNFFNYLRDNKGY A+ QYVLSYQFGTEPFTGSGTLNV SWTASIN	234

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/BLAST/blastp\\_suite-2sequences/Formatting Results - TUS9YET3113](#)

[Formatting options](#)

[Download](#)

## Blast 2 sequences

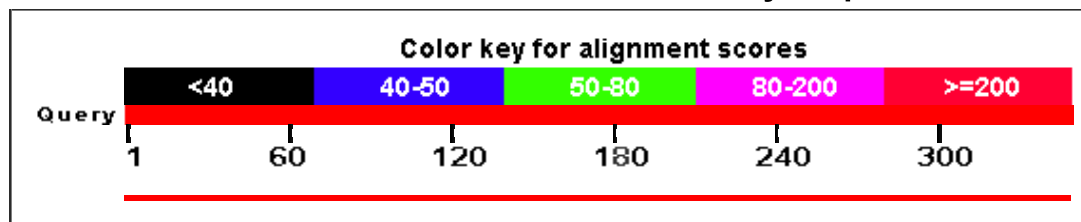
## EGIV T. harzianum IOC-3844

**Query ID** |c|25571  
**Description** EGIV T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 348

**Subject ID** 25573  
**Description** EGIV T. reesei CAA71999.1  
**Molecule type** amino acid  
**Subject Length** 344  
**Program** BLASTP 2.2.26+

Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [PubChem](#) [BioAssay](#)

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
25573	EGIV T. reesei CAA71999.1	<a href="#">534</a>	534	99%	0.0	76%	

## Alignments

>lcl|25573 EGIV T. reesei CAA71999.1  
Length=344

Score = 534 bits (1375), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 264/347 (76%), Positives = 309/347 (89%), Gaps = 3/347 (1%)

Query	1	MAQKLSNLLATALTVATGVVGHGHVNNIIVNGVYQGYDPTSFPYMPNPPIVIGWTAADL	60
		M QKLSNLL TAL VATGVVGHGH+N+I++NGV+YQ YDPT+FPY NPPIV+GWTAADL	
Sbjct	1	MIQKLSNLLVTALAVATGVVGHGHINDIVINGVWYQAYDPTTFPYESNPPIVVGWTAADL	60
Query	61	DNGFVAPDAYQGPDI VCHKNATNAKGHASVKAGDSILLQWLPLPWPHPGPIVDYLANCNG	120
		DNGFV+PDAYQ PDI+CHKNATNAKGHASVKAGD+IL QW+P+PWPHPGPIVDYLANCNG	
Sbjct	61	DNGFVSPDAYQNPDI ICHKNATNAKGHASVKAGDTILFQWVPVPWPHPGPIVDYLANCNG	120
Query	121	DCETVDKTTLEFFKIDGVGLISGADPGNWAQNVLVTNNNTWVVQIPKDLQGTGNYVLRHEL	180
		DCETVDKTTLEFFKIDGVGL+SG DPG WA +VL++NNNTWVV+IP +L GNYVLRHE+	
Sbjct	121	DCETVDKTTLEFFKIDGVGLLSGGDPGTWASDVLI SNNNTWVVKIPDNLAPGNYVLRHEI	180
Query	181	IALHSSGQPDGAQNYPQCFNLAVTGTGSLQPTGVLATDLYHESDPGILVNIYTSSLAYTM	240
		IALHS+GQ +GAQNYPQCFN+AV+G+GSLQP+GVL TDLYH +DPG+L+NIYTS L Y +	
Sbjct	181	IALHSAGQANGAQNYPQCFNIAVSGSGSLQPSGVLGTDLYHATDPGVLINIYTSPLNYII	240
Query	241	PGPSLVAGLPSSVAQGN SAATATSSATIPGGGGSSSSSRASSASATKTSSGATSRPTSSSI	300
		PGP++V+GLP+SVAQG+SAATAT+SAT+PGGG S S + T ++ A+SRP+S+	
Sbjct	241	PGPTVVSGLPTSVAQGSSAATATASATVPGGG---SGPTSRTTTTARTTQASSRPSSTPP	297
Query	301	RTTSAPAGGPTQTLYGQCGGSGYSGPTICASPAVCTTLNPYYAQCVS	347
		TTSAPAGGPTQTLYGQCGGSGYSGPT CA PA C+TLNPYYAQC++	
Sbjct	298	ATTSAPAGGPTQTLYGQCGGSGYSGPTRCAPPATCSTLNPYYAQCCLN	344

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/ BLAST/ blastp suite-2sequences/ Formatting Results - TUSEUK19113](#)

[Formatting options](#)

[Download](#)

## Blast 2 sequences

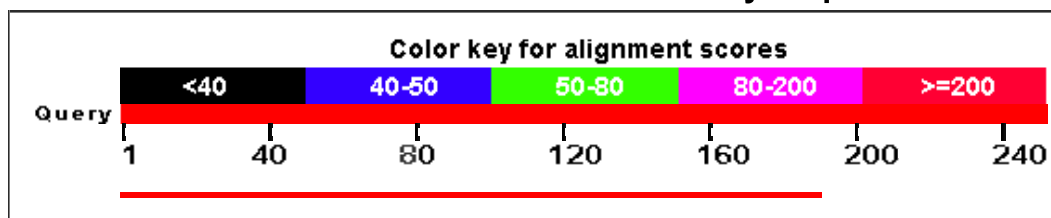
## EGV T. harzianum IOC-3844

**Query ID** |c|13963  
**Description** EGV T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 247

**Subject ID** 13965  
**Description** EGV T. reesei CAA83846.1  
**Molecule type** amino acid  
**Subject Length** 242  
**Program** BLASTP 2.2.26+

Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
13965	EGV T. reesei CAA83846.1	<a href="#">312</a>	312	75%	2e-111	84%	

## Alignments

>lcl|13965 EGV T. reesei CAA83846.1  
Length=242

Score = 312 bits (799), Expect = 2e-111, Method: Compositional matrix adjust.  
Identities = 157/187 (84%), Positives = 164/187 (88%), Gaps = 1/187 (1%)

```

Query 1      MKAPLLIALGIAGAASAFKATTTRYDQEGACGCGSSSSGGAASWQLGIGNGVYTAAGSQ 60
           MKA L++ I GA SA+KATTTRYDQEGACGCGSSSSG A WQLGIGNGVYTAAGSQ
Sbjct 1      MKATLVLGSLIVGAVSAYKATTTRYDQEGACGCGSSSSG-AFPWQLGIGNGVYTAAGSQ 59

Query 61     AFFDTAGASWCGAGCGKCYQLTSTGEAPCSTCGKGGVAGQSIIVMVTNLCPNNGNAQWCP 120
           A FDTAGASWCGAGCGKCYQLTSTG+APCS+CG GG AGQSIIVMVTNLCPNNGNAQWCP
Sbjct 60     ALFDTAGASWCGAGCGKCYQLTSTGQAPCSSCGTGGAAGQSIIVMVTNLCPNNGNAQWCP 119

Query 121    TVGGTNQYGYSYHFDIMAQNEIFGDNVVVDFEPVVCPGQATSDWETCLCIDMQETDTPV 180
           VGGTNQYGYSYHFDIMAQNEIFGDNVVVDFEP+ CPGQA SDW TCLC+ QETD TPV
Sbjct 120    VVGGTNQYGYSYHFDIMAQNEIFGDNVVVDFEPIACPGQAASDWGTCLCVGQQETDTPV 179

Query 181    GLGDPGS 187
           D GS
Sbjct 180    LGNDTGS 186
    
```

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/ BLAST/ blastp suite-2sequences/ Formatting Results - TUSHV7Z611R](#)

[Formatting options](#)

[Download](#)

## Blast 2 sequences

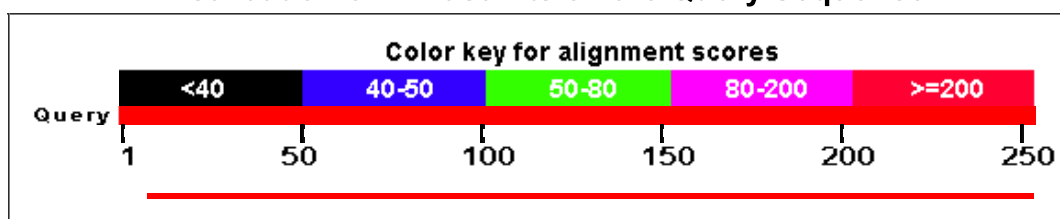
## EGVII T. harzianum IOC-3844 (252 letters)

**Query ID** Icl|1245  
**Description** EGVII T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 252

**Subject ID** 1247  
**Description** EGVII T. reesei AAP57753.1  
**Molecule type** amino acid  
**Subject Length** 249  
**Program** BLASTP 2.2.26+

Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
1247	EGVII T. reesei AAP57753.1	<a href="#">446</a>	446	96%	5e-164	86%	

## Alignments

>lcl|1247 EGVII T. reesei AAP57753.1  
Length=249

Score = 446 bits (1147), Expect = 5e-164, Method: Compositional matrix adjust.  
Identities = 211/244 (86%), Positives = 228/244 (93%), Gaps = 0/244 (0%)

Query	9	VQKSWGCLAGSALGHGQIQNFTLNGAYNQGFILDYYYQKVNTGHFPNVAGWYAEDLDLGF	68
		+ + GCLAGS LGHGQ+QNFT+NG YNQGFILDYYYQK NTGHFPNVAGWYAEDLDLGF	
Sbjct	6	ILAALGCLAGSVLGHGQVQNFTINGQYNQGFILDYYYQKQNTGHFPNVAGWYAEDLDLGF	65
Query	69	ISPDAYTTPDIICHKNAAPGAATATAAAGSTIVFSWGPWPWHPYGPILTYVAQCSGSCT	128
		ISPD YTTPDI+CHKNAAPGA +ATAAAGS IVF WGP WHPYGPPI+TYV +CSGSCT	
Sbjct	66	ISPDQYTTDPDIVCHKNAAPGAISATAAAGSNIVFQWGPVWHPYGPPIVITYVVECSGSCT	125
Query	129	TVDKTSLRWVKIQESGINYNTQVWAQQDLINQGSKWTVKIPSSLKPGNYVFRNEILAAHG	188
		TV+K +LRWVKIQE+GINYNTQVWAQQDLINQG+KWTVKIPSSL+PGNYVFR+E+LAAHG	
Sbjct	126	TVNKNLNRWVKIQEAGINYNQVWAQQDLINQGNKWTVKIPSSLRPGNYVFRHELLAAHG	185
Query	189	ASSQNGMQNYPQCVNIVVTGSGTKSLPAGTPATSLYKPTDPGILFNPYTTITNYTIPGPA	248
		ASS NGMQNYPQCVNI VTGSGTK+LPAGTPAT LYKPTDPGILFNPYTTIT+YTIPGPA	
Sbjct	186	ASSANGMQNYPQCVNIAVVTGSGTKALPAGTPATQLYKPTDPGILFNPYTTITSYTIPGPA	245
Query	249	LWQG 252	
		LWQG	
Sbjct	246	LWQG 249	



## BLAST®

## Basic Local Alignment Search Tool

[NCBI/BLAST/blastp\\_suite-2sequences/Formatting Results - TUSMFKMG11R](#)[Formatting options](#)[Download](#)

## Blast 2 sequences

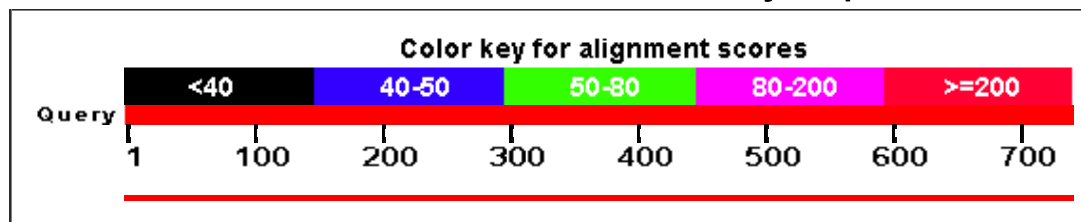
## BGL1 T. harzianum IOC-3844 (734 letters)

**Query ID** lc|22113  
**Description** BGL1 T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 734

**Subject ID** 22115  
**Description** BGL1 T. reesei AAA18473.1  
**Molecule type** amino acid  
**Subject Length** 744  
**Program** BLASTP 2.2.26+

[Graphic Summary](#)

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [PubChem](#) [BioAssay](#)

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
22115	BGL1 T. reesei AAA18473.1	<a href="#">1338</a>	1338	100%	0.0	88%	

## Alignments

>lcl|22115 BGL1 T. reesei AAA18473.1  
Length=744

Score = 1338 bits (3462), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 656/744 (88%), Positives = 695/744 (93%), Gaps = 10/744 (1%)

Query	1	MRYTSAAALAI AAPFVRADSK-----VVPAGTPWGVAYGKAQAALAKLTLQDKV	51
		MRY +AAALA+A PF RADS VVPAGTPWG AY KA+AAALAKL LQDKV	
Sbjct	1	MRYRTAAALALATGPFARADSHSTSGASAEAVVPPAGTPWGTAYDKAKAALAKLNLQDKV	60
Query	52	GIVSGVGWNQGPCVGNTPVSKIGYPSLCLQDGPLGVRFSSGSTAFTPGVQAASSTWDLNL	111
		GIVSGVGWN GPCVGNTP SKI YPSLCLQDGPLGV+S+GSTAFTPGVQAASSTWD+NL	
Sbjct	61	GIVSGVGWNGGPCVGNTPASKISYPSLCLQDGPLGVRYSTGSTAFTPGVQAASSTWDVNL	120
Query	112	IRERGGQFIGQEVVDTGIHVTLGVPVAGPLGKTPQGGRNWEGFSVDPYLTGLAMAQTINGIQ	171
		IRERGGQFIG+EV+ +GIHV LGPVAGPLGKTPQGGRNWEGF VDPYLTG+AM QTINGIQ	
Sbjct	121	IRERGGQFIGEEVKASGIHVILGVPVAGPLGKTPQGGRNWEGFGVDPYLTGIAMGQTINGIQ	180
Query	172	SVGQVQATAKHYYILNEQERNRESMSNADDRTLHELHYAWPFADAVQANVASVMCSYNKINT	231
		SVGQVQATAKHYYILNEQE NRE++SSN DDRTLHELHY WPFADAVQANVASVMCSYNK+NT	
Sbjct	181	SVGQVQATAKHYYILNEQELNRETISSNPDDDRTLHELHYTWPFADAVQANVASVMCSYNKVNT	240
Query	232	TWACEDQYTLQTLTKDQLGFPYVMTDWNQAHSTVQAATAGLDMSMPGTFDNGNRLWGP	291
		TWACEDQYTLQTLTKDQLGFPYVMTDWNQAH+TVQ+A +GLDMSMPGTFDNGNRLWGP	
Sbjct	241	TWACEDQYTLQTVLTKDQLGFPYVMTDWNQAHSTTVQANSGLDMSMPGTFDNGNRLWGP	300
Query	292	ALTNVAVNSGQVPSRVDDMVTRILAAWYLAGQDAAGYPSLSLSRVQVGHKTNVRSIARD	351
		ALTNVAVNS QVP+SRVDDMVTRILAAWYL QD AGYPS ++SRVQVGHKTNVRIARD	
Sbjct	301	ALTNVAVNSQVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNISRVQVGNHKTNVRAIARD	360
Query	352	GIVLLKNDGNILPLKPPASIAVIGSAAIIGAHASNSGSCGDKGCDNGALGMGWGSGAVNY	411
		GIVLLKND NILPLKPPASIAV+GSAAIIG HA NS SC DKGCD+GALGMGWGSGAVNY	
Sbjct	361	GIVLLKNDANILPLKPPASIAVVGSAAIIGNHARNSPSCNDKGCDNGALGMGWGSGAVNY	420
Query	412	PYFVAPYDAINTRATSQGTQVTLSTNTDNTSSGASAAAGKDVAVIVFITADSGEGYITVEGN	471
		PYFVAPYDAINTRA+SQGT+VTLSTNTDNTSSGASAA GKDVAIVFITADSGEGYITVEGN	
Sbjct	421	PYFVAPYDAINTRASSQGTQVTLSTNTDNTSSGASAAARGKDVAVIVFITADSGEGYITVEGN	480
Query	472	AGDRNDLNAWHNGNALVQAVAGSNQNVIVVVHVSVAIILEQIILALPQVKAIVWAGLPSQE	531
		AGDRN+L+ WHNGNALVQAVAG+N NVIVVVHVSVAIILEQI+ALPQVKA+VWAGLPSQE	
Sbjct	481	AGDRNLDLPWHNGNALVQAVAGANSNVIVVVHVSVAIILEQIILALPQVKAIVWAGLPSQE	540
Query	532	SGNALVDVLWGDVSPSGKLVYTIKSPNDYNTRITS-GDDSFSEGLFIDYKHFDDAGITP	590
		SGNALVDVLWGDVSPSGKLVYTIKSPNDYNTRI S G DSFSEGLFIDYKHFDDA ITP	
Sbjct	541	SGNALVDVLWGDVSPSGKLVYTIKSPNDYNTRIVSGGSDSFSEGLFIDYKHFDDANITP	600
Query	591	RYEFGFGLSYTKFNYSRSLSVLSTAKSGPATGAVVPGGSDLFQNVATITVDITNSGAVTG	650
		RYEFG+GLSYTKFNYSRSLSVLSTAKSGPATGAVVPGGSDLFQNVAT+TVDI NSG VTG	
Sbjct	601	RYEFGYGLSYTKFNYSRSLSVLSTAKSGPATGAVVPGGSDLFQNVATVTVDIANSQVVTG	660
Query	651	AEVAQLYLTYPSSAPRTPPKQLRGFQVSLTAGQTGTATFNIRRRDLSYWDTSQKVVVP	710
		AEVAQLY+TYPSSAPRTPPKQLRGFQVSLTAGQTGTATFNIRRRDLSYWDTSQKVVVP	
Sbjct	661	AEVAQLYITYPSSAPRTPPKQLRGFQVSLTAGQTGTATFNIRRRDLSYWDTSQKVVVP	720
Query	711	SGSFGISVGASSRDIRLTGSLSVS 734	
		SGSFGISVGASSRDIRLT +LSV+	
Sbjct	721	SGSFGISVGASSRDIRLTSTLSVA 744	



## BLAST®

## Basic Local Alignment Search Tool

[NCBI/BLAST/blastp\\_suite-2sequences/Formatting Results - TUSPZ8KN11R](#)

[Formatting options](#)

[Download](#)

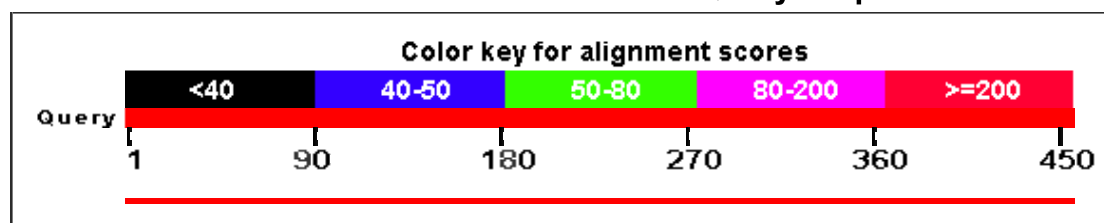
## Blast 2 sequences

## BGL2 T. harzianum IOC-3844 ABN50090.1

<b>Query ID</b>	lc 54733	<b>Subject ID</b>	54735
<b>Description</b>	BGL2 T. harzianum IOC-3844 ABN50090.1	<b>Description</b>	BGL2 T. reesei BAA74959.1
<b>Molecule type</b>	amino acid	<b>Molecule type</b>	amino acid
<b>Query Length</b>	453	<b>Subject Length</b>	466
		<b>Program</b>	BLASTP 2.2.26+

## Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
54735	BGL2 T. reesei BAA74959.1	<a href="#">854</a>	854	100%	0.0	89%	

## Alignments

>lcl|54735 BGL2 T. reesei BAA74959.1  
Length=466

Score = 854 bits (2206), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 406/454 (89%), Positives = 429/454 (94%), Gaps = 1/454 (0%)

Query	1	MLPKDFQWGFATAAYQIEGAIDKDGGRPSIWDTFCAIPGKIADGTSGVTACDSYNRTAED	60
Sbjct	1	MLPKDFQWGFATAAYQIEGA+D+DGRGPSIWDTFCA PGKIADG+SGVTACDSYNRTAED	60
Query	61	IALLKSLGAKSYRFSISWSRIIPKGGRRDDPVNQLGIDHYAQFVDDLLEAGITPFITLFWH	120
Sbjct	61	IALLKSLGAKSYRFSISWSRIIP+GGR D VNQ GIDHY +FVDDLL+AGITPFITLFWH	120
Query	121	DLPEELHQRYGGLLNRTFPLDFENYARVMFKALPKVRNWITFNEPLCSAIPGYGSGTFA	180
Sbjct	121	DLPE LHQRYGGLLNRTFPLDFENYARVMF+ALPKVRNWITFNEPLCSAIPGYGSGTFA	180
Query	181	PGRQSTTEPWIVGHNLLVAHGRAVKVYRDEFKDLN-DGQIGIVLNGDFTYPWDSSDPLDR	239
Sbjct	181	PGRQST+EPW VGHN+LVAHGRAVK YRD+FK + DGQIGIVLNGDFTYPWD++DP D+	240
Query	240	EEAERRLEFFTAWAYADPIYLGDPASMRKQLGDRLPEFTPEEKAFVLGSNDFYGMNHYTS	299
Sbjct	241	EEAERRLEFFTAW+ADPIYLGDPASMRKQLGDRLP FTPEE+A V GSNDFYGMNHYTS	300
Query	300	NYIRHRTSPATADDTVGNVDVLFYFNKEGQCIGPETESSWLRPCAGFRDFLWISKRYNY	359
Sbjct	301	NYIRHR+SPA+ADDTVGNVDVLF NK+G CIGPET+S WLRPC AGFRDFLWISKRY Y	360
Query	360	PKIYVTENGTSKLGENDLPKEKILEDDFRVNYNEYIRAMFTAATLDGVNVKGYFAWSLM	419
Sbjct	361	PPIYVTENGTSIKGESDLPKEKILEDDFRVKYNEYIRAMVTAVELDGVNVKGYFAWSLM	420
Query	420	DNFEWADGYVTRFGVTYVDYENGQQRFPKKSAKS	453
Sbjct	421	DNFEWADGYVTRFGVTYVDYENGQ+RFPKKSAKS	454

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/ BLAST/ blastp suite-2sequences/ Formatting Results - TUSTF3ZU11R](#)

[Formatting options](#)

[Download](#)

## Blast 2 sequences

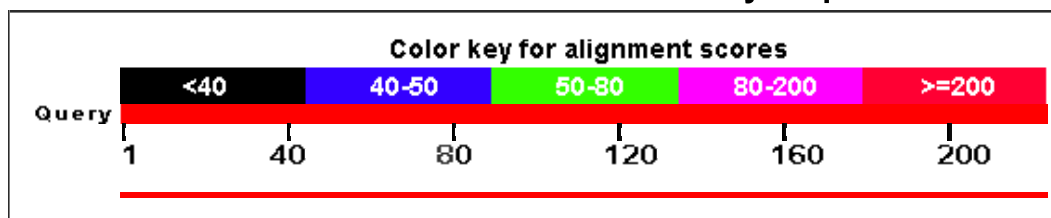
## Xyn2 T. harzianum IOC-3844

**Query ID** |cl|17085  
**Description** Xyn2 T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 221

**Subject ID** 17087  
**Description** Xyn2 T. reesei CAA49293.1  
**Molecule type** amino acid  
**Subject Length** 222  
**Program** BLASTP 2.2.26+

Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence



## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
17087	Xyn2 T. reesei CAA49293.1	<a href="#">387</a>	387	100%	7e-142	88%	

## Alignments

>lcl|17087 Xyn2 T. reesei CAA49293.1  
Length=222

Score = 387 bits (995), Expect = 7e-142, Method: Compositional matrix adjust.  
Identities = 196/222 (88%), Positives = 202/222 (91%), Gaps = 1/222 (0%)

Query	1	MVAFTSLLAGFA-AIAGVLSAPTESSVEVEKRQTIGPGTGYSNGYYSYWNDDGHAGVTTYT	59
		MV+FTSLLA + A A SV VEKRQTI PGTGY+NGY+YSYWNDDGH GVTYT	
Sbjct	1	MVSFTSLLAASPPSRASCRPAAEVESVAVEKRQTIQPGTGYNNGYFYSYWNDDGHGGVTTYT	60
Query	60	NGGGGSFTVNWSNSGNFVGGKGWQPGTKNKVINFGSYNPNNGNSYLSIYGWSRNPLIEYY	119
		NG GG F+VNWSNSGNFVGGKGWQPGTKNKVINFGSYNPNNGNSYLS+YGWSRNPLIEYY	
Sbjct	61	NGPGGQF+VNWSNSGNFVGGKGWQPGTKNKVINFGSYNPNNGNSYLSVYGWSRNPLIEYY	120
Query	120	IVENFGTYNPSTGATKLGEVTSDGSVYDIYRTQRVNQPSIIGTATFYQYWSVRRNHRSSG	179
		IVENFGTYNPSTGATKLGEVTSDGSVYDIYRTQRVNQPSIIGTATFYQYWSVRRNHRSSG	
Sbjct	121	IVENFGTYNPSTGATKLGEVTSDGSVYDIYRTQRVNQPSIIGTATFYQYWSVRRNHRSSG	180
Query	180	SVNTANHFNAWASHGLTLGTMDYQIVAVEGYFSSGSASITVS	221
		SVNTANHFNAWA GLTLGTMDYQIVAVEGYFSSGSASITVS	
Sbjct	181	SVNTANHFNAWAQQGLTLGTMDYQIVAVEGYFSSGSASITVS	222

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/BLAST/blastp\\_suite-2sequences/Formatting Results - TUSXK4V911R](#)

[Formatting options](#)

[Download](#)

## Blast 2 sequences

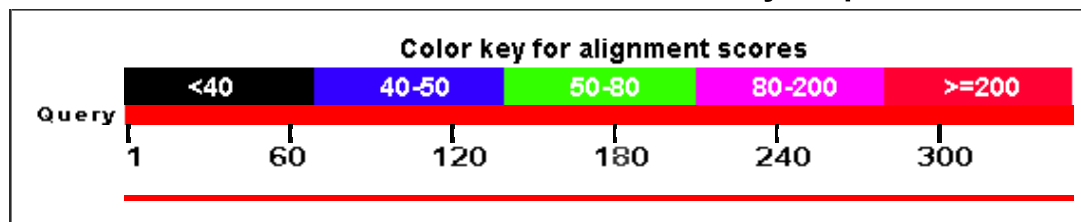
## Xyn3 T. harzianum IOC-3844

**Query ID** |c|51961  
**Description** Xyn3 T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 347

**Subject ID** 51963  
**Description** Xyn3 T. reesei BAA89465.2  
**Molecule type** amino acid  
**Subject Length** 347  
**Program** BLASTP 2.2.26+

Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence





## Dot Matrix View



## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
51963	Xyn3 T. reesei BAA89465.2	<a href="#">571</a>	571	100%	0.0	83%	

## Alignments

>lcl|51963 Xyn3 T. reesei BAA89465.2  
Length=347

Score = 571 bits (1471), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 288/347 (83%), Positives = 311/347 (90%), Gaps = 0/347 (0%)

Query	1	MKANIIFCLLAPLVAALPAEFIRLEPELAARRVNITERMADLEDROASVSIDQLFKRKGGK	60
Sbjct	1	MKAN+I CLLAPLVAALP E I L+PELAA R N+TER ADL DRQAS SIDQL KRKGGK	60
Query	61	VYFGTATDRGLLQRE+NAAIIQA+LGQVTPENSMKWQSLEPNQGOYNWGDADYLVNFAQQ	120
Sbjct	61	LYFGTATDRGLLQREKNAAIIQADLGQVTPENSMKWQSLENNQGOYNWGDADYLVNFAQQ	120
Query	121	NGKTIRGHTLIWHSQLPNWNINNADTLRQAIRTHVLTVVGRYKGGKIRAWDVVNEIFNE	180
Sbjct	121	NGK+IRGHTLIWHSQLP WVNINNADTLRQ IRTHV TVVGRYKGGKIRAWDVVNEIFNE	180
Query	181	DGTLRSSVFSRLLGEEFVSI AFRAAREADPSARLYINDYNLDVAGSSKVNLMRYVVDKWI	240
Sbjct	181	DGTLRSSVFSRLLGEEFVSI AFRAARDADPSARLYINDYNLDRANYGKVNGLKTYVSKWI	240
Query	241	SQGVPIDGIGTQTHIGAGGGASVQALQQLATAPVTELAITELDIVGAPSNDSYSAVVRGC	300
Sbjct	241	SQGVPIDGIGS+Q+H+ GGG+ GALQQLAT PVTELAITELDI GAP+ DY+ VV+ C	300
Query	301	LNVPKCLGITVWGISDKDSWRSGSSPLLFDNFNPKAAYNSIVSILQ	347
Sbjct	301	L+V KC+GITVWGISDKDSWR+ ++PLLFD+NFNPK AYSIV ILQ	347
		LSVSKCVGITVWGISDKDSWRASNPLLFDANFNPKPAYNSIVGILQ	347

## BLAST®

## Basic Local Alignment Search Tool

[NCBI/BLAST/blastp\\_suite-2sequences/Formatting Results - TUT0KUV011R](#)

[Formatting options](#)

[Download](#)

## Blast 2 sequences

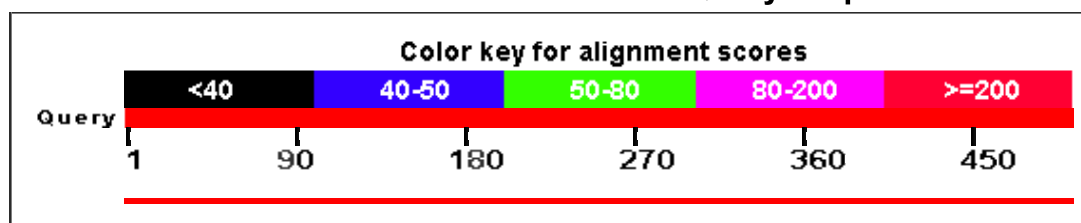
## SWOI T. harzianum IOC-3844

**Query ID** lc|21039  
**Description** SWOI T. harzianum IOC-3844  
**Molecule type** amino acid  
**Query Length** 499

**Subject ID** 21041  
**Description** SWOI T. reesei CAB92328.1  
**Molecule type** amino acid  
**Subject Length** 493  
**Program** BLASTP 2.2.26+

Graphic Summary

## Distribution of 1 Blast Hits on the Query Sequence



# Dot Matrix View



## Descriptions

Legend for links to other resources: **U** UniGene **E** GEO **G** Gene **S** Structure **M** Map Viewer **Pub** PubChem BioAssay

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
21041	SWOI T. reesei CAB92328.1	<a href="#">849</a>	849	100%	0.0	87%	

## Alignments

>lcl|21041 SWOI T. reesei CAB92328.1  
Length=493

Score = 849 bits (2193), Expect = 0.0, Method: Compositional matrix adjust.  
Identities = 436/499 (87%), Positives = 458/499 (92%), Gaps = 6/499 (1%)

Query	1	MARKLSLLALASLISLSVQONCAALFGQC	GNGWTGSTCCVSGAQCTYVND	FYSQCLAST	60
Sbjct	1	MA KL L+ALASL+SLS+QONCAALFGQC	GW+G+TCCV+GAQC++VND+YSQCLAST		60
Query	61	GGGSTTSRSTSSSSVSRSSSTSSVSRSSS	SAGSSPGGGSP	TGGASTYTTTDTATVAPHSQ	120
Sbjct	61	GG +S+S S + S+S+S S S S	PTG ASTYTTTDTATVAPHSQ		114
Query	121	SPYPSIAASSCGSWTLVDNVCCPSYC	ATDDTSEFCTGGCTCATPPSADCKSGTMYPEVHH		180
Sbjct	115	SPYPSIAASSCGSWTLVDNVCCPSYCA	DDTSE C+G TC TPPSADCKSGTMYPEVHH		174
Query	181	VTSNETWHYSRSTHFGLTSGGACGFGLYGL	CTKGSVTASWTDPMLGSTCDAFCTAYPLLC		240
Sbjct	175	VSSNESWHYSRSTHFGLTSGGACGFGLYGL	CTKGSVTASWTDPMLGATCDAFCTAYPLLC		234
Query	241	KDPTGTTLRGNFAAPNGDYYSQFWSSL	PGALDNYLSCGECIELIQTKPDGTDYAVGEAGY		300
Sbjct	235	KDPTGTTLRGNFAAPNGDY+QFWSSL	PGALDNYLSCGECIELIQTKPDGTDYAVGEAGY		294
Query	301	TDPITILEIVDSCPCSANSKWCCPGADHC	GEIDFKYGCPLPADSIHLDLSDIAMGRLQGN		360
Sbjct	295	TDPITILEIVDSCPCSANSKWCCPGADHC	GEIDFKYGCPLPADSIHLDLSDIAMGRLQGN		354
Query	361	GSLTNGVIPTRYKRVPCPKLGNVYIWL	RNNGGGPYFALTAVNTNGPGSVTKIEIKGAGTD		420
Sbjct	355	GSLTNGVIPTRY+RV CPK+GN YIWL	RNNGGGPYFALTAVNTNGPGSVTKIEIKGA TD		414
Query	421	TWVAMEHDPNYTSSLPTRRYG	SWVIPQGS	GFNLPVGIRLTSPTGEQIVNEQAIKTFTPP	480
Sbjct	415	NWVALVHDPNYTSSRPQERYG	SWVIPQGS	GFNLPVGIRLTSPTGEQIVNEQAIKTFTPP	474
Query	481	ATADPNFYIIDIGVQFSQN		499	
Sbjct	475	ATDPNFYIIDIGVQFSQN		493	

## Apêndice H: Valores de expressão calculados para as figuras 5, 6 e 7.

Os valores de expressão listados abaixo revelam o número de vezes que o transcrito é mais ou menos representado em relação ao alvo calibrador.

FIGURA 5 (Induzido X não induzido)			FIGURA 6 (Induzido X não induzido 8h)			FIGURA 7 (Induzido X Actina Induzida)		
Transcrito / Tempo	Expressão	P-valor	Transcrito / Tempo	Expressão	P-valor	Transcrito / Tempo	Expressão	P-valor
<b>CBHI</b>			<b>CBHI</b>			<b>CBHI</b>		
8h	19,307	Up P=0,00	8/8h	1	P=0,328	8h	-4640,29	Down P=0,00
			24/8h	9896,868	Up P=0,00	24h	1,039	P=0,667
36h	164498,23	Up P=0,00	36/8h	204621,76	Up P=0,00	36h	64,222	Up P=0,00
			48/8h	150693,25	Up P=0,00	48h	56,689	Up P=0,00
60h	15929,413	Up P=0,00	60/8h	100968,5	Up P=0,00	60h	40,504	Up P=0,00
<b>CBHII</b>			<b>CBHII</b>			<b>CBHII</b>		
8h	98,76	Up P=0,00	8/8h	1	P=0,328	8h	-4721,403	Down P=0,00
			24/8h	12420,421	Up P=0,00	24h	1,117	P=0,175
36h	147746,79	Up P=0,00	36/8h	228490,52	Up P=0,00	36h	35,629	Up P=0,00
			48/8h	176296,03	Up P=0,00	48h	34,297	Up P=0,00
60h	7509,902	Up P=0,00	60/8h	116916,28	Up P=0,00	60h	24,818	Up P=0,00
<b>EGI</b>			<b>EGI</b>			<b>EGI</b>		
8h	1,072	P=0,675	8/8h	1	P=0,328	8h	-76596,986	Down P=0,00
			24/8h	21161,127	Up P=0,00	24h	-6,364	Down P=0,00
36h	112798,46	Up P=0,00	36/8h	259999,33	Up P=0,00	36h	4,874	Up P=0,00
			48/8h	214701,57	Up P=0,00	48h	4,532	Up P=0,00
60h	44421,607	Up P=0,00	60/8h	171748,87	Up P=0,00	60h	3,918	Up P=0,00
<b>EGII</b>			<b>EGII</b>			<b>EGII</b>		
8h	3,774	Up P=0,00	8/8h	1	P=0,328	8h	-3590,577	Down P=0,00
			24/8h	1353,447	Up P=0,00	24h	-6	Down P=0,00
36h	12924,378	Up P=0,00	36/8h	19967,056	Up P=0,00	36h	6,277	Up P=0,00
			48/8h	8427,781	Up P=0,00	48h	3,171	Up P=0,00
60h	4941,959	Up P=0,00	60/8h	5711,075	Up P=0,00	60h	2,297	Up P=0,00
<b>EGIII</b>			<b>EGIII</b>			<b>EGIII</b>		
8h	0,996	P=1,00	8/8h	1	P=0,341	8h	-103552,3	Down P=0,00
			24/8h	5494,913	Up P=0,00	24h	-34,535	Down P=0,00
36h	72984,755	Up P=0,00	36/8h	326894,98	Up P=0,00	36h	3,317	Up P=0,00
			48/8h	330017,44	Up P=0,00	48h	4,257	Up P=0,00
60h	5641,159	Up P=0,00	60/8h	273061,23	Up P=0,00	60h	3,811	Up P=0,00
<b>EGVII</b>			<b>EGVII</b>			<b>EGVII</b>		
8h	12,752	Up P=0,00	8/8h	1	P=0,328	8h	-30467,848	Down P=0,00
			24/8h	3048,049	Up P=0,00	24h	-29,446	Down P=0,00
36h	535978,98	Up P=0,00	36/8h	454014,04	Up P=0,00	36h	9,448	Up P=0,00
			48/8h	583834,87	Up P=0,00	48h	15,189	Up P=0,00
60h	232243,76	Up P=0,00	60/8h	406098,96	Up P=0,00	60h	11,511	Up P=0,00
<b>SWO</b>			<b>SWO</b>			<b>SWO</b>		
8h	18,046	Up P=0,00	8/8h	1	P=0,328	8h	-803,414	Down P=0,00
			24/8h	28,185	Up P=0,00	24h	-82,424	Down P=0,00
36h	6909,486	Up P=0,00	36/8h	1133,243	Up P=0,00	36h	1,083	P=0,175
			48/8h	809,866	Up P=0,00	48h	-1,409	Down P=0,00
60h	2456,22	Up P=0,00	60/8h	481,818	Up P=0,00	60h	-1,959	Down P=0,00