



UNIVERSIDADE DE BRASÍLIA
REDE CENTRO-OESTE DE PÓS-GRADUAÇÃO, PESQUISA E INOVAÇÃO
DOUTORADO EM BIOTECNOLOGIA E BIODIVERSIDADE

EXPRESSÃO DE CELULASES RECOMBINANTES DE INSETOS
PRAGA EM *PICHLIA PASTORIS*TM

ANA ELIZABETH OLIVEIRA DE ARAÚJO ALVES

Brasília
Distrito Federal - Brasil
© 2020

ANA ELIZABETH OLIVEIRA DE ARAÚJO ALVES

**EXPRESSÃO DE CELULASES RECOMBINANTES DE INSETOS
PRAGA *EMPICHIA PASTORIS*TM**

Orientador: Dr.^a Maria Fátima Grossi-de-Sa, PhD

Coorientador: Dr. Érico Augusto Rosas de Vasconcelos, PhD

Tese apresentada ao Programa de Pós-Graduação em Biotecnologia e Biodiversidade da Rede Pró-Centro-Oeste do Instituto de Ciências Biológicas da UnB como requisito parcial para obtenção do título de Doutor em Biotecnologia e Biodiversidade.

Brasília
Distrito Federal - Brasil
© 2020

ANA ELIZABETH OLIVEIRA DE ARAÚJO ALVES

O desenvolvimento das atividades experimentais da tese “**Expressão de Celulases Recombinantes de Insetos Praga em *Pichia pastoris*TM**”, apresentada ao Programa de Pós-Graduação em Biotecnologia e Biodiversidade da Rede Centro-Oeste de Pós-Graduação, Pesquisa e Inovação, Universidade de Brasília (UnB), contou com a colaboração direta do Laboratório de Interação Molecular de Planta-Praga (LPPI), do Empresa Recursos Genéticos e Biotecnologia.

Ademais, as atividades desenvolvidas tiveram o apoio financeiro das agências de fomento à pesquisa: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Banca Examinadora:

Dr.^a Maria Fátima Grossi-de-Sa
EMBRAPA-CENARGEM

Orientadora

Dr. Érico Augusto Rosas de Vasconcelos
UNIEURO
Co-orientador

Dr. Osmindo Rodrigues Pires Júnior
UnB
Examinador Interno

Dr. Guilherme Souza Prado
UCB
Examinador Externo

Dr. Osmundo Brillhante de Oliveira Neto
UNIEURO
Examinador Externo

Dr. Túlio César de Lima Lins
UNICEPLAC
Suplente

Exclusivamente a Deus. Meu Pai
protetor, fiel, consolador.

AGRADECIMENTOS

Agradeço a Deus por existir, pela minha vida e o respirar de todos os dias;

Agradeço à minha família e peço perdão pela minha ausência em tantos momentos especiais. Meu esposo Adriano, meus filhos Maria Eduarda e João Pedro, minha mãezinha Elza, meu pai Antonio;

Ao meu grande amigo e colega de trabalho Dr. Érico Rosas Vasconcelos, onde tudo começou. Grande pesquisador, rei do Western blot;

À Dr.^a Maria Fátima Grossi, que autorizou minha participação e orientação;

Ao Dr. Antonino, chamado carinhosamente de Dijair, o rei da clonagem;

À Elínea, que me ensinou passos muito importantes nesta tese;

Ao Joel, que, além de amigo, tem um coração grandioso e me ajudou com as quantificações;

À Daiane, pequenina, mas um gigante no combate aos desafios. Me ensinou a fazer SDS;

À Liz Nathália, amiga paraguaia, mas que é super brasileira, rainha dos ensaios;

Ao Dr. Guilherme e Dr.^a Vanessa pela dedicação nas correções e participação na banca de qualificação;

À Ana Gabriela, Hudson e Clídia, com suas palavras de conforto e carinho;

A toda a equipe do laboratório, uma vez que ninguém faz nada sozinho;

À Dr.^a Roberta, Dr. Peyman, Dr.^a Cristina, Dr. Leonardo, Dr. Osmundo;

Agradeço a todos os companheiros dessa jornada;

À Capes e CNPq pelo apoio financeiro;

Ao Prof. César Grisólia, meu orientador do Mestrado, que me deu o conselho mais importante da minha carreira;

Aos meus amigos da Uniceplac, verdadeiros e eternos companheiros;

A todos que, de alguma maneira, me ajudaram para a realização deste trabalho.

SUMÁRIO

1. INTRODUÇÃO	1
2. REVISÃO BIBLIOGRÁFICA	4
2.1. Biomassa vegetal e biocombustíveis.....	4
2.1.1. Panorama do etanol 2G	5
2.2. Embrapa: principais contribuições para a produção de etanol 2G.....	7
2.3. Insetos praga: fonte para produção de celulases.....	9
2.4. Transcriptomas.....	15
2.5. Celulases.....	16
2.5.1.β-glicosidase industrial	18
2.5.2. Endoglucanase industrial	19
2.6. Sistema <i>Pichia pastoris</i> TM (<i>Komagataella phaffii</i>) utilizado para expressão de proteínas recombinantes.....	20
3. JUSTIFICATIVA	22
4. OBJETIVOS	23
4.1. Objetivo Geral	23
4.2. Objetivos Específicos	23
5. MATERIAL E MÉTODOS	24
5.1. Seleção de contigs dos transcriptomas de <i>Telchin licus licus</i> e <i>Anthonomus grandis</i>	24
5.2. Extração do RNA total de insetos e isolamento dos genes de interesse.....	24
5.3. Quantificação da expressão dos transcritos por qRT-PCR.....	27
5.4. Vetor de clonagem.....	27

5.5. Preparação e transformação de células bacterianas.....	28
5.6. Preparação do DNA plasmidial.....	29
5.7. Vetor de Expressão.....	30
5.8. Transformação de <i>Pichia pastoris</i> TM (<i>K.phaffi</i>).....	31
5.9. Expressão de proteínas no sistema <i>P.pastoris</i> TM	33
5.10. SDS-PAGE, Western blot e ELISA	33
5.11. Ensaio de difusão radial.....	34
5.12. Dosagem de açúcar redutor com DNS.....	35
6. RESULTADOS E DISCUSSÃO.....	36
6.1. Análise dos Transcriptomas de <i>T. licus licus</i> e <i>A. grandis</i>	36
6.2. Clonagem e subclonagem	41
6.3. Expressão recombinante/heteróloga	43
6.4. Caracterização funcional das enzimas TlicBgluc1 e AgraGH45-1.....	45
7. CONCLUSÃO	52
REFERÊNCIAS	53
ANEXO 1- Curva de calibração padrão de BSA.....	66
ANEXO 2- Curva de calibração do ensaio com açúcares redutores.....	66
ANEXO 3- Artigo publicado.....	67

LISTA DE FIGURAS

Figura 1. Esquema generalizado do tubo digestivo de insetos.....	10
Figura 2. Ciclo de vida da broca-gigante (<i>T. licus licus</i>)	12
Figura 3. <i>Anthonomus grandis</i> adulto.....	13
Figura 4. Degradação da biomassa da cana-de-açúcar por lagartas	14
Figura 5. Representação esquemática da degradação da celulose.....	16
Figura 6. Representação esquemática da hidrólise na ligação β -1,4	18
Figura 7. Representação esquemática do vetor de clonagem pGEM®-T Easy	28
Figura 8. Representação esquemática do vetor de expressão.....	31
Figura 9. Alinhamento das sequências gênicas.....	39
Figura 10. qRT-PCR dos transcritos <i>TlicBgluc1</i>	40
Figura 11. qRT-PCR dos transcritos <i>AgraGH45-1</i>	41
Figura 12. Análise de restrição do vetor pGEM®-T Easy com os genes candidatos...	42
Figura 13. Presença do gene <i>TlicBgluc1</i> no genoma da levedura transformada.....	43
Figura 14. Presença do gene <i>AgraGH45-1</i> no genoma da levedura transformada.....	43
Figura 15. Análise da expressão da proteína recombinante por SDS-PAGE e <i>Western Blotting</i> de <i>TlicBgluc1</i>	44
Figura 16. Análise da expressão da proteína recombinante por SDS-PAGE e <i>Western Blotting</i> de <i>AgraGH45-1</i>	44
Figura 17. Atividade celulolítica qualitativa da enzima recombinante β -glicosidase <i>TlicBgluc1</i> em substratos de carboximetilcelulose (CMC) e celobiose.....	46
Figura 18. Atividade celulolítica qualitativa da enzima recombinante β -1,4-endoglucanase <i>AgraGH45-1</i> em substrato de carboximetilcelulose (CMC).....	47
Figura 19. Atividade enzimática β -glicosidase <i>TlicBgluc1</i> no ensaio de quantificação de açúcar redutor mediado por DNS.....	48
Figura 20. Avaliação da atividade enzimática de β -1,4-endoglucanase <i>AgraGH45-1</i>	49

LISTA DE TABELAS

Tabela 1. Dados da produção brasileira de etanol, açúcar e bagaço da cana-de-açúcar das safras de 2014 a 2018.....	04
Tabela 2. Principais microrganismos utilizados na produção de endoglucanase por fermentação em estado sólido.	20
Tabela 3. Oligonucleótidos para clonagem no das ORFs dos genes candidatos.....	41
Tabela 4. Componentes dos mixes de PCR, volumes utilizados e concentração final...26	
Tabela 5. Condições das reações de PCR, primeiro e segundo ciclos.....	26
Tabela 6. Sítios de restrição adicionados aos <i>primers</i> para subclonagem.....	42

RESUMO

As celulases desempenham um papel significativo no processo industrial de degradação da celulose. Neste estudo, foi relatado a clonagem e a expressão heteróloga de uma β -glicosidase recombinante a partir do cDNA de *T. licus licus* (*TlicB-glu1*) e uma β -1,4-endoglucanase recombinante do cDNA de *A. grandis* (*AgraGH45-1*). Os genes foram expressos em sistema heterólogo de *Pichia pastoris*TM. Baseado na predição de proteínas e busca por domínios conservados, foram clonadas as sequências gênicas *TlicB-glu1* com identidade na família GH1 (Glicosil Hidrolase 1) e *AgraGH45-1* com identidade na família GH45 (Glicosil Hidrolase 45). Os transcritos do gene clonado *TlicB-glu1* apresentaram uma expressão gênica diferencial na fase larval adulta e no intestino médio em relação a carcaça e outras fases da vida do inseto. O perfil proteico do sobrenadante da colônia recombinante apresentou uma proteína diferencial de peso molecular estimado de aproximadamente 50 kDa. A enzima recombinante expressa mostrou-se ativa em pH 4,5 a 45 °C, de acordo com a faixa de pH e temperaturas das enzimas comerciais utilizadas nas indústrias. Também apresentou atividade β -glicosídica na degradação da carboximetilcelulose (CMC) e celobiose. No ensaio de açúcares redutores, observou-se a liberação de glicose no meio (4,76 mM) duas vezes superior em relação ao controle (2,13 mM). Os transcritos do gene clonado *AgraGH45-1* apresentaram maior expressão gênica no intestino em ambos os estágios adulto e larval. O perfil protéico apresentou uma proteína de peso molecular estimado de aproximadamente 35 kDa. A enzima recombinante expressa mostrou-se ativa em pH 5,0 a 50 °C na degradação de CMC e Hidroxietilcelulose (HEC). As enzimas recombinantes produzidas neste estudo mostraram-se eficientes na degradação de celuloses amorfas. Estes dados reforçam a importância de identificar a bioquímica digestiva desses insetos-praga, visando o potencial na produção de ferramentas biotecnológicas.

Palavras-chave: *Anthonomus grandis*, clonagem, enzimas digestivas, expressão heteróloga, transcriptoma, *Telchin licus licus*.

ABSTRACT

Cellulases play a significant role in the industrial process in the degradation of cellulose. In this study, we report the cloning and heterologous expression of a recombinant β -glucosidase from the *Telchin licus licus* cDNA (*TlicB-gluc1*), and a recombinant β -1,4-endoglucanase from the *A. grandis* cDNA (*AgraGH45-1*). The genes were expressed in a heterologous *Pichia pastoris*TM. Based on protein prediction and search for conserved domains, the *TlicB-gluc1* gene sequences with identity in the GH1 family (Glycosyl Hydrolase 1) and *AgraGH45-1* with identity in the GH45 family (Glycosyl Hydrolase 45) were cloned. The transcripts of the cloned *TlicB-gluc1* gene showed differential gene expression in the adult larval stage and in the midgut in relation to the carcass and other stages of the insect's life. The protein profile of the recombinant colony supernatant showed a differential protein with an estimated molecular weight of approximately 50 kDa. The recombinant enzyme expressed was active at pH 4,5 at 45 ° C according to the pH range and temperatures of the commercial enzymes used in the industries. It also showed β -glycosidic activity in the degradation of carboxymethylcellulose (CMC) and cellobiose. In the reducing sugar assay, glucose was released into the medium (4,76 mM) twice as high as the control (2,13 mM). The transcripts of the cloned *AgraGH45-1* gene showed higher gene expression in the intestine in both the adult and larval stages. The protein profile showed a protein with an estimated molecular weight of approximately 35 kDa. The expressed recombinant enzyme was active at pH 5,0 at 50 °C in the degradation of CMC and Hydroxyethylcellulose (HEC). The recombinant enzymes produced in this study proved to be efficient in the degradation of amorphous celluloses. These data reinforce the importance of identifying the digestive biochemistry of these insect-pests, aiming its application in the production of biotechnological tools.

Keywords: *Anthonomus grandis*, cloning, digestive enzymes, heterologous expression, transcriptome, *Telchin licus licus*.

1.INTRODUÇÃO

Segundo dados do Banco Mundial e do Fundo Monetário Internacional (FMI), a demanda mundial por energia será o dobro até 2030, principalmente no setor de transportes, induzindo assim, a produção de novas fontes de combustíveis como alternativas ao petróleo (ENGDAHL, 2014; ZILBERMAN & TIMILSINA, 2014). A busca por fontes de energia, a partir de resíduos lignocelulósicos é uma realidade mundial (NANDA; MOHAMMAD; REDDY, 2014). A biomassa ou resíduo lignocelulósico é qualquer material orgânico de origem animal ou vegetal que podem ser transformados ou não industrialmente. Essa matéria-prima parece ser a chave para um fornecimento contínuo de combustíveis renováveis (YUE; YOU; SNYDER, 2014), sem concorrência direta com culturas utilizadas para alimentação (BRETHAUER; STUDER, 2015).

O Brasil é uma referência mundial no que diz respeito à utilização de energia renovável (GANDOLFI et al, 2012). O país ocupa uma posição privilegiada por apresentar grande potencial de cultivo de matérias-primas renováveis, culturas agrícolas de grande extensão (como a cana-de-açúcar), fatores ambientais e pioneirismo na produção de etanol (SANTOS; BORSCHIVER; COUTO, 2011). Cerca de 45% da energia produzida, e 18% dos combustíveis consumidos no Brasil já são renováveis (ANP, 2016). Contudo, a produção de etanol combustível (1ª geração ou 1G) no Brasil, apesar de possuir incentivos fiscais, está sujeita a flutuações de mercado, uma vez que a fonte para a sua produção é a mesma utilizada para a produção de açúcar refinado, ou seja, caldo ou melaço da cana-de-açúcar (PROTASIO et al, 2013).

O etanol 2G, biocombustível de 2ª geração, por sua vez, é produzido a partir da degradação do bagaço da cana-de-açúcar ou de outras fontes de biomassa, com a utilização de enzimas que transformam os açúcares liberados na quebra da biomassa em etanol celulósico. No entanto, os métodos que são utilizados para tratar a biomassa (pré-tratamento), bem como sua conversão em etanol (pós-tratamento), aumentam os custos da produção e, com isso, diminuem a rentabilidade (KOHLHEPP, 2010). De fato, há um crescente interesse na produção de etanol 2G. No entanto, uma eficiente degradação enzimática da celulose em glicose representa o principal entrave a ser alcançado pelas indústrias na comercialização em larga escala deste biocombustível (FERNANDES et al, 2015). Devido ao caráter recalcitrante da celulose, é necessário o pré-tratamento da matéria-prima e de etapas de hidrólise adaptadas, para obter monômeros de açúcar

fermentáveis. Esse processo vem passando por otimização contínua, a fim de aumentar o rendimento e reduzir o custo econômico. Portanto, a descoberta de novas enzimas com alta atividade lignocelulítica ou novas estratégias é extremamente importante (SAADEDDIN, 2014).

Para os processos industriais, a limitação da hidrólise enzimática da celulose ocorre, principalmente, devido à heterogeneidade da parede vegetal e recalcitrância da lignina e da própria celulose, que são fatores que aumentam os custos de todo o processo. Por essas características, torna-se obrigatória a etapa de pré-tratamento da biomassa para facilitar a ação das celulasas no substrato (GALBE; ZACCHI, 2012).

Um dos processos mais importante para a produção do etanol 2G seria a utilização de um coquetel de celulasas de baixo custo e eficientes, que contenham as principais enzimas do complexo celulolítico da família Glicosil Hidrolisase (GH): endoglucanase, exoglucanases (celobiohidrolases) e β -glicosidasases, atuando sinergisticamente na degradação da celulose (BRIJWANI; OBEROI; VADLANI, 2010), ou, futuramente, em enzimas secretadas pelo procarioto *Caldicellulosiruptor bescii*, que demonstraram ter um desempenho superior às misturas de exo- e endoglucanases, comercialmente disponíveis *in vitro* (CHUNG et al, 2015). Membros desse gênero bacteriano termofílico anaeróbico são candidatos emergentes ao bioprocessamento consolidado (CBP), etapa para produção de etanol 2G, porque são capazes de crescer eficientemente em biomassa sem pré-tratamento convencional, tornando-o um cenário atraente para examinar os efeitos da engenharia metabólica (CHA et al, 2013; KIM et al, 2017; LEE et al, 2020).

De fato, as celulasas foram consideradas a terceira maior classe de enzimas utilizadas industrialmente (YOON et al, 2014). Na natureza, as celulasas, por sua vez, são encontradas principalmente nos micro-organismos como fungos e bactérias, que se alimentam de biomassa vegetal, além de insetos como lagartas, besouros, cupins e nematoides de hábitos endofíticos, capazes de degradar a parede vegetal (ZHANG & LYND, 2004). Esses microrganismos/pragas carregam genes, que foram transferidos por transferência horizontal, ocorrido ao longo da evolução, oriundos de bactérias e fungos. Nos genomas de insetos da ordem Lepidoptera e Coleoptera, por exemplo, já foram descritos genes codificadores de enzimas degradadoras de biomassa vegetal (GLASS et al, 2013; KUMAR et al, 2008).

Por essa razão, recentemente os transcriptomas de *Telchin licus licus* e *Anthonomus grandis* (bicudo-do-algodoeiro) foram seqüenciados e podem representar

uma importante fonte de genes que codificam enzimas, com potencial na degradação da parede vegetal. Nesse sentido, torna-se importante elucidar a bioquímica metabólica destes insetos, tendo em vista o potencial para produção de enzimas para fins biotecnológicos (FONSECA et al, 2015). Com base nisso, o presente estudo descreve a clonagem e expressão heteróloga dos genes que codificam uma β -glicosidase recombinante de *T. licus licus* (*TlicB-glu1*) e uma β -1,4-endoglucanase recombinante de *A. grandis* (*AgraGH45-1*).

2. REVISÃO BIBLIOGRÁFICA

2.1. Biomassa vegetal e biocombustíveis

Os biocombustíveis são produzidos a partir de organismos vivos, diferentemente dos combustíveis fósseis. Essa classe de fonte energética renovável pode ser dividida em três categorias: (a) biocombustíveis de primeira geração (1G), produzidos principalmente a partir de açúcares fermentáveis e amido; (b) biocombustíveis de segunda geração (2G), produzidos a partir de biomassa vegetal não-fermentável (celulose e hemicelulose); e (c) biocombustíveis de terceira geração (3G), produzidos a partir de algas e outros micro-organismos (CARDONA & SÁNCHEZ, 2007).

Em relação à produção de etanol 1G, os Estados Unidos da América (EUA) e o Brasil são os maiores produtores mundiais, com um rendimento anual de 36 e 24,7 bilhões de litros, respectivamente (CHEN et al, 2016). A maior fonte de produção de etanol no Brasil é a cana-de-açúcar (*Saccharum spp.*), enquanto o milho (*Zea mays*) é a principal fonte nos EUA (BROWN & BROWN, 2013). Na safra brasileira de 2018, foram produzidos cerca de 632,9 milhões de toneladas de bagaço da cana-de-açúcar (Tabela 1) (CONAB, 2019). Em uma tonelada de cana-de-açúcar, é possível obter 140 Kg de substrato seco, contendo cerca de 43% de celulose, 25% de hemicelulose e 23% de lignina (BRETHAUER & WYMAN, 2010).

Tabela 1. Dados da produção brasileira de etanol, açúcar e bagaço da cana-de-açúcar das safras de 2014 a 2018.

Ano-Safra	Total de Etanol 1G (m ³)	Açúcar (Toneladas)	Bagaço da Cana-de-Açúcar (Toneladas)
2013/2014	28.916.281	35.603.958	637.714.365
2015/2016	30.492.728	37.594.829	666.304.044
2017/2018	26.685.074	33.508.980	632.879.440

Fonte: CONAB, 2019.

O etanol tradicional 1G, produzido no Brasil, é produzido a partir do caldo da cana-de-açúcar. A sacarose contida nesse líquido é fermentada pela levedura *Saccharomyces cerevisiae*, que a transforma em etanol. No entanto, no bagaço e na palha da cana-de-açúcar também contém açúcares, que podem ser aproveitados. O

maior desafio é o acesso a esses polímeros, pois estão associados à estruturas complexas da biomassa (CHEN et al, 2016).

Diferente do etanol 1G, o etanol 2G não é obtido do caldo, mas do bagaço e da palha da cana ou de qualquer outro material vegetal que contenha celulose, no entanto esses materiais não podem ser fermentados diretamente. É preciso passar por um processo de pré-tratamento, para liberar celulose e hemicelulose dos outros componentes da biomassa e, então, promover o processo de hidrólise para gerar açúcar fermentado a etanol. É durante a hidrólise que os coquetéis enzimáticos realizam sucessivas quebras nas longas cadeias de moléculas da celulose, até transformá-las em açúcares simples. Aumentar a eficiência e reduzir o custo desses coquetéis é, portanto, um dos grandes desafios para alavancar a indústria do etanol 2G (BUCKERIDGE & DE SOUZA, 2014; BUCKERIDGE, 2018).

2.1.1. Panorama do etanol 2G

A procura por tecnologias e rotas que possam facilitar e tornar economicamente viável a produção do etanol 2G, em escala comercial, tornou-se um desafio mundial. Entretanto, barreiras agrícolas, tecnológicas e industriais ainda atrasam todas as expectativas para alcançar a meta de produção. Como, por exemplo, os EUA, que não alcançaram 10% de sua capacidade total de produção (LORENZI & ANDRADE, 2019).

O quantitativo de pesquisas científicas e registro de patentes que envolveram estudos com o objetivo de viabilizar essas barreiras, nos últimos anos, cresceram 31% sob a liderança dos EUA e China (LORENZI & ANDRADE, 2019). De acordo com o Centro de Gestão e Estudos Estratégicos (CGEE), vinculado ao Sistema Nacional de Ciência, Tecnologia e Inovação (SNCTI), o etanol 2G foi anunciado como o futuro do fornecimento de fontes de energia para substituir os combustíveis fósseis. Matérias-primas baratas como bagaço de cana-de-açúcar, gramíneas, lascas de madeira, palha de trigo, palha de milho, podem ser utilizados para fins biotecnológicos. Com certeza o açúcar contido nesses materiais lignocelulósicos é o ponto de partida da indústria do futuro (ZILBERMAN & TIMILSINA, 2014).

A transformação dessa biomassa em produtos biotecnológicos apresentou seus primeiros registros em 2004 (PULZ & GROSS, 2004), destacando-se principalmente a produção do etanol 2G, a partir do bagaço de cana-de-açúcar e da palha de milho. Para uma usina produzir por ano 140 milhões de litros de etanol 2G, ela precisa de 350 mil

toneladas de biomassa (ZILBERMAN & TIMILSINA, 2014). Na produção do etanol 1G a partir do milho, por exemplo, utiliza-se uma solução homogênea com enzimas que facilitam a quebra dos açúcares do amido, que foi dissolvido em água. Já na biomassa utilizada na produção do etanol 2G, os açúcares são sólidos e insolúveis, e as enzimas devem atuar nestes substratos. Por essa razão, o sistema 2G é muito mais complexo (JAISWAL et al, 2017).

Com o avanço tecnológico, o consumo de etanol saltará dos atuais 22 bilhões para 68 bilhões de litros em 2020. Segundo a União da Indústria de Cana-de-Açúcar (Unica), a maior organização representativa do setor de açúcar e bioetanol do Brasil, somente com a produção do etanol 1G, essa demanda não será suprida (LORENZI & ANDRADE, 2019).

Já foram anunciadas 61 usinas de etanol 2G no mundo, dentre estas, três brasileiras: GraalBio, Centro de Tecnologia Canavieira (CTC) e Raízen (BUSSAMRA et al, 2015). Em 2013, o relatório publicado pelo Conselho Avançado Americano de Etanol elencou os países com avanços biotecnológicos e capazes de produzir etanol 2G em larga escala: EUA, Canadá, China, Dinamarca, Itália, Alemanha e Espanha (ZILBERMAN & TIMILSINA, 2014). No Relatório de Prospecção do Etanol 2G 2030 (2018), a maioria dos países com compromissos mais ousados são países em desenvolvimento latino-americanos (Argentina, Brasil, México, Paraguai, e Uruguai), e asiáticos (Índia, Indonésia e Filipinas), que apontam para um cenário de forte inserção de etanol 2G no setor de transporte em 2030 (SILVÉRIO et al, 2018).

O preço ideal para a produção do etanol 2G brasileiro é de R\$ 0,80 o litro, alinhado ao custo do etanol 1G (PETROBRAS, 2017). Em 2017, no seminário “Sugar & Ethanol Brazil”, em parceria com o Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE), divulgou que o custo de produção do litro de etanol 1G era próximo de R\$ 1,10, enquanto o de 2G, próximo a R\$ 1,50. A expectativa para 2020 é que fique entre R\$ 0,70 e R\$ 0,50 o litro do etanol 2G (LORENZI & ANDRADE, 2019). Em um panorama geral, as usinas brasileiras não alcançaram suas metas de produção e pouco se sabe sobre os próximos passos dessas empresas. Observa-se que a busca por avanços biotecnológicos é um desafio mundial, uma vez que a produção do etanol 2G alcançou uma produção, não alcançando ainda as metas desejáveis pelas indústrias (NYKO, 2018).

Outro desafio é o período de entressafra, que não permite a operação da usina nos 365 dias do ano, já que a média anual é de 240 dias. A Universidade Federal do

Paraná (UFPR) desenvolveu etanol 2G em pequena escala, a partir do bagaço do eucalipto apenas para estudo experimental, constituindo uma opção de biomassa nos períodos de entressafra da cana (dezembro a março) com objetivo de não deixar as usinas inativas durante esse período (MARTINS et al, 2018).

Outro problema enfrentado nas usinas é o tempo necessário para o processo de hidrólise, o escalonamento, uma barreira que impede a maturação comercial do etanol 2G, segundo dados do Instituto de Biociências de Energia (EBI) da Califórnia. As usinas funcionam de maneira eficiente em escalas muito grandes. Por isso, é fundamental a disponibilidade da biomassa na quantidade suficiente, para aproveitar toda a infraestrutura (LORENZI & ANDRADE, 2019).

Atualmente, as usinas conseguem produzir etanol 2G, a partir do hidrolisado de celulose. A produção do etanol 2G, a partir do hidrolisado de hemicelulose ainda está em desenvolvimento e demanda mais custos (ANTUNES et al, 2017). Por essa razão, pesquisas desenvolvidas nessa área são realizadas na tentativa de desenvolver uma levedura geneticamente modificada, com capacidade de fermentar a hemicelulose (LORENZI & ANDRADE, 2019).

Das rotas tecnológicas mais estudadas, a mais eficaz é a hidrólise enzimática. As enzimas comercializadas pela Novozymes® e Genecor®, e utilizadas nessa etapa são onerosas, além da necessidade de grande quantidade para a maior eficiência. A falta de continuidade do processo industrial faz com que se utilizem muitas enzimas, que não são reutilizadas, devido a ausência de uma técnica de imobilização dessas enzimas (SOUZA et al, 2017). As rotas enzimáticas têm apresentado vantagens importantes sobre as rotas químicas, embora o processo de hidrólise ainda precise de melhorias (CALIL NETO et al, 2013; ANDRADE et al, 2019).

2.2. Embrapa: principais contribuições para a produção de etanol 2G

Após levantamento dos dados publicados no portal da EMBRAPA, órgão ligado ao Ministério da Agricultura, Pecuária e Abastecimento (MAPA), observou-se um verdadeiro esforço para a produção do etanol 2G brasileiro.

A Embrapa tornou-se uma grande aliada na evolução do etanol 2G, desenvolvendo possíveis soluções para os entraves agrícolas, como a produção de plantas transgênicas, identificação de genes responsáveis pela dureza da parede celular e

aumento de digestibilidade da biomassa; filtros para a extração da lignina, pesquisas com outras fontes de biomassa para substituir a cana-de-açúcar nos períodos de entressafra (MACHADO et al, 2016). Na área tecnológica, a Embrapa está investindo na produção de enzimas mais eficientes e específicas, na busca por enzimas celulolíticas de microrganismos não-cultiváveis em laboratório, na produção de organismos geneticamente modificados (OGM), na isomerização de enzimas para digestão dos produtos da hemicelulose, além da imobilização e reutilização das enzimas durante o processo industrial de produção (DAMASO et al, 2012; CARVALHO et al, 2015; CAVALCANTE, 2019).

A Embrapa Agroenergia, onde se encontra um volume maior de pesquisas que visam à produção do etanol 2G, desenvolve estudos, envolvendo o desenvolvimento e inovação de processos de conversão da biomassa, além de uma planta piloto e um núcleo de apoio a culturas energéticas (NACE). Pesquisas voltadas para a produção de biocombustíveis, degradação de biomassa, transformação dos produtos gerados, bem como prospecção, seleção, caracterização e cultivo de microrganismos e desenvolvimento de coquetéis enzimáticos (ARAÚJO et al, 2012; COSTA, 2015). Em 2019, desenvolveu um estudo com metabolômica de leveduras fermentadoras de xilose na identificação de alvos metabólicos para aumento da produção de etanol 2G (CAVALCANTE, 2019).

Desde 2013, a Embrapa Agroenergia incentiva a pesquisa com microrganismos para o desenvolvimento de etanol 2G. Por essa razão, vem reforçando as pesquisas que integram ferramentas de Genômica Funcional, Transcriptômica e Metabolômica (CAMPOS, 2015). Em 2014, anunciou o desenvolvimento de algas geneticamente modificadas (GM) com a utilização de genes da biblioteca genômica desse grupo. Também desenvolveu protocolos de transformação das microalgas acoplados às biorrefinarias, uma vez que o gás carbônico gerado no próprio processo de produção do etanol pode ser injetado no reator e utilizado como fonte de carbono para o crescimento das microalgas. Os cientistas já estruturaram uma coleção com 50 cepas de microalgas e a equipe tem como um dos focos o sequenciamento do genoma das linhagens mais promissoras e na caracterização da biomassa por elas produzida (SANTANA, 2016).

A Embrapa Agroenergia também estuda, desde 2016, as moléculas monoxigenases e as expansinas. As monoxigenases são produzidas apenas por fungos e algumas bactérias, e atuam como auxiliares na degradação da celulose. A degradação da

biomassa com as enzimas do tipo celulases é realizada por reações químicas que utilizam água (hidrólise), enquanto as monoxigenases rompem as cadeias de celulose por oxidação. As expansinas são encontradas nas plantas, em microrganismos e até mesmo em alguns animais. São capazes de afrouxar as fibras sem rompê-las, facilitando o acesso das enzimas à celulose. Foram realizados testes para medir a quantidade de açúcar gerado, aumentando ou reduzindo a quantidade de expansinas e monoxigenases nos coquetéis. Outros testes foram realizados na Embrapa Instrumentação (São Carlos/SP), na tentativa de aperfeiçoar a produção (SANTOS, 2017).

A Embrapa Instrumentação apresentou uma nova proposta para uma rota mais simplificada, mais rápida e menos onerosa. A rota utilizou um coquetel enzimático mais eficiente, com o uso misto de dois fungos: *Trichoderma reesei* e *Aspergillus oryzae*. A rota simplificada é vista pelo grupo como um avanço tecnológico importante para aumentar a eficiência do uso da biomassa vegetal, uma vez que foi possível eliminar três etapas do processo (extração, filtração e centrifugação) (RABELO et al, 2014).

Outro problema relatado pelas usinas de etanol 2G é a falta da reutilização das enzimas, pois no final do processo estas são descartadas (ALVIRA, et al 2010). A Embrapa Agroenergia também trabalhou com pesquisas de imobilização enzimática na tentativa de reaproveitá-las (ALMEIDA et al, 2015).

Em 2018, a Embrapa Agroenergia divulgou uma pesquisa de impacto global. Em parceria com o Reino Unido e EUA, foi identificado um gene endógeno responsável pela feruloilação. A supressão desse gene aumentou a liberação de açúcares em até 60%. Esse resultado poderá contribuir com o desenvolvimento de plantas com paredes celulares mais fáceis de serem digeridas para a produção de etanol 2G (SOUZA et al, 2018).

2.3. Insetos praga: fonte para produção de celulases

Ainda dentro das pesquisas realizadas na Embrapa, o Laboratório de Interação Molecular Planta-Praga, lidera projetos que buscam encontrar fontes de genes que expressam enzimas associadas à degradação da biomassa vegetal. Os genomas abordados neste estudo foram de *T. licus licus* e *A. grandis*, pois é relatado, por estudos filogenéticos, a presença de famílias gênicas que degradam celulose. A praga *T. licus licus* é considerada o principal inseto praga da cana-de-açúcar no Nordeste brasileiro enquanto que *A. grandis* é a principal praga das plantas de algodão, considerada um dos

insetos-praga de maior relevância na cotinocultura (HASUNUMA et al, 2013; (NEGRISOLI JUNIOR, 2015; RIVASTAVA et al, 2018).

A broca-gigante da cana-de-açúcar, *T. licus licus*, é uma lagarta que atinge até cerca de 10 cm centímetros de comprimento. O ciclo de vida do inseto é longo, podendo chegar a dez meses, se desenvolvendo no interior do colmo da cana-de-açúcar, alimentando-se de resíduos lignocelulósicos. Em plantas jovens, as larvas chegam ao rizoma (maior conteúdo de sacarose) e ocasionam o apodrecimento e posterior secagem dos tecidos da planta (PINTO et al, 2006). Esse inseto-praga representa um grande problema fitossanitário em países das Américas do Sul e Central, incluindo o Brasil, Bolívia, Colômbia, Panamá, Trinidad e Tobago e Guiana. No Brasil, *T. licus licus* contribui para as maiores perdas na produção de cana-de-açúcar da região Nordeste, especialmente nos estados de Alagoas, Pernambuco, Paraíba e Rio Grande do Norte (NEGRISOLI JUNIOR, 2015).

O sistema digestivo desses insetos-praga é dividido em 3 partes: **intestino anterior**, que inclui cavidade bucal, faringe, esôfago, papo, pró-ventrículo; **intestino médio**, que inclui o ceco gástrico, ventrículo, membrana peritrófica, espaço ectoperitrófico e endoperitrófico; e **intestino posterior**, que inclui túbulos de Malpighi, íleo, cólon, reto e ânus (Figura 1) (TERRA & FERREIRA, 1994; SUN et al, 2010).

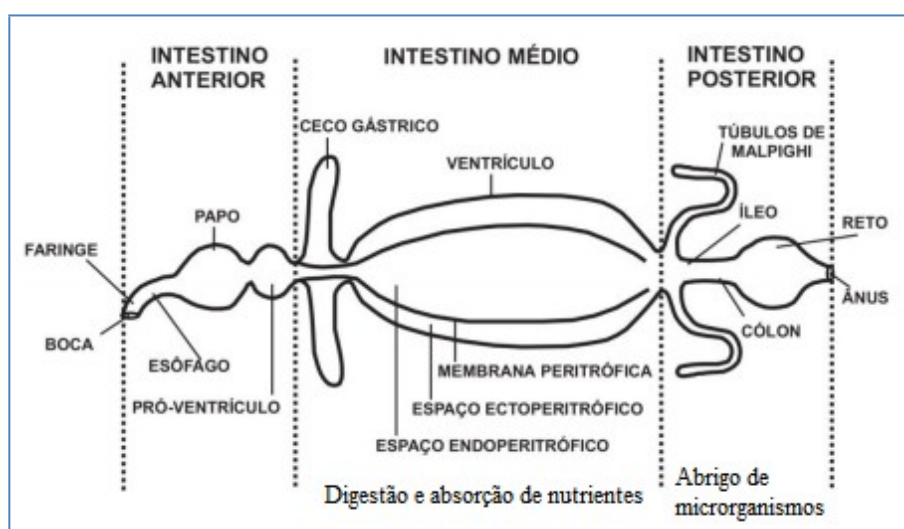


Figura 1. Esquema generalizado do tubo digestivo de insetos (Terra & Ferreira, 1994).

O pró-ventrículo funciona como um triturador, fornecendo um controle da válvula de entrada de alimentos no intestino médio, que é o principal local de digestão e absorção de nutrientes. O intestino médio consiste em um simples tubo ou ventrículo. Na maioria dos insetos, o intestino médio é revestido de uma estrutura quitinosa que separa o espaço endoperitrófico (interior da membrana) do ectoperitrófico (exterior da membrana). O esfíncter separa o intestino médio do intestino posterior. A excreção é realizada pelos túbulos de Malpighi. O intestino posterior, envolvidos na absorção de água e íons, desembocam ânus, podendo abrigar microrganismos que podem auxiliar na digestão da celulose (TERRA & FERREIRA, 1994).

O controle da infestação na cana-de-açúcar é dificultado principalmente pela localização das diferentes fases larvais da praga nas plantas e devido ao fato de o inseto se ocultar entre os tecidos vegetais, formando galerias (Figura 2A). Os danos provocados pelo inseto ocorrem desde a sua forma larval, que tem coloração branca, com uma única mancha castanha e de formato irregular no pronoto, primeiro segmento torácico de insetos (Figura 2B) (MENDONÇA et al, 1996).

Os ovos apresentam cinco arestas salientes, possuem uma estrutura poliédrica de 4 mm de comprimento e coloração variada (Figura 2A1). Ao eclodirem, as larvas penetram no solo para perfurar a base da cana, provocando uma abertura ascendente no colmo, e sua destruição pode chegar a até 1 m de comprimento. Ainda na base do colmo, a lagarta constrói um casulo com as fibras da cana e transforma-se em pupa (Figura 2B1). Após emergirem, os adultos copulam e as fêmeas fazem a postura dos ovos, continuando o ciclo (MENDONÇA et al, 1996).



Figura 2. Ciclo de vida da broca-gigante (*T. licus licus*). A) Fase larval no colmo da cana-de-açúcar (galeria). B) Pronoto. A₁) ovos. B₁) pupa. C₁) orifício de saída. Fonte: MENDONÇA et al, 1996.

A atividade intestinal é a principal interface do inseto com seu ambiente, e as propriedades das enzimas digestivas se correlacionam à sua filogenia. Os processos digestivos desses polímeros no intestino ocorrem em três fases: fase inicial, com o catabolismo de polímeros, por ação das despolimerases (amilases, celulases, hemicelulases e proteinases); fase intermediária, em que os oligômeros resultantes da despolimerização sofrem hidrólise por oligomerases, gerando dímeros de maltose (no caso do amido), celobiose (no caso da celulose) e dipeptídeos (no caso das proteínas); fase final, em que esses dímeros são clivados em monômeros. As β -glicosidases atuam no último passo, quando a celobiose é hidrolisada, produzindo dois monômeros de D-glicose. Além disso, as β -glicosidases têm um papel importante na interação entre inseto e planta, ou seja, uma importante atuação na digestão de carboidratos (MARANA et al, 2000; BENOLIEL, et al 2010).

O intestino médio é o compartimento no qual se observa a ação de enzimas lignocelulolíticas também confirmadas nos coleópteros, principalmente as β -1,4-endoglucanase, que já se encontram registrados no Banco de dados genômico (GenBank) (GIRARD & JOUANIN, 1999; MARTIN, 1987; ZHOU et al, 2007).

O inseto-praga *A. grandis*, bicudo-do-algodoeiro, chama a atenção devido ao seu

elevado poder de destruição e alta capacidade reprodutiva (BELOT, 2015). Este inseto foi detectado no Brasil em fevereiro de 1983, e provoca desde então, enormes prejuízos para os agricultores, nas cadeias produtivas da fibra e seus subprodutos (GONDIM et al, 1999; SILVIE et al, 2001). Este inseto mede até nove milímetros de comprimento e até sete milímetros de envergadura, de coloração castanho-ferruginosa quando jovem e cinza quando mais adulto. A cabeça é alongada para frente, prolongando-se em um rostro que tem metade do seu comprimento fino e recurvado (Figura 3)(BUSOLI et al, 1994). Esta espécie se alimenta de pólen, pois as fêmeas só conseguem desenvolver seus ovos com este substrato, destruindo a planta no estágio reprodutivo. A fase de pupa também se dá nos botões florais e flores (GALLO et al, 1988). O uso de defensivos químicos tem apresentado maior eficácia no controle da praga, mas é oneroso e poluente (MARTIN et al, 1987; WOLFENBERGER et al, 1997).

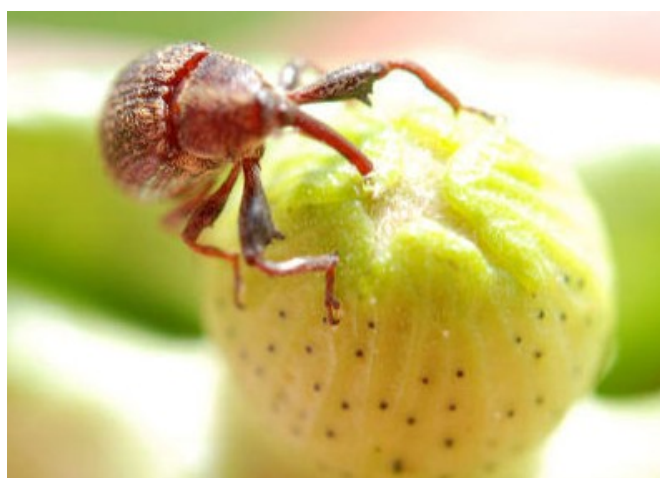


Figura 3. *Anthonomus grandis*- Inseto adulto. Fonte: Sebastião Araújo/Embrapa, 2017.

O sistema digestivo desses insetos vem sendo alvo de muitas pesquisas para a caracterização das enzimas, pois a maioria desses insetos digere com eficiência a parede celular vegetal. Jones e colaboradores (2013) expressaram uma enzima digestiva de *Spodoptera frugiperda* e cuja sequência polipeptídica é de uma β -glicosidase. A atividade glicosídica foi muito bem-descrita em insetos no processo digestivo. Terra e colaboradores (2012) analisaram em *Diabrotica virgifera virgifera* (Coleoptera) a atividade de uma α -glicosidase. Sua ação foi observada por hidrolisar resíduos não-redutores α -1,4 terminais de dextrinas, dissacarídeos e oligossacarídeos, promovendo a liberação de α -D-glicose.

Em estudos de relação taxonômica, foi demonstrado que as famílias de genes

GH1, *GH16*, *GH27* e *GH31*, expressam as enzimas β -1,3-glucanase, α -glicosidase, α -galactosidase e β -1,4-glicosidase presentes em insetos das ordens Diptera, Lepdoptera, Hymenoptera e Hemiptera (PAUCHET et al, 2009).

Segundo Terra e Ferreira (2005), as β -glicosidases dos insetos são muito ativas, uma vez que atuam na digestão intermediária e final da celulose e hemicelulose, e estão envolvidas no aproveitamento dos açúcares, presentes nos glicolipídeos. O coleóptero *Tenebrio molitor*, por exemplo, possui três tipos de β -glicosidases; *S.frugiperda* possui dois tipos; e *D. saccharalis* possui três tipos. Alguns possuem genes que codificam endoglucanases, como foi identificado em coleópteros fitófagos, pertencentes à superfamília Chrysomeloidea (PAUCHET et al, 2009).

Em comparação aos danos provocados pela *D. saccharalis*, o potencial de *T.licus licusem* degradar a biomassa do colmo da cana-de-açúcar é bem maior (Figura4) (FONSECA et al, 2015). Nesse sentido, torna-se importante elucidar a bioquímica metabólica deste inseto, tendo em vista o potencial de degradação de biomassa vegetal a ser explorado biotecnologicamente, adicionalmente ao fato de que somente algumas sequências de DNA se encontram disponíveis para esta espécie no GenBank (<https://www.ncbi.nlm.nih.gov/nuccore/>).



Figura 4: Degradação da biomassa da cana-de-açúcar por lagartas. A) *D. saccharalis* B) *T. licus licus*. C) Colmo da cana-de-açúcar consumido pela lagarta *T. licus licus* (galeria). Fonte: MENDONÇA et al, 1996.

2.4. Transcriptomas

Esforços tem sido aplicado para o desenvolvimento de alternativas biotecnológicas para o controle de pragas. Dentre estes, o estabelecimento de um banco de dados genômico. Como o repertório das enzimas digestivas de *T. licus licus* era desconhecido, Fonseca e colaboradores (2015) desenvolveram uma biblioteca completa de todos os genes expressos pelo inseto, por meio do sequenciamento completo de *expressed sequence tags* (ESTs) obtidas a partir de diferentes estágios do desenvolvimento. As informações geradas foram disponibilizadas para diversos estudos, dentre os quais a busca por enzimas digestivas de *T. licus licus* que atuam na degradação de biomassa vegetal (FONSECA et al, 2015). Com esta análise, foi possível prever que *T. licus licus* é uma fonte adequada de enzimas celulolíticas para uso na degradação de biomassa com fins industriais.

Baseado nos transcriptomas do intestino de *A. grandis* (FIRMINO et al, 2013), uma busca por sequências nucleotídicas conservadas que codificam enzimas celulolíticas foi conduzida sob o taxon Arthropoda na base de dados NCBI-GenBank. Sequências conservadas foram usadas como consulta em busca de sequências similares no banco de dados do cDNA de *A. grandis*. Com bibliotecas de cDNA e ESTs, muitas enzimas de insetos foram descobertas e avaliadas. O desenvolvimento de transcriptomas tem permitido identificar genes envolvidos na degradação da biomassa vegetal (WILLIS et al, 2010).

2.5. Celulases

A diversidade das aplicações de celulases na indústria vem se destacando na reciclagem de papel, produção de detergente, extração de sucos, geração de produtos dermatológicos e pigmentos, extração de óleos vegetais, processamento de algodão, processamento de alimentos, aplicações farmacêuticas e produção de biocombustíveis (HUANG et al, 2015).

Para utilizar seus componentes, a biomassa passa por um pré-tratamento, que consiste em procedimentos mecânicos, químicos e térmicos, e tem como objetivo maximizar a exposição dos compostos da biomassa às celulases, que realizam seu processamento. Um pré-tratamento eficiente pode ter um rendimento de até 80% na liberação de monossacarídeos por celulases e xilanases. A sacarificação enzimática é

a etapa que requer a liberação do maior número possível de monossacarídeos para fermentação. As enzimas requisitadas nesta etapa devem, portanto, apresentar eficiência e estabilidade (JORDAN et al, 2012). Contudo, no processo de obtenção biotecnológica de etanol 2G observam-se dois principais entraves: os métodos de pré-tratamento impostos à biomassa e o alto custo de produção das celulases (SINGHANIA et al, 2013). A celulose da biomassa vegetal é um homopolissacarídeo linear de glicose cujos resíduos se encontram unidos por ligações glicosídicas do tipo β -(1,4). Os resíduos de glicose se alternam verticalmente formando uma unidade repetitiva fundamental da celulose, o dissacarídeo celobiose (Figura 5). A celulose encontra-se num estado dito “cristalino”, responsável pela sua resistência à hidrólise enzimática (SINGHANIA et al, 2013).

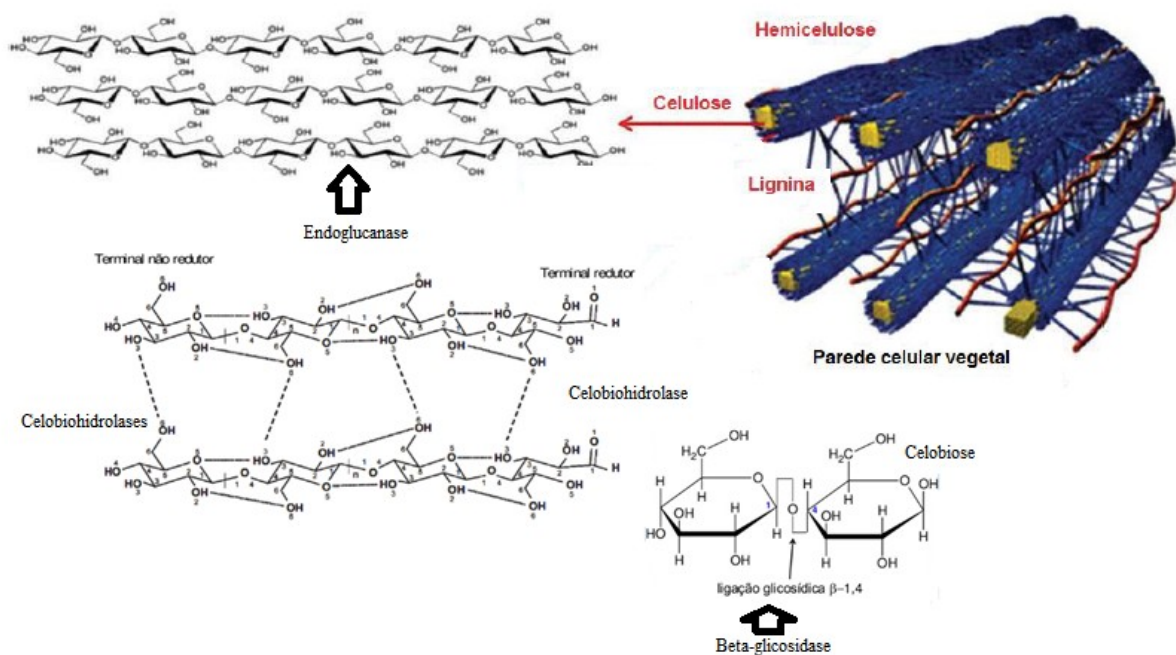


Figura 5. Representação esquemática da degradação da celulose. Endoglucanases clivam internamente o polímero de celulose; Celobiohidrolases (exoglucanases) clivam as segundas ligações glicosídicas a partir das extremidades da celulose parcialmente digerida; β -glicosidases clivam a ligação β -1,4 que unem os monômeros de glicose. Adaptado de VOLYNETS; EIN-MOZAFFARI; DAHMAN (2017).

A destruição de toda a estrutura cristalina da celulose até a obtenção de um produto fermentável (glicose) pode ser alcançada a partir da atividade de três celulases: celobiohidrolases (exocelulases), endoglucanases e β -glicosidases. As endoglucanases

realizam clivagens na parte interna da fibra de celulose, liberando fragmentos menores formados por poucas unidades de glicose, os oligossacarídeos. As exoglucanases realizam clivagens nas extremidades das fibras e liberam celobiose. As β -glicosidases realizam clivagens entre duas unidades de glicose que formam a celobiose, liberando unidades de glicose livre (JORDAN et al, 2012).

Durante a fase inicial da produção dos biocombustíveis, as indústrias realizam a suplementação dessas enzimas para a conversão total em etanol. Dessa forma, surge a necessidade de prospectar enzimas específicas para a hidrólise das ligações β -glicosídicas da celulose, o que constitui outro fator limitante, embora nos últimos anos a produção eficiente de celulasas tenha aumentado mais de 10 vezes, diminuindo os custos do processo de sacarificação (JORDAN et al, 2012).

As enzimas comerciais como a Novo188 (Novozymes, EUA), uma β -glicosidase de *Aspergillus niger*, apresenta atividade industrial em faixa de pH 5,0 a 6,0 a 50 °C.; a β -glicosidase G4511 de amêndoas e de *Escherichia coli* (Sigma-Aldrich) possui atividade em faixa de pH 4,0 a 5,0 a 37 °C; com a liberação de 1 μ mol de glicose por minuto. Atualmente, essas enzimas não atendem, portanto, as expectativas na produção do etanol 2G em escala comercial (SINGHANIA et al, 2012).

A maioria das celulasas não é estável em altas temperaturas e muitas pesquisas estão sendo realizadas com bactérias para obter celulasas termostáticas (YEOMAN et al, 2010). As bactérias *Caldicellulosiruptor bescii* utilizam os carboidratos da lignocelulose para seu crescimento. São tecnologicamente importantes porque produzem hidrolases glicosídicas de múltiplos domínios, eficientes na degradação de celulose microcristalina e hemiceluloses. Elas se adaptaram com sucesso para se ligar especificamente a substratos lignocelulósicos por meio de domínios de homologia de camada superficial (SLH) associados à hidrolases de glicosídeo e proteínas de ligação únicas (tapirinas) presentes apenas nessas bactérias. Várias espécies desse gênero podem degradar prontamente a lignocelulose, no entanto, são necessárias melhorias na quantidade do carboidrato liberado para alcançar o potencial biotecnológico desses organismos (LEE et al, 2020).

O CelA é a enzima mais abundante secretada por essa bactéria, e demonstrou ter um desempenho superior às misturas de exo- e endoglucanases comercialmente disponíveis *in vitro*. O CelA contém uma endoglucanase da família GH9 e uma exoglucanase da família GH48 conhecidas por serem sinérgicas em sua atividade. São conectadas por três domínios de ligação à celulose. Repetidos resíduos de aspartato

foram introduzidos nas extremidades *N*- terminais dos domínios CelA GH9 e GH48 para melhorar a eficiência de secreção e eficiência catalítica. Apresentou uma maior atividade em CMC ($0,81 \pm 0,03$ mg/mL) (KIM et al, 2017).

2.5.1. β -glicosidase industrial

As glicosidases são enzimas capazes de hidrolisar ligações glicosídicas, sejam estas a α (α -glicosidases) ou β (β -glicosidases) (Figura 6). As glicosidases atuam em variadas funções e fazem parte da família das glicosil-hidrolases (GH). São utilizadas em processos biotecnológicos e alvos de pesquisas para desenvolvimento de novas tecnologias para a indústria alimentícia (clareamento de sucos de frutas, aumentar a qualidade nutritiva de alimentos fermentados, remoção da parede celular para liberação de fragrâncias e para extração de ágar a partir de algas), médica (prevenção de certas doenças como câncer, desordens cardiovasculares e sintomas pós-menopausa), agroindustrial (retirada de microfibrilas ou para restaurar o brilho em tecidos de algodão) e de biocombustíveis (BAYER et al, 1998); (HUANG et al, 2015). Há também registros de suas atividades nas reações de transglicosilação de ligações β -glicosídicas em conjugados glicosídicos. Hidrolisam substratos com glicose, galactose, fucose, manose, xilose e são nomeadas de acordo com os monômeros liberados. Também são denominadas myrosinase, thioglucosidase e sinigrinase (VOLYNETS; EIN-MOZAFFARI; DAHMAN 2017).

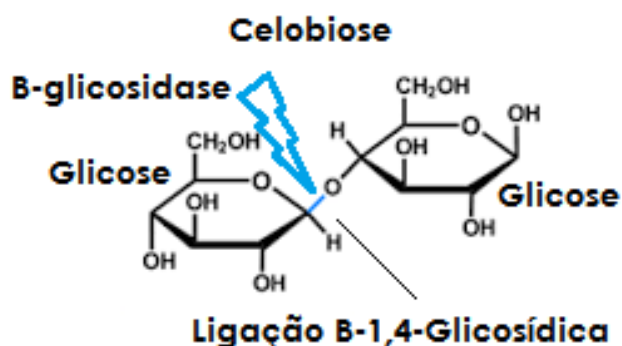


Figura 6. Representação esquemática da hidrólise da ligação β -1,4 da celobiose. Adaptado de VOLYNETS; EIN-MOZAFFARI; DAHMAN (2017).

Os coquetéis enzimáticos utilizados na indústria contêm celulasas do tipo β -glicosidases, que são fundamentais para a hidrólise completa da celulose (SINGHANIA et al, 2010). A demanda por tais enzimas é elevada porque as mesmas são específicas e não produzem resíduos tóxicos. A utilização de enzimas em processos industriais tem aumentado ao longo dos últimos anos. Entretanto, seu alto custo representa, ainda, um obstáculo para o avanço de alguns processos industriais (ALVIRA, et al 2010). Os coquetéis são fornecidos pelas empresas Genencor® e Novozymes®, esta com uma fábrica no Brasil (Curitiba-PR), Novozymes Latin America Ltda (ZÚÑIGA, 2010).

Outra abordagem alternativa com o uso de celulasas seria o uso do bioprocessamento consolidado (CBP), no qual o organismo fermentativo também é responsável pela produção das enzimas solubilizantes. Bactérias *C.besccii* são capazes de fermentar todos os açúcares C5 e C6 primários a partir da biomassa vegetal e são as bactérias celulolíticas mais termofílicas conhecidas, com temperatura de crescimento ideal entre 78 ° C e 80 ° C. Podem crescer e degradar a biomassa contendo alto teor de lignina, bem como a celulose altamente cristalina sem pré-tratamento convencional, aumentando a possibilidade de melhoria econômica adicional da produção de biocombustíveis a partir da biomassa vegetal, reduzindo ou eliminando a etapa de pré-tratamento. No entanto, ainda são necessários avanços de engenharia metabólica para melhorar a produção de biocombustíveis nesse gênero procaríoto (CHA et al, 2013).

2.5.2. Endoglucanase industrial

A Endo-1,4- β -glucanase (EG) é a enzima do complexo celulolítico responsável por iniciar a hidrólise da molécula de celulose. Hidrolisa randomicamente as regiões internas da estrutura amorfa da fibra, liberando oligossacarídeos e novos terminais, sendo um redutor (quando a glicose possui uma hidroxila heterosídica livre) e um não redutor (quando a hidroxila heterosídica da molécula da extremidade participa de ligação com a glicose adjacente). Geralmente sua ação é avaliada pela redução da viscosidade na solução de CMC (LYND et al, 2002).

Um das técnicas que são utilizadas na produção de celulasas e outras enzimas é a fermentação em estado sólido (FES), que consiste na produção de enzimas por fungos filamentosos (SAQIB et al, 2012). Essa produção de endoglucanase por fungos foi amplamente estudada, principalmente nos fungos filamentosos *Trichoderma*, *Aspergillus*, *Fusarium* e *Penicillium* observados na tabela 2 (CHELLAPANDI, 2008).

Tabela 2: Principais microrganismos utilizados na produção de endoglucanase por fermentação em estado sólido.

Resíduos	Microrganismos	Fermentação (horas)	Temperatura de Incubação (° C)	Referências
Bagaçõ-da-cana	<i>Myceliophthora sp.</i>	18	50	Zanelato et al, 2012
Pera espinhosa	<i>Aspergillus niger</i>	72	32	Pirota et al, 2013
Ervas chinesas	<i>Trichoderma virride</i>	72	30	Longwei et al, 2014
Farelo de trigo	<i>Aspergillus oryzae</i>	72	35	Pirota et al, 2016

2.6. Sistema *Pichia pastoris*TM (*Komagataella phaffii*) utilizado para expressão de proteínas recombinantes

O sistema procarioto de expressão é útil por produzir proteínas recombinantes, embora muitas vezes as proteínas eucariotas produzidas sejam instáveis e não apresentem atividade biológica. A *Escherichia coli* é um sistema bacteriano amplamente utilizado, mas não realiza modificações pós-traducionais. Por essa razão, a preferência de sistema de expressão para enzimas nas indústrias são os modelos eucariotos. Além do crescimento rápido, elas podem secretar e realizar modificações pós-traducionais, além dos altos níveis de expressão proteica obtidos (WYSOCKA-KAPCINSKA, et al 2010).

Há mais de 20 anos, a expressão heteróloga baseada em leveduras tem sido utilizada na produção e secreção de proteínas recombinantes para uso humano (ex.: insulina e vacina da hepatite B), animais, plantas, bactérias, outros (LIU et al, 2012). A capacidade de algumas leveduras crescerem em metanol foi descoberta e a partir de então essas leveduras foram utilizadas para produzir proteínas celulares simples com objetivo de alimentar animais. Nessa época, a indústria Phillips Petroleum desenvolveu meios de cultura e métodos para crescer *Pichia pastoris* em metanol de maneira contínua e com altas densidades celulares (LIU et al, 2012). Portanto são leveduras facultativas metilotróficas, capazes de usar compostos de carbono único (por exemplo, metanol) ou compostos multi-carbono sem ligações carbono-carbono (éter dimetílico) como fonte de alimento. São do grupo dos ascomicetos, homotáticas (autoférteis) e haploides (KURTZMAN et al, 2011).

Após análises filogenéticas, o gênero *Pichia* foi transferido para o gênero *Komagataella* e divididas em três espécies: *K. pastoris*, *K. phaffii* e *K. pseudopastoris*.

(DE SCHUTTER et al, 2009). Nesse estudo o nome *P.pastoris* está presente para indicar a marca comercial da levedura.

Assim, o sucesso da expressão em *P.pastoris* é devido à sua alta taxa de crescimento, baixo custo e acurácia de modificações pós-traducionais, como clivagem proteolítica, glicosilação e formação de ligações dissulfeto. As densidades celulares podem alcançar até 200g/L de peso seco durante uma fermentação. Além disso, a produção de proteínas recombinantes extracelulares facilita o processo de purificação da proteína de interesse (LIU et al, 2012). No entanto, a manutenção da proteína heteróloga na levedura não é duradoura e não é possível observar os clones transformantes. É necessário o uso de outras estratégias moleculares para confirmar a inserção do gene de interesse no genoma da levedura.

3. JUSTIFICATIVA

As demandas biotecnológicas vêm avançando a cada dia, e os processos industriais precisam acompanhar o crescimento tecnológico, impulsionando cada vez mais para a produção de novas enzimas, para atender o mercado industrial, assim como o aprimoramento das técnicas envolvidas. Setores industriais que utilizam a biomassa vegetal como matérias-primas necessitam de novas fontes enzimáticas ou a readequação das etapas de produção para alcançar sua total capacidade de produção, uma vez que, a disponibilidade de biomassa para esse fim, cada dia se torna mais acessível.

Diante disso, a busca por novas fontes de genes para a produção de enzimas celulolíticas é essencial para contribuir com os avanços biotecnológicos. Além dos modelos fúngicos, que já são utilizados na indústria, os insetos praga também apresentam potencial enzimático na degradação da celulose. Aliado a ferramentas moleculares e estudos genômicos, é possível prever a quantidade de sequências gênicas, presentes no genoma desses insetos, que estão envolvidos na degradação da biomassa vegetal. Essa investigação para identificar as enzimas produzidas por insetos praga é viável, devido à disponibilidade de tecnologia envolvendo transcriptomas, expressão heteróloga de proteínas, e substratos sintéticos, utilizados na detecção da atividade enzimática. A adaptação das enzimas digestivas dos insetos praga pode ser um diferencial na geração de celulases recombinantes.

4. OBJETIVOS

4.1. Geral

O presente estudo teve como objetivo realizar a expressão de celulases recombinante dos insetos praga, *T. licus licus* e *A. grandis*, em sistema heterólogo de *P. pastoris*TM, além de caracterizar suas atividades biológicas.

4.2. Específicos

Como objetivos específicos, destacam-se:

- Construir vetores de expressão recombinantes com sequências gênicas relacionadas com a degradação da biomassa vegetal de *T. licus licuse A. grandis*;
- Transformar a levedura *P.pastoris*TM, e selecionar os clones que expressam celulases recombinantes;
- Analisar a secreção das celulases no sobrenadante, por meio de gel desnaturante e *western blot*;
- Comparar a atividade celulolítica recombinante com a atividade da levedura transformada com vetor vazio;

5. MATERIAL E MÉTODOS

5.1. Seleção de contigs dos transcriptomas de *T. licus licus* e *A. grandis*

Com base na biblioteca genômica dos insetos-praga foi realizada uma busca por sequências de genes codificadores de enzimas, utilizando bibliotecas gênicas de artrópodos, envolvidos com a degradação de biomassa vegetal, nos bancos de dados do *National Center for Biotechnology Information* - NCBI (<http://www.ncbi.nlm.nih.gov>). As sequências selecionadas foram analisadas no servidor ORF FINDER (www.ncbi.nlm.nih.gov/orffinder), traduzidas *in silico* na ferramenta ExPASy Translate Tool (<https://web.expasy.org/translate/>), e a análise de similaridade aos genes buscados no GenBank foi conduzida no BLASTX (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) e Clustal Omega (<http://www.clustal.org/omega/>).

Após selecionar as sequências, foram desenhados os iniciadores/*primers* (<http://perlprimer.sourceforge.net/download.html>), *forward* e *reverse*, no servidor *Primer Generator* (http://www.bioinformatics.org/sms/rev_comp.html) e preditas as suas temperaturas de anelamento na ferramenta *Oligo Analyzer* (<https://www.idtdna.com/calc/analyzer>). Adicionalmente, a presença do peptídeo-sinal nativo foi predita no recurso SIGNALP 4.1 SERVER (<http://www.cbs.dtu.dk/services/SignalP/>).

Os *primers* para a clonagem da *Open Reading Frame* (ORF) dos genes selecionados, a partir do cDNA genômico de *T. licus licus* e *A. grandis*, foram utilizados para amplificação via PCR (*Polymerase Chain Reaction*) (Tabela 3).

5.2. Extração do RNA total dos insetos e isolamento dos genes de interesse

Larvas e Insetos adultos de *T. licus licus* foram cedidos pela Embrapa Tabuleiros Costeiros-Alagoas. As fêmeas foram mantidas em gaiolas entomológicas, e as larvas foram individualizadas e alimentadas com fragmentos de cana-de-açúcar. Os insetos de *A. grandis* foram obtidos de uma colônia mantida no Laboratório de Criação de Insetos da Embrapa Recursos Genéticos e Biotecnologia, em Brasília, DF. Ovos, larvas, pupas e adultos foram mantidos sob condições controladas de temperatura (26 ± 2 °C), umidade relativa do ar ($70 \pm 10\%$) e 12 h: 12 h luz: fotoperíodo escuro. Os insetos foram alimentados diariamente com dieta artificial.

Com uma pinça e bisturi, foram retirados os intestinos e as carcaças.

Imediatamente, os intestinos e as carcaças foram submersos em nitrogênio líquido e separados em tubos de microcentrífuga, estocados a -80°C . Após maceração com almofariz e pistilo, para cada 50 a 100 mg de tecido pulverizado com nitrogênio líquido adicionou-se 1 mL do reagente de trizol e 10 μL de β -mercaptoetanol, homogeneizando lentamente. O sistema foi centrifugado a 12.000 g por 10 minutos a 4°C . Transferiu-se o sobrenadante para outros tubos de microcentrífuga, que ficaram incubados por 5 minutos a temperatura ambiente. Acrescentaram-se 200 μL de clorofórmio, agitou-se o sistema no vortex, incubando-o por mais 3 minutos a temperatura ambiente. O sistema foi centrifugado a 12.000 g por 15 minutos a 4°C , transferindo-se 500 μL da fase aquosa para novos tubos de microcentrífuga e adicionando-se o mesmo volume de álcool isopropílico. Homogeneizou-se o sistema por 10 minutos a temperatura ambiente.

Centrifugou-se novamente o sistema a 12.000 g por 10 minutos, descartando o sobrenadante, e o *pellet* foi lavado com 1 mL de etanol a 75% por 2 vezes. Centrifugou-se a 7.500 g por 10 minutos, descartando o sobrenadante. O *pellet* repousou por 15 minutos a temperatura ambiente na câmara de fluxo laminar VECO até se evaporar todo o etanol. Ressuspendeu-se o *pellet* em 30 μL de água Milli-Q e retiraram-se 3 μL para quantificar a amostra em espectrofotômetro NanoVue Plus® (GE Healthcare, EUA), de acordo com o protocolo do fabricante. Aliquotas de 25 μL foram estocadas a -80°C .

Aplicou-se uma amostra de 2 μL do extrato de RNA total em gel de agarose 1% (Gel Doc XR-Foto documentador de géis), com marcador 1 kb Ladder Plus (Gibco) para observar a qualidade e a integridade do RNA via eletroforese. Confirmada a integridade das amostras, foi realizada a síntese do cDNA a partir do RNA mensageiro via transcrição reversa. As reações foram catalisadas pela enzima transcriptase reversa (RT).

Assim, utilizaram-se 2 μg do RNA total adicionando-se 2,5 μL do *primer* oligo (dT) a 10 mM e água Milli-Q estéril. As reações foram conduzidas em tubo de 0,2 mL contendo 2 $\mu\text{g}/\mu\text{L}$ do RNA total, 10 μL *Buffer* 5X (Invitrogen: 250 mM tris-HCl, 375 mM KCl, 15 mM cloreto de magnésio, pH 8,3), 0,2 M DTT (100 mM), 0,5 mM de dNTP (10 mM) e 200 U da enzima MMLV (200 U/ μL), completando o volume final para 50 μL com água Milli-Q. Em seguida, foi realizada uma PCR para confirmar a qualidade do cDNA, separado em gel de agarose 1% via eletroforese. As PCRs foram realizadas em tubos de 0,2 mL para cada amostra, contendo o cDNA e os componentes da reação de amplificação (Tabela 4), completando o volume final para 15 μL com água Milli-Q. O segundo ciclo de PCR consistiu em 1 μL (diluído 1:20) do molde do primeiro ciclo de

PCR com os sistemas de reações preparados, as PCRs foram conduzidas em termocicladores T100 Thermal Cycler (Bio-Rad) sob as condições descritas na Tabela 5, sendo as mesmas para ambos os ciclos.

Tabela 4: Componentes dos mixes de PCR, volumes utilizados e concentração final.

Componentes	Volume	Concentração Final
Buffer 10X (200 mM tris-HCl, 500 mM KCl, 20 mM MgCl ₂ , pH 8,5)	1,5 µL	1X
dNTP	0,3 µL	0,2 mM
MgCl ₂ (<i>Taq</i> DNA Polimerase 500U) ou MgSO ₄ (<i>Taq</i> DNA Polimerase Platinum®)	0,9 µL	1,5 mM/µL
Primer Forward específico do contigou oligo(dT)	0,3 µL	10 mM
Primer Reverse específico do contig	0,3 µL	10 mM
<i>Taq</i> DNA Polimerase Platinum High Fidelity® ou <i>Taq</i> DNA Polimerase 5	0,15 µL	2U/µL
cDNA molde	1 µL	20 ng/µL
H ₂ O Milli-Q autoclavada	q.s.p. 10,55	-
Volume Final	15 µL	-

Tabela 5: Condições das reações de PCR, primeiro e segundo ciclos.

Amplicon	Etapas	Temperaturas
<i>T. licus licus</i>	Desnaturação Inicial	95°C por 5 minutos
	Desnaturação	95°C por 30 segundos
	Anelamento	58°C por 30 segundos
	Extensão	72°C por 1 minuto e 40 segundos
		30 ciclos
	Extensão Final	72 °C por 5 minutos
<i>A. grandis</i>	Desnaturação Inicial	95 °C por 5 minutos
	Desnaturação	95 °C por 30 segundos
	Anelamento	60 °C por 30 segundos
	Extensão	72 °C por 1 minuto e 40 segundos
		30 ciclos
	Extensão Final	72 °C por 5 minutos

Os produtos de PCR foram submetidos à corrida eletroforética em cuba horizontal a 80V, por 2 horas, em gel de agarose a 1%, imerso em TBE 1X, (TBE 10X: 80 mM Tris-base, 40 mM ácido bórico anidro e 2,5 mM de EDTA para 1 L de água MilliQ, pH 8,3), corado com brometo de etídio. Em seguida, os fragmentos de PCR foram analisados no software Gel Doc XR, por meio do fotodocumentador de géis Bio-Rad. A identificação de cada banda foi realizada por comparação com o marcador de massa molecular de 100 pb Ladder Plus (Gibco).

5.3. Quantificação da expressão dos transcritos por qRT-PCR

Foram dissecadas as carcaças separando-se do intestino médio, de cada fase dos insetos: ovo, estágio larval inicial, estágio larval final, pupa e adulto. O RNA total foi extraído com o reagente Trizol e em 2 µg de cada RNA foi adicionado 1 µL de DNase I (Thermo Fisher Scientific, USA). Confirmada a integridade da amostra, foi realizada PCR quantitativa de transcrição reversa (qPCR-RT). A análise da expressão gênica foi realizada em triplicatas biológicas e técnicas, obtidas por amostras independentes de cDNA. Os *primers* (Tabela 3) foram escolhidos como genes normalizadores endógenos e a expressão relativa foi calculada usando o método $2^{-\Delta Ct}$. Para as reações de amplificação de *T. licus licus*, os seguintes passos foram adaptados: 2 minutos a 50 °C, 10 minutos a 95 °C e 40 ciclos a 94 °C durante 15 segundos e 60 °C durante 1 minuto. Para a amplificação de *A. grandis* foram usados *primers* específicos e de dois genes de referência (GAPDH e β -tubulina) (Tabela 3). Cada reação RT-qPCR incluiu 5,0 µL da Mistura Principal qPCR GoTaq® 2X (Promega, EUA) 2,6 µL de água livre de nuclease, 0,2 µL (0,2 µM) de *primers* forward e reverse e 2,0 µL de cDNA (diluído 1:20), para um volume total de 10 µL. Os ensaios de RT-qPCR foram realizados em um termociclador CFX96 (Bio-Rad, EUA) sob as seguintes condições: incubação a 95 °C por 20 s, seguida de 40 ciclos a 95 °C por 3 s e 60 °C por 30 s.

5.4. Clonagem dos genes *TlicBgluc1* e *AgraGH45-1*

Uma vez observados os fragmentos amplificados nos tamanhos esperados e também sua qualidade, os amplicons sem o peptídeo-sinal nativo foram excisados como fragmentos do gel e recuperados com auxílio de lâmina de bisturi de aço carbono estéril (Wiltex Plus), sendo purificados utilizando o *Kit Promega™ Wizard™ Genomic DNA Purification*. Após, a eletroforese foi repetida para checagem da presença DNA

purificado, que em seguida foi quantificado em espectrofotômetro NanoVue Plus® (GE Healthcare, EUA), de acordo com o protocolo do fabricante. Para a reação de ligação do inserto no vetor, foram utilizados 1 µL do vetor (30 ng µL⁻¹) *pGEM-T-Easy Vector Systems* (Promega) (Figura 7), 6 µL do *Buffer 2X*, 1 µL da enzima T4 DNA Ligase (6 U) e 3 µL do inserto (60 ng µL⁻¹), completando com água Milli-Q para 10 µL de volume final.

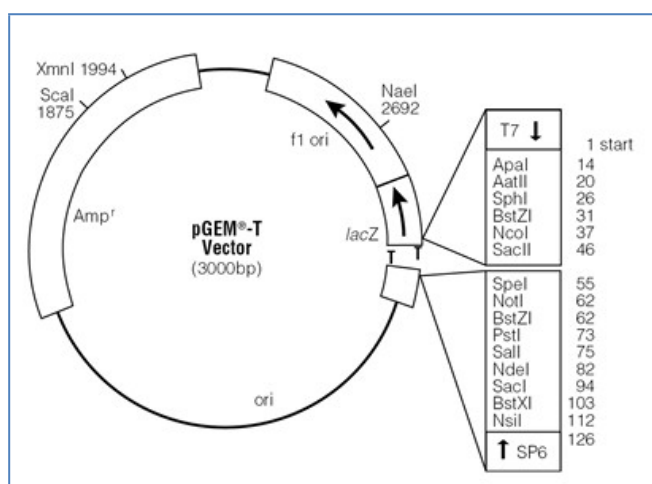


Figura 7: Representação Esquemática do Vetor de clonagem pGEM®-T Easy (Promega). Fonte: Technical Manual Promega.

5.5. Preparação e transformação das células bacterianas

Células termocompetentes de *E.coli* da linhagem *XL1blue* foram cultivadas em 5mL de meio LB(Luria Bertani) líquido (extrato de levedura a 0,5%, peptona de caseína a 1% e NaCl a 1%,pH 7,2) estéril incubadas a 37 °C por 12 horas sob agitação de 250 rpm.

Após crescimento, foi retirado 1 mL do pré-inóculo e adicionado em 30 mL de meio LB líquido, incubando-se a 37 °C sob agitação de 250 rpm. Ao atingir a OD₆₀₀= 1,1 ~ 1,3, as células foram centrifugadas a 3.000 g por 10 minutos a 4°C descartando-se o sobrenadante e ressuspensas em 10 mL de CaCl₂ a 100mM estéril e gelado. Outra centrifugação foi realizada sob as mesmas condições, e as células foram ressuspensas em solução estéril de CaCl₂ a 100 mM e glicerol 1%. Aliquotas de 50 µL foram preparadas e utilizadas para transformação bacteriana via choque térmico. Para transformação, utilizaram-se 50 µL das células competentes previamente preparadas.

Foram adicionados 10 µL da ligação gene-vetor e o sistema mantido em gelo por 30 minutos para cada gene. Para realizar o choque térmico, as células foram incubadas a 42 °C por 45 segundos em banho-maria, e imediatamente transferidas para o gelo, mantendo-as por mais 2 minutos. Adicionaram-se 500 µL de meio LB líquido previamente aquecido a 42 °C no tubo onde estavam as células submetidas à transformação, incubando-se o sistema em shaker por 1 hora a 37 °C e 200 rpm. Passado esse tempo, as células foram plaqueadas em meio LB sólido (extrato de levedura a 0,5 %, peptona de caseína a 1 %, NaCl a 1%, pH 7,2 e agar bacteriológico 1,5%) estéril, acrescido de ampicilina a 50 µg/mL, também se adicionou 100 µL de IPTG (100 mM) indutor isopropil-β-D-tiogalactopiransídeo (Sigma-Aldrich) e 20 µL de X-Gal (1mg/mL)(Sigma-Aldrich) para seleção de clones transformantes. Placas foram incubadas a 37 °C por 16 h. Com o crescimento de colônias brancas e azuis, fez-se a repicagem das colônias brancas em outra placa com meio LB contendo o mesmo antibiótico descrito, mas sem IPTG e X-Gal, incubando-se sob as mesmas condições de crescimento.

5.6. Preparação de DNA plasmidial e sequenciamento

Após crescimento das colônias, procedeu-se com a amplificação e extração plasmidial para a excisão dos insertos e posterior subclonagem no vetor de expressão. Em 5 mL de meio LB líquido com o mesmo antibiótico e concentração, em tubos de 50 mL, cada colônia repicada foi inoculada e cultivada a 37 °C por 16h. Os tubos foram centrifugados a 5.000 rpm a 4 °C por 5 minutos. Descartou-se o sobrenadante e o *pellet* foi ressuspensionado em 300 µL de solução P1 gelada (pH 8,0; 100 µg/mL RNase A; EDTA 10 mM), e transferidos individualmente para tubos de 1,5 mL devidamente identificados. Adicionaram-se 300 µL de solução P2 (pH 5,5; 200 mM NaOH, 1% SDS) a temperatura ambiente. Os tubos foram invertidos de 5 a 10 vezes mantidos em repouso por 5 minutos em temperatura ambiente. Adicionaram-se 300 µL de solução P3 gelada (pH 5,5; acetato de potássio 3 M). Os tubos foram invertidos novamente de 5 a 10 vezes e incubados em gelo por mais 30 minutos. Os tubos foram centrifugados a 14.000 rpm a 4°C por 30 minutos e os sobrenadantes (800 µL) transferidos para outros tubos, nos quais foram adicionados 600 µL de clorofil (clorofórmio e álcool isopropílico, 24:1) e agitados em vórtex por 1 minuto. Foi realizada a mesma centrifugação e os sobrenadantes (700 µL) transferidos para outros tubos. Adicionaram-

se 600 µL de clorofil e agitou-se o sistema em vórtex por 1 minuto, centrifugando a 11.000 rpm a 4°C por 5 minutos. Os sobrenadantes foram novamente coletados e precipitados com a adição de 600 µL de isopropanol, seguido de incubação em gelo por 30 minutos e centrifugados. Os sobrenadantes foram descartados e os *pellets* lavados com 300 µL de etanol gelado a 70%. Após centrifugação a 11.000 rpm a 4°C por 15 minutos, desprezou-se o etanol e os *pellets* foram liofilizados em *speed vac*RVC2-18 CD plus. Após a evaporação de todo etanol, *opellet* foi ressuspensão em 30 µL de água Milli-Q acrescidos de 1 µL de RNase mantendo o sistema a 37 °C por 16 horas.

Para confirmar a clonagem foi realizada a digestão com as endonucleases de restrição específicas para cada gen e utilizados 1 µL do DNA plasmidial, acrescido de 0,5 µL de cada enzima de restrição e 4 µL de tampão 10X (Thermo Scientific), completando o volume final para 40 µL com água Milli-Q, incubadas a 37 °C por 3 horas. Os vetores também foram quantificados e sequenciados. Os fragmentos derivados das digestões foram separados em gel de agarose a 1%, imerso em TBE 1X e corado com brometo de etídio. A identificação de cada banda ocorreu por comparação com marcador de massa molecular de 100 pb Ladder Plus (Gibco). As reações de sequenciamento foram realizadas com 100 a 200ng de plasmídeo, quantificado por espectrofotometria utilizando o GeneQuant RNA/DNA Calculator Pro (Biochrom, UK), e as sequências foram obtidas no sequenciador automático MegaBACE 1000 (Molecular Dynamics) utilizando-se o kit MegaBACE Dye Terminator. A análise da qualidade do sequenciamento foi realizada pelos programas PHRED e PHRAD (www.unb.br/ib/cel/biomol), e a análise comparativa das sequências obtidas com as sequências depositadas em banco de dados foi executada no recurso BLASTX (<http://www.ncbi.nlm.nih.gov/BLAST/>), de modo a confirmar sua similaridade com genes que estão envolvidos em vias metabólicas de degradação de biomassa vegetal, bem como identificação da família enzimática à qual a enzima codificada pertencia.

5.7. Subclonagem dos genes *TlicBgluc1* e *AgraGH45-1*

Em seguida, novos *primers* com os sítios de restrição para subclonagem dos genes amplificados no vetor de expressão pGAPZα-B foram desenhados (Tabela 6). A inserção dos sítios de restrição nos *primers* ocorreu após a análise das enzimas de restrição no servidor NEBcutter V2.0: (<http://nc2.neb.com/NEBcutter2/>) para não interromper a inserção da construção no sistema de subclonagem. Novas PCRs foram

realizadas utilizando os referidos *primers*. Os amplicons observados foram recuperados, purificados e quantificados, conforme descrito no tópico 5.7, e ligados ao vetor pGAPZ α -B (Figura 8), por meio da reação de ligação composta por: 15 ng do inserto (DNA), 1 μ L do vetor pGAPZ α -B, 2 μ L da enzima T4 Ligase (Invitrogen™) e 4 μ L de tampão 5X, completando o volume com água Milli-Q para 20 μ L e concentração final de inserto:vetor de 90 ng/ μ L.

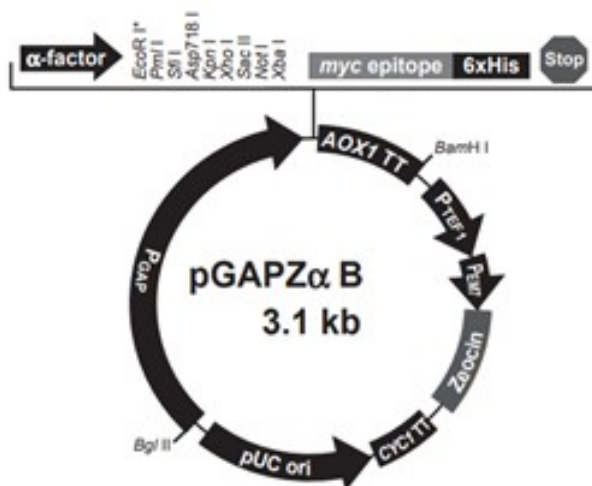


Figura 8. Representação esquemática do vetor de expressão pGAPZ α -B (*P. pastoris* Expression Vectors). Fonte: Technical Manual.

5.8. Transformação de *P. pastoris*TM (*K. phaffii*)

Para conduzir a expressão heteróloga em *P. pastoris*, o vetor contendo os genes de interesse e linearizado com a enzima de restrição *Bgl* II para *T. licus licus* e *Avr*II para *A. grandis*, utilizados para transformar células eletrocompetentes de *P. pastoris*, preparadas conforme o procedimento padrão, descrito no manual Expression Kit (ThermoFisher Scientific). Para isso, adicionaram-se 2,5 μ L (10 U) da enzima de linearização, 15 ng/ μ L⁻¹ do inserto, 1 μ L tampão (10X), e completou-se o volume final com água Milli-Q para 10 μ L. Após incubação a -20 °C por 18 horas, o sistema foi submetido à centrifugação a 10.000 rpm, 30 minutos e 4°C. O sobrenadante foi descartado e adicionaram-se 300 μ L de etanol a 80%, novamente centrifugando por 10.000 rpm, 10 minutos e 4°C. O sobrenadante foi descartado e esse passo foi

repetidoduas vezes. O precipitado foi liofilizado a 37°C, por 1 hora, e ressuspensão em 10µL de água Milli-Q. A construção foi quantificada conforme descrito no item 5.4.

Células de *P. pastoris* foram cultivadas em meio YPD sólido estéril (extrato de levedura a 1%, peptona de caseína a 2%, D-glicose a 2% e ágar bacteriológico a 2%), a 28 °C, para o isolamento de colônias (quatro dias). Posteriormente, preparou-se um pré-inóculo com 10 mL de meio YPD líquido estéril e 1 colônia de *P.pastoris*, incubando-se a 28 °C, por 18 horas. Foram adicionados 150 mL do pré-inóculo em 500 mL de meio YPD. Quando o inóculo atingiu OD₆₀₀ de 1,4, este foi centrifugado a 4.000 rpm, por 5 min a 4°C. O sedimento foi ressuspensão em 200 mL de água destilada gelada. Foi realizada outra centrifugação nas mesmas condições e o sedimento ressuspensão em 100 mL de água estéril gelada, repetindo-se o procedimento. Após mais uma centrifugação, nas mesmas condições, as células foram ressuspensas em 10 mL de 1M sorbitol gelado e estéril. Foi realizada nova centrifugação e as células foram ressuspensas em 500 µL de 1M sorbitol. Aliquotas de 80 µL foram preparadas e mantidas em gelo até a utilização.

Para o procedimento de transformação, foram misturados 80µL de células eletrocompetentes com aproximadamente 10 µg dos vetores linearizados. A mistura foi colocada em cubetas de 0,2 cm e incubada por 5 minutos em gelo. A eletroporação foi realizada com os seguintes parâmetros: 1500 V, 25 µF, 400 Ω em eletroporador (Bio Rad gene Pulser). Em seguida, foram adicionados 680 µL de sorbitol estéril e gelado a 1M na cubeta e o conteúdo foi transferido para tubo estéril de 1,5 mL, incubando-se a 30 °C sem agitação por 1 hora, e plaqueando-se 150 µL do sistema de transformação em meio YPD sólido acrescido de zeocina a 50 µg/mL. As placas foram incubadas por 4 dias a 28 °C até o aparecimento das colônias.

As colônias emergentes e cuja integração do plasmídeo foi confirmada por digestão enzimática foram selecionadas para expressão heteróloga em pequena escala. A digestão ocorreu com 1 µL do DNA plasmidial, acrescido de 0,5 µL de cada enzima de restrição e 4 µL de tampão 10X (Thermo Scientific), completando o volume final para 40 µL com água Milli-Q, incubando-se a 37 °C por 3 horas, e os fragmentos gerados separados por eletroforese, visualizados por meio do foto-documentador de géis Bio-Rad.

5.9. Expressão de proteínas no sistema *P. pastoris*TM

Em 5 mL de meio YPD contendo zeocina a 50 µg/mL foram inoculadas as colônias recombinantes, mantidas por 4 dias a 28 °C e 200 rpm para expressão. Após os tubos foram centrifugados a 4 °C por 5 minutos a 200 rpm, retirando-se 500 µl do sobrenadante para conduzir os ensaios de atividade enzimática e para análise do perfil proteico por eletroforese em gel desnaturante de poliacrilamida (SDS-PAGE) a 12%. A quantificação do sobrenadante foi estimada pelo método de Bradford (1976), utilizando-se albumina de soro bovino (BSA- Sigma) como padrão (ANEXO 1). Dessa forma o restante do sobrenadante foi colocado em novos tubos e as células foram descartadas. Foi adicionado 2X o volume do sobrenadante de ácido tricloroacético (TCA) e acetona (1:3), estocando-se o sistema a -20 °C por 18 horas para precipitação das proteínas como descrito por Niu et al, 2018. Logo após, os tubos foram novamente centrifugados a 9.000 rpm por 30 minutos a 4 °C. O sobrenadante foi descartado e ao sedimento acrescentou 1 mL de acetona a 100 %, e o conteúdo foi transferido para novos tubos de microcentrífuga, sendo centrifugados a 11.000 rpm por 20 minutos a 4 °C. A centrifugação foi repetida mais duas vezes para retirar e evaporar toda acetona.

5.10. SDS-PAGE, *Western Blotting* e ELISA

As proteínas expressas e quantificadas foram solubilizadas em 25 µl tampão de amostra de proteínas (Tris-HCl 1 a M pH 6,8, SDS a 4 %, β-Mercaptoetanol a 4 %, Glicerol a 20 %, Azul de bromofenol a 0,1 %) para serem corridas 30 µg no gel de SDS-PAGE a 12%. As amostras foram desnaturadas a 100°C por 10 minutos e centrifugadas a 5.000 rpm por 5 minutos. As proteínas foram separadas no gel juntamente com o marcador de peso molecular BenchMark™ Prestained (ThermoFisher Scientific) para 10 a 220 KDa. A separação foi conduzida a 220 V, 40 mA e 10 W, com a miliamperagem constante, e em tampão de corrida 1X (Trizma base a 125 mM, glicina a 0,96 M, SDS a 0,5 %) por 4 horas. Para a revelação de bandas, foi utilizada a solução corante de Coomassie brilliant blue G-250 (Bio Rad) a 0,25% por 12 horas e a solução decolorante de metanol a 30% com ácido acético a 7 %, até o surgimento de bandas visíveis.

Também foi realizada a transferência das proteínas de um gel-espelho para uma membrana de nitrocelulose para conduzir o *Western blotting*. Para isso, a membrana foi ativada com a solução de transferência 1X (Tris base a 125 mM, glicina a 960 mM, pH

8,3) na proporção de 5 mL de solução de transferência 10X, 35 mL de água Milli-Q e 10 mL de metanol por 5 minutos em placa de Petri. A transferência ocorreu a 5V por 50 minutos. Após a transferência, utilizou-se 20 mL de solução de bloqueio a 3%, contendo tampão TBS-T 1x (Tris base a 0,5 M, NaCl a 1,5 M, Tween 20 a 0,1%), 0,5% de PVA (poli-vinil álcool - Sigma Aldrich, DE) e 3% de gelatina, previamente aquecida em micro-ondas por 1 minuto e filtrada com papel filtro comum. A membrana permaneceu incubada por 12 horas a 48 rpm. Descartou-se a solução de bloqueio, a membrana foi lavada com 15 mL de TBST 1X por 5 minutos a 48 rpm por 2 vezes, sendo a terceira lavagem com o tampão TBS 1x (Tris base a 0,5 M, NaCl a 1,5 M). Descartou-se o tampão e adicionaram-se 15mL de TBS 1x contendo 1% de gelatina, previamente aquecida em micro-ondas e filtrada com papel filtro comum. Em seguida adicionaram-se 3 µl do anticorpo anti-His conjugado com fosfatase alcalina (Thermo Fisher Scientific) (1:5.000) a suspensão, incubando novamente a membrana por 4 horas.

Depois disso, a membrana foi lavada, como já descrito, para retirar todo o anticorpo não-ligado ao antígeno-alvo. Para a revelação da membrana, utilizou-se a solução de revelação AP (Bio Rad), composta por 50µL da solução A, 50µL da solução B e 200 µL do tampão AP Color 25x, completando o volume com água Milli-Q para 5 mL.

5.11. Ensaio de difusão radial

Alíquotas de 500 µL do sobrenadante da expressão das leveduras, transformadas com os vetores pGAPZ α -B vazio, pGAPZ α -B/TlicBgluc1 e pGAPZ α -B/ AgraGH45-1, foram utilizadas para conduzir os ensaios de atividade enzimática. As placas foram preparadas em meio YPG sólido, enriquecido com 1% de carboximetilcelulose (CMC) e vermelho congo a 1%, em diferentes faixas de pH 4,0 a 6,0 (ajustado com Acetado de Sódio 50 mM), assim como as placas com Celobiose (Sigma) para os ensaios com a β -glicosidase recombinante de *T. licus licus*, e CMC para a β -1,4-endoglucanase de *A. grandis*. No meio solidificado, foram realizadas perfurações (5 mm de diâmetro/10 mm de profundidade) para colocar 50 µL da expressão controle (vetor vazio) e das expressões recombinantes separadamente, incubando-se a 45 °C, por 16 horas. Para a revelação, foi adicionando ácido acético a 10% na superfície, mantendo-se em contato com o material inoculado, por 1 hora, a temperatura ambiente. A hidrólise do CMC foi analisada pela formação de halos brancos ao redor do poço. O conteúdo de proteínas do

sobrenadante foi determinado pelo método de Bradford (1976). A atividade das enzimas recombinantes foi detectada, conforme descrito por Jimenez e colaboradores (CALDERÓN-CORTÉS et al., 2010). Como controles negativo foram utilizados 50 µL de sobrenadante de cultura celular inoculado com *P. pastoris*TM, transformada com o vetor vazio pGAPZα-B.

5.12. Quantidade de açúcares redutores liberados pela atividade ceulolítica

Células de *P.pastoris*TM (colônias) transformadas com o vetor pGAPZα-B vazio e pGAPZα-B/TlicBgluc1 foram inoculadas separadamente em 5ml por 4 dias a 28 °C em 200 rpm em meio YPG suplementado com 1% de CMC. A atividade enzimática foi determinada por ensaio com ácido 3,5-dinitrossalicílico (DNS). O sobrenadante foi isolado por centrifugação a 4.000 g, à temperatura ambiente, durante 10 minutos. Uma alíquota de 100 µL do sobrenadante foi adicionada a 100 µl de DNS a 1% e incubada a 100 ° C por 10 minutos. A quantidade de açúcares redutores liberados pela atividade da glicosidase foi determinada usando glicose como padrão, de acordo com a metodologia descrita por Miller (1959). A absorbância das reações foi medida a 550 nm e os valores foram utilizados para o cálculo do teor de açúcares redutores.

O DNS é um composto aromático que reage com açúcares redutores e outras moléculas redutoras para formar o ácido 3-amino-5-nitrossalicílico, que absorve fortemente a luz a 550 nm. Uma unidade de atividade da enzima (U) foi definida como a quantidade de enzima que libera 1 µmol de produto formado por minuto nas condições do ensaio. Os tubos controle continham as mesmas quantidades de reagentes do tubo de reação, porém um com o vetor vazio e outro somente com o meio de cultura não inoculado (branco).

6. RESULTADOS E DISCUSSÃO

Não há dúvidas de que a celulose, presente nos resíduos lignocelulósicos, é o ponto de partida da indústria para o futuro. A expectativa das indústrias é diminuir as etapas iniciais de produção, e tornar o processo mais rentável e menos oneroso. A utilização de enzimas celulolíticas, com melhor desempenho, pode, portanto, reduzir a suplementação das celulasas para a conversão total de celulose em glicose durante a fermentação (JORDAN et al,2012). Para obter uma enzima melhorada, pesquisadores já adotam várias abordagens, como a expressão heteróloga de celulasas, de diversas fontes, com a intenção de produzir essa enzima em grande escala (SINGH et al., 2016). Desta forma, o estudo aqui desenvolvido inclui a clonagem e expressão, em *P. pastoris*TM, de um gene para uma β -glicosidase do inseto *T. licus licus* e de um gene para uma β -1,4-endoglucanase de *A. grandis*, visando avaliar o potencial dessas enzimas em futuras aplicações biotecnológicas.

6.1. Análise dos Transcriptomas de *T. licus licus* e *A. grandis*

Os transcriptomas de *T. licus licus* e *A. grandis* forneceu dados sobre os genes relacionados à degradação da celulose em insetos praga. Neste estudo, uma ORF de 1473 pb foi identificada na biblioteca de cDNA de *T. licus licus*, codificando uma β -glicosidase de 490 resíduos de aminoácidos (denominada Tlic-B-glu1). Outra ORF de 654 pb, foi identificada na biblioteca de *A. grandis*, codificando uma β -1,4-endoglucanase de 217 resíduos de aminoácidos (denominada AgraGH45-1). Ambos os genes foram clonados e as enzimas codificadas foram expressas com sucesso em *P. pastoris*TM.

A análise do transcriptoma é uma ferramenta necessária para entender melhor a biologia desses insetos-praga. Não há relatos de um sistema de criação, com todo o ciclo, em laboratório de *T. licus licus*, o que dificulta um estudo completo de seu desenvolvimento. A falta de dados sobre a fisiologia e biologia molecular de *T. licus licus* motivou, portanto, o sequenciamento de seu transcriptoma. Relatos recentes confirmam a presença de genes relacionados à degradação da celulose em insetos Coleopteros (IBARRA et al, 2019), assim como na ordem Lepdoptera, já publicado por Pauchet e seus colaboradores em 2009.

Durante a elucidação do transcriptoma, foi realizada uma análise mais aprofundada, comparando-se as sequências do intestino médio para identificar transcritos de genes envolvidos na digestão de celulose. Nesse sentido, os transcriptomas do intestino médio tornaram-se o principal sistema de estudos dessas espécies para a prospecção de enzimas com fins biotecnológicos (FIRMINO et al, 2013; FONSECA et al, 2015). Atualmente, os dados disponíveis para estudos residem no transcriptoma de vários estágios de desenvolvimento dos insetos (ovos, larvas em estágio inicial, larvas em estágio final, pré-pupas, pupas e adultos machos e fêmeas), carcaça e intestino médio (FONSECA et al, 2015). Como já descrito anteriormente neste estudo, a espécie *T. licus licus* apresenta hábito endofítico, alimentando-se do colmo da cana-de-açúcar. Essa característica pode ter influenciado na expressão diferencial das enzimas digestivas, tornando-as mais ativas, em comparação ao nível de degradação, observada em outra broca-da-cana, o inseto-praga, *Diatraea saccharalis* (Lepidoptera).

É importante salientar que, durante o sequenciamento, as sequências de contaminantes (procarióticas, virais e mitocondriais) foram removidas após análise. Assim, parece não haver influência significativa de contaminação por DNA procarioto na biblioteca de cDNA de *T. licus licus* e *A. grandis* (FIRMINO et al, 2013; FONSECA et al, 2015).

Para iniciar a confirmação e a identificação dos genes de interesse, o alinhamento múltiplo de resíduos de aminoácidos foi uma ferramenta norteadora. A sequência de aminoácidos TlicBgluc1 foi identificada na família Glicosil Hidrolase 1 (GH1) (Figura 9A). Observou-se que os motivos catalíticos NEPW (resíduos 184-187) e TENG (resíduos 391-384) foram conservados em comparação a outras β -glucosidases GH1. É provável que os dois ácidos glutâmicos conservados E187 e E399 em TlicBgluc1 atuem como sítio catalítico e nucleófilo, respectivamente, como descrito em outras β -glucosidases GH1 (LIU et al, 2011).

A sequência de aminoácidos prevista de *A. grandis*, AgraGH45-1, compartilhou uma identidade com outras glicosil-hidrolases da família 45 (GH45), como a Rf-GH45 de *Rhynchophorus ferrugineus* (74%), e Oa EGase II de *Oncideres albomarginata* (56%). Todos os aminoácidos são conservados nos domínios catalíticos de β -1,4-endoglucanase. No entanto, as três sequências mostram pequenas diferenças entre os aminoácidos nos locais de glicosilação. Os resultados do BLASTx também demonstraram uma fraca similaridadecom a glicosil hidrolase de *Aspergillus niger*

(GH61), mostrando apenas 22% de identidade com AgraGH45-1 (Figura 9B).

Esses dados corroboram com os dados já conhecidos em coleópteros e lepidópteros, que possuem número variável de glicosil hidrolases em seus genomas (GH1, GH5, GH9, GH45 e GH48) (DANCHIN et al, 2012; SRIVASTAVA et al, 2018; KIRSCH et al, 2019).

1 A	Bglucl	<u>MKSALVLSLMVLACSAT</u> --TIKQQRFPDDFLFGTATASYQIEGGWNEDGKGENIWDYMT	58
	AAC06038.1_Sfru	<u>MKLLVLSLVAACNAS</u> --IVRQQRFPDDFLFGTATASYQIEGAWDEDGKGENIWDYMV	58
	XP_011557578.1_Pxyl	<u>MKRFLVILCAVLLCGHAA</u> AVVKQSGRSFPKDFLFGTATASYQIEGAYNEDGKGENIWDRLT	60
		** :*: .: *.: : .: * ** .*****:*****:***** :.	
	Tlic-Bglucl	HNTPEVIKDFSNGLDIAADTYNNYKRDVEMMRELGLDAYRFSMSWARILPNGFSNEINEAG	118
	AAC06038.1_Sfru	HNTPEVIRDLNSGLDIAADSYHNYKRDVEMMRELGLDAYRFSLSWARILPTGMANEVNPAG	118
	XP_011557578.1_Pxyl	HQVPSPKDKQSSGLDIAADSYHKVERDVEEMRELGLDAYRFSLSWSRILPTGFSNHINPAG	120
		:.. *: * .*****:***: *****:*****:***:***.***:*** **	
	Tlic-Bglucl	VQFYNNYIDEMIKYNIMPMVTLYHWDLPQKLQELGGFQNPLISDFWFEDYARVAYEKFGR	178
	AAC06038.1_Sfru	IAFYNNYIDEMIKYNITPLITLYHWDLPQKLQELGGFANPLISDFWFEDYARVVFNFGDR	178
	XP_011557578.1_Pxyl	VDYNNRLINEMIKYNIQPLVTLTYHWDLPQPLQDLGGFASPLFGQWFEDYARVVYTNFGDR	180
		: :*. *.:*.:*** *.:***** **:* ** .***:*****: :****	
		▼	
	Tlic-Bglucl	VKHWITFNEPREICYEGYGSSTKAPILNATGIGEYICAKNLVMSHAKAYRAYNKDFKSSQ	238
	AAC06038.1_Sfru	VKMFITFNEPREICFEGYGSATKAPILNATAMGAYLCAKNLVTAHAKAYLYDREFRPVQ	238
	XP_011557578.1_Pxyl	VKFFITFNEPREICNEGYGGTGFAPVFNATGIGEYICAKHLVTAHAKAYHLYNNEFRASQ	240
		** :***** ***: .: **:* ** .***:*** :***** *.:* : *	
	Tlic-Bglucl	GGQCGITISVNWFGPETDSEEDQAAELFRQCQWGMYAEPISFVEGGFPKELAEVAQKS	298
	AAC06038.1_Sfru	GGQCGITISVNWFGPATPTPEDEMAAELRQGEWGIYAHPIFSAEGGFPKELSDKIAEKS	298
	XP_011557578.1_Pxyl	GGKVGITISVNWFEAETDSEDDKLAIIWFRQGSWGIWFSEPIFGLGGFPKELAEVAEKS	300
		** : ***** * : :* ** * .***:***. *****:***:***	
	Tlic-Bglucl	AEQGYRSRMPFEFTYEERILVRGAYDFFGVNHYYTYLVSASNYLTTNPIPSMLDDVNVGT	358
	AAC06038.1_Sfru	AQQGYPWSRLPEFTEEEKAFVRGTSDLIGVNHYYTFLVSATERKGPYPVPSLLDDVDVTS	358
	XP_011557578.1_Pxyl	AQQGFPRSRLEPFTDEEREYIRGTSDFFGINHYFNARVSATYTPVNPVPSLNDDVGVDL	360
		*:**: **:* ** : **:* **:* ** ***: :**:* **...	
		▼	
	Tlic-Bglucl	VVPEDWPQSASTWLYQAPNSIYNALTHLKSXYNDPIFYITENGWSTFT--PGGLEDDRI	416
	AAC06038.1_Sfru	WADDSWLKSASAWLTLAPNSIHTALHLLNLYNKPVFYITENGWSTDESRENSLIDDDRI	418
	XP_011557578.1_Pxyl	IFESKRAESTSGPV---VKLSNVMLLLKELYGEWDFYITENGMSSA---PGVDDTARG	412
		. . :*: * : .: .: * :. *.. ***** * : .: * *	
	Tlic-Bglucl	RYYRAALESALDSLDAVRLKGYMAWSLMDNFEWMEGYTERFGLYQADFEDPARTRTPRK	476
	AAC06038.1_Sfru	QYYRASMESLLNCLDDGINLKGMAWSLMDNFEWMEGYIERFGLYEVDVDFSDPARTRTPRK	478
	XP_011557578.1_Pxyl	LYYKAALENCLDAIDGIIKLGMAWSLMDNFEWMOGYTSRFGLYAVDFEDPNRPTPKH	472
		:*:*. *.:* * :.***:*** ** .***** .**.* * ***: :	
	Tlic-Bglucl	SAFVYKQIIKTRV----- 489	
	AAC06038.1_Sfru	AAFVYKHIIKRVVDYEPET--MVMTIDEGH 509	
	XP_011557578.1_Pxyl	SAFVYKRIIKDRFIDHAYVPDASVSVMTIDEGH 505	
		:*****:*** *	

B

<i>Oa-Ease II</i>	MKVIIILVLAVLYTAQAEKHYHSIPVKYGVSGSGVT <u>TRYWDC</u> CCPSCGWSENLNNPSGTFV	60
<i>Rf-GH45</i>	MKAFIPIILALATIAY--ASEIQTVVNGISGQGT <u>TRYWDC</u> CKPSCAWIENVNTPV-N	56
<i>AgraGH45-1</i>	MKSFVGAIAVL-----SVVAVAYGLSGSGT <u>TRYWDC</u> CKPSCSWKENVGTLDPV-D	50
	** : : : * * * * * : * * * * * : * * * * * : * * * * * : * * * * *	
<i>Oa-Ease</i>	▼ TCTADGSTAAPNAESSC-IGGTAYMCSSQPRMVNSSFALGYVAASFSGGADTSMCCAC	119
<i>Rf-GH45</i>	SCAADGVTIVNASVHSGCDNDGTSYVCNNQPWAVSDTLAYGFVAASFSGGADNSHCCIC	116
<i>AgraGH45-1</i>	SCAADGETKLDASVTSVSGCDSDGTSYVCNNLQPWAVNDTLAYGFVAASFSGGVDNSQCCVC	110
	: * : * * * * : : * * * * * : * * * * * : * * * * * : * * * * * : * * * * *	
<i>Oa-Ease</i>	IRLDFQGDLSGKSMIVQVTNTGGDLGNSNHFDIAIPGGGVGIFTEGCSSQWGPCSTGWGDQ	179
<i>Rf-GH45</i>	LKLFQNALNGKSMVVQVINTGSDLGNSNHFDIQVPGGGVGIPTRGCCQTQWNAPSWGQQ	176
<i>AgraGH45-1</i>	LKLTFFKNFLAGKIMIVQNVNTGGDLGNSNQFDIQIPGGGVGIFTRGCCQTQWNAPSWGQQ	170
	: : * * : . * * * * : * * * * * : * * * * * : * * * * * : * * * * * : * * * * *	
<i>Oa-Ease</i>	YGGVSSAECSQLPSALQS [▼] GCAFRFQFLGGVSNPPVNFQVECPGDLT [▼] GATGCTY--	234
<i>Rf-GH45</i>	YGGVSSDAECSQLPSELQEGCHFRFNWFQ [▼] NADNPQVYFEQVTC [▼] PSELTSLSGCSA--	231
<i>AgraGH45-1</i>	YGGVTSDAECDELPELQAGCHFRFGWYENADNPQVDFEQITC [▼] PTELTSLTGCVNDN	227
	*****:*.***.~** : ** ** ** : . . . ** * ** : ** : ** . **	

Figura 9. Alinhamento das sequências gênicas.

A) Alinhamento da seqüência de aminoácidos de β -glicosidases em Lepidopteras. β -glicosidases listadas: TlicBgluc1 (este estudo), *S. frugiperda* (AAC06038.1) e *Plutellaxylostella* (XP_011557578.1). A linha tracejada sublinhada: local de clivagem do peptídeo sinal nativo. As seqüências sublinhadas contêm os resíduos de ácido glutâmico (marcado com uma ponta de seta escura), responsáveis pela atividade glicosil-hidrolase. As letras em negrito correspondem aos resíduos de aminoácidos envolvidos na ligação ao substrato.

B) Alinhamento da seqüência de aminoácidos de β -1,4-endoglucanases de Coleopteros. β -1,4-endoglucanases listadas: AgraGH45-1 (este estudo), Rf-GH45 de *Rhynchophorus ferrugineus* e Oa EGase II de *Oncideres albomarginata*. O peptídeo de sinal está sublinhado no extremo N-terminal. Quadrante: local catalítico conservado, incluindo ácido aspártico catalítico. Os locais potenciais de N-glicosilação estão sublinhados abaixo do alinhamento. As setas escuras indicam resíduos de cisteína conservada.

Essa classificação das glicosil hidrolases (GH) pode ocorrer por dois métodos. Uma baseada na especificidade em relação ao substrato, e outra, baseada na seqüência de nucleotídeos identificados. A classificação baseada na especificidade ao substrato não é muito informativa, uma vez que existem múltiplos substratos que não podem ser determinados para uma enzima individual (SRIVASTAVA et al, 2018). Neste estudo, classificamos a enzima TlicBgluc1, como GH1 e a enzima AgraGH45-1, como GH45, com base na seqüência nucleotídica identificada. Este sistema de classificação é amplamente aceito por fornecer informações estruturais sobre outros membros da mesma família, por meio de ferramentas de bioinformática (SRIVASTAVA et al, 2018). Na família GH, as endoglucanases são as enzimas responsáveis pela rápida solubilização do polímero celulósico (redução do grau de polimerização), assim como são responsáveis por iniciar a hidrólise da molécula de celulose. Por isso, sua classificação também é baseada na especificidade em relação ao substrato. Desde 2004,

o substrato preferencial para verificar a atividade das endoglucanases é a carboximetilcelulose-CMC (ZHANG et al, 2004). Identificada então, as famílias dos genes candidatos, os seus transcritos, *TlicBgluc1* e *AgraGH45-1*, foram analisados por qPCR, confirmada assim, um elevado nível de expressão no estágio larval final e no intestino médio de *T.licus licus* (Figura 10), e de *A.grandis* (Figura 11), mais expresso no intestino médio do que na carcaça. Segundo Terra e Ferreira (1994), o intestino médio é a parte mais importante da digestão de carboidratos no sistema digestivo de insetos.

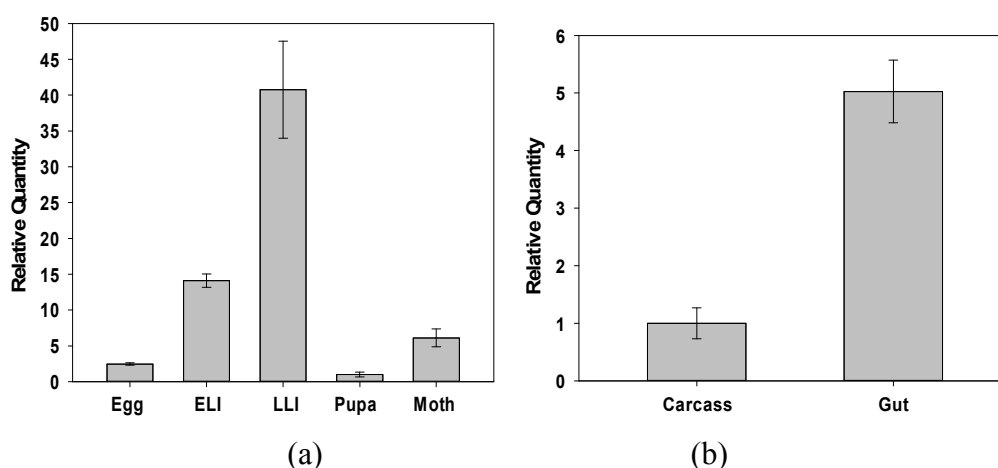


Figura 10. qRT-PCR dos transcritos *TlicBgluc1*. Padrão de expressão do gene *TlicBgluc1* em diferentes estágios da vida. (A) A expressão do novo gene *TlicBgluc1* foi avaliada por qPCR em diferentes estágios da vida dos insetos (ovo, estágio larval inicial [ELI], estágio larval tardio [LLI], pupa, adulto) para determinar o estágio em que o gene exibe maior expressão. De acordo com os resultados, o gene foi amplamente expresso ao longo do ciclo de vida do inseto, mas a expressão aumentou significativamente no LLI, quando o inseto exigiu a maior utilização de glicose. (B) No estágio larval, a expressão do gene *TlicBgluc* foi avaliada no intestino médio e na carcaça. Os transcritos do gene foram mais detectados no intestino. *TlicGAPDH* e *TlicRPS18* foram escolhidos como genes normalizadores endógenos e a expressão relativa foi calculada usando o método $2^{-\Delta Ct}$. Barras representam 5% de erro padrão.

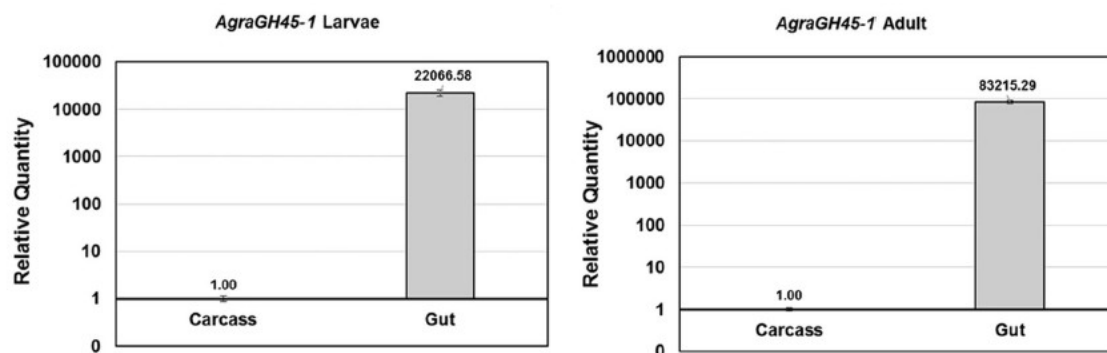


Figura 11: qRT-PCR dos transcritos *AgraGH45-1*. Perfil de transcrição do *AgraGH45-1* em tecidos de adultos e larvas de *A. grandis*, incluindo o intestino médio e a carcaça. A carcaça compreende todos os outros tecidos de insetos, mais o intestino. Em ambos os estágios, o gene *AgraGH45-1* é muito mais expresso no intestino médio do que na carcaça.

6.2. Clonagem e subclonagem

Após o desenho dos pares de *primers* para a clonagem, estes foram sintetizados (Tabela 3). Foi realizada a extração do RNA total intestinal de *T. licus licus*, assim como de *A. grandis*, bem como a amplificação do cDNA via RT-PCR.

Tabela 3. Oligonucleotídeos para clonagem das ORFs dos genes candidatos.

Primer	Primer Sequência 5' – 3'	Ta (° C)	Fragmento esperado
<i>TlicBgluc1Fw</i>	5' _ATGGTGTGGCCTGTAGTGCC 3'	58°	1.473 pb
<i>TlicBgluc1Rv</i>	5' _TTAGTCGACTGTAAGTCTGGTCTTTATTATCTG 3'	58°	1.473 pb
<i>TlicBgluc1 qPCR-Fw</i>	CCACAAAAACTTCAAGAGTTAGGC	-	-
<i>TlicBgluc1 qPCR-Rv</i>	ACTCTGTCACCGAACTTTTCGT	-	-
<i>AgraGH45-1Fw</i>	5' ATGTGACGAGCTCCCTACAGA 3'	60°	654 pb
<i>AgraGH45-1Rv</i>	5' TTCAGTTGGGCAGGTAATTTG 3'	60°	654 pb
<i>AgraGAPDH_qPCR_F</i>	AGATCGTCGAGGGTCTGATG	-	-
<i>AgraGAPDH_qPCR_R</i>	AAGGCGGAATGACTTTACC	-	-
<i>AgraBtub_qPCR_F</i>	GGTTGCGACTGTTTACAAGG	-	-
<i>AgraBtub_qPCR_R</i>	GCACCACCGAGTAAGTGTTCC	-	-

As sequências codantes foram analisadas *in silico* no servidor NEBCutter V2.0 (BioLabs) para mapear sítios de restrição existentes e inexistentes dos genes de interesse. Após esta análise foram desenhados novos *primers* para a subclonagem no vetor de expressão, onde se adicionou nas extremidades 5' e 3' as endoglucanases de

restrição (Tabela 6).

Tabela 6. Sítios de restrição adicionais aos *primers* para subclonagem.

Gene	Primers com os sítios de restrição sublinhados	Enzimas de restrição	Orientação
<i>TlicBglu1</i>	5' <u>TA</u> ACTGCAGATACCATAAAACAGCAGAGAA 3'	<i>Pst</i> I	Forward
<i>TlicBglu1</i>	5'TTAGTCGACTGTA <u>ACT</u> CTGGTCTTTATTATCTG3'	<i>Sal</i> I	Reverse
<i>AgraGH45-1</i>	5'TAAGAATTC <u>CA</u> CTTAGCGGTCTGGA <u>ACT</u> ACT 3'	<i>Eco</i> RI	Forward
<i>AgraGH45-1</i>	5' TTAGTCGACCATATTGTCATTTACACAACC 3'	<i>Sal</i> I	Reverse

Os vetores de clonagem *pGEM®-T Easy* foram amplificados e os plasmídeos extraídos dos clones transformantes foram digeridos com as enzimas de restrição específicas para cada inserto. As análises confirmaram a presença dos insertos, assim como a integridade do vetor de clonagem (Figura 12).

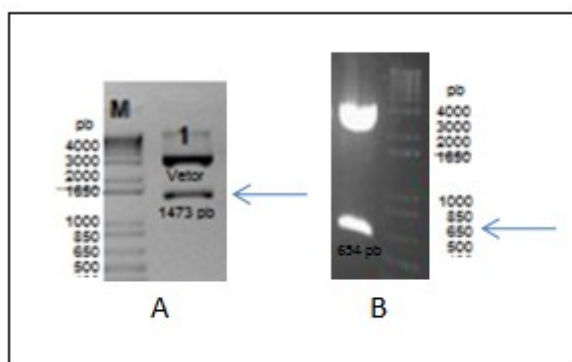


Figura 12. Análise de restrição do vetor *pGEM®-T Easy* com os genes candidatos. (A) Digestão do vetor *pGEM®-T Easy* (3000 pb) com endonucleases *Pst*I e *Sal*I do gene *TlicBglu1* (1473pb-seta). (B) Digestão do vetor *pGEM®-T Easy* (3000 pb) com endonucleases *Eco*RI e *Sal*I do gene *AgraGH45-1* (654pb-seta). M) Marcador Ladder Plus 1 kb (Gibco).

Com os novos *primers* para a subclonagem no vetor de expressão, foram realizadas outras PCRs nas mesmas condições descritas nas Tabelas 4 e 5. As bandas amplificadas com estes *primers* foram excisadas e utilizadas para subclonagem no vetor de expressão pGAPZ α -B expresso com uma tag His6 C-terminal em *P. pastoris*TM. Os vetores contendo os genes de interesse foram linearizados e utilizados para transformação de *P.pastoris*TM por eletroporação com as construções *TlicBglu1*-pGAP α -B-His4-Zeo, *AgraGH45-1*-pGAP α -B-His4-Zeo e pGAP α -B-His4-Zeo (controle negativo). Após a ligação gênica e transformação, o DNA plasmidial das construções foram digeridas com endonucleases específicas para o gene *TlicBglu1* (Figura 13) e *AgraGH45-1* (Figura 14) confirmando os fragmentos esperados.

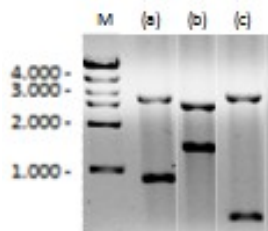


Figura 13. Presença do gene *TlicBgluc1* no genoma da levedura transformada confirmada com as digestões: **(a)** *PstI* e *EcoR I* (900 pb); **(b)** *PstI* e *SalI* (1473 pb): gene *TlicBgluc1*; **(c)** *EcoRI* e *SalI* (500 pb); Eletroforese em gel de agarose 1% corado com brometo de etídeo. M) Marcador 1 kb Ladder Plus (Gibco).

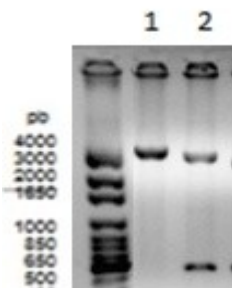


Figura 14. Presença do gene *AgraGH45-1* no genoma da levedura transformada confirmada com as digestões. 1) DNA plasmidial sem digestão. 2) Digestão com *EcoRI* e *SalI* (650 pb): gene *AgraGH45-1*. Eletroforese em gel de agarose 1% corado com brometo de etídeo. M) Marcador 1 kb Ladder Plus (Gibco).

6.3. Expressão recombinante/heteróloga

A preferência, nas indústrias, é por sistemas de expressão protéica utilizando modelos eucariotos (WYSOCKA-KAPCINSKA, et al 2010). Essas proteínas celulósicas industriais apresentam de forma geral, melhores atividades em temperaturas compreendidas entre 45 e 65 °C e em valores de pH mais ácidos, na faixa de 4,0 a 5,5 (MARTINS et al, 2018). Para verificar esses parâmetros, as enzimas β -glicosidase *TlicBgluc1* e β -1,4-endoglucanase *AgraGH45-1* deste estudo, foram expressas como proteínas solúveis, analisada por SDS-PAGE e *Western blotting*. A expressão da β -glicosidase *TlicBgluc1* foi confirmada por ensaio de imunotransferência identificada com peso molecular de aproximadamente 50 kDa, consistente com o estimado a partir da sequência de aminoácidos (Figura 15). A expressão da β -1,4-endoglucanase *AgraGH45-1* também foi confirmada no ensaio, identificada com peso molecular de aproximadamente 35 kDa (Figura 16).

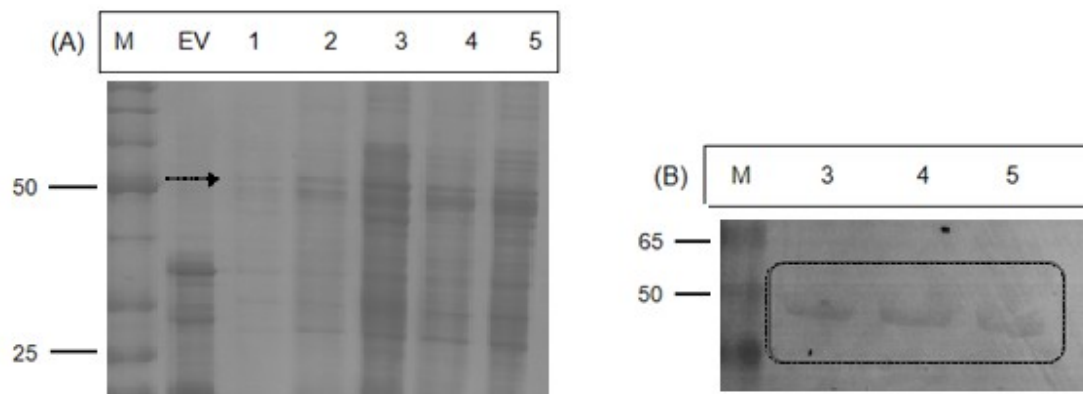


Figura 15. Análise da expressão da proteína recombinante por SDS-PAGE e *Western blotting* de TlicBgluc1.

(A) Análise por SDS-PAGE de colônias com vetor vazio e colônias de *P. pastoris*TM transfectadas com *TlicBgluc1*-pGAPZ α -B. O fragmento transfectado com vetor vazio (EV) não exibe banda correspondente de 50 kDa de acordo com o mw previsto. As colônias 1 e 2 exibem baixa expressão proteica e as colônias 3 a 5 exibem alta expressão proteica prevista por MW. (B) A análise por *Western blotting* confirma a presença da preteína TlicBgluc1 no sobrenadante da expressão da colônia 3 a 5 com uma única banda imunomarcada com o correspondente aproximado.

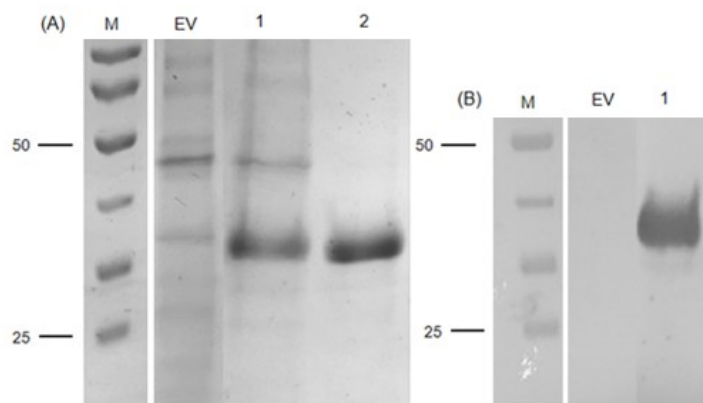


Figura 16: Análise da expressão da proteína recombinante por SDS-PAGE e *Western blotting* de AgraGH45-1.

(A) Análise por SDS-PAGE de colônias com vetor vazio (EV) e colônias de *P. pastoris* transfectadas com *AgraGH45-1*-pGAPZ α -B (1 e 2). O fragmento transfectado com vetor vazio (EV) não exibe banda correspondente de 35 kDa de acordo com o mw previsto. As colônias 1 e 2 exibem alta expressão proteica prevista por MW.

(B) A análise por *Western blotting* confirma a presença da proteína AgraGH45-1(1) no sobrenadante da expressão com uma única banda imunomarcada com o correspondente aproximado.

A maioria das β -glicosidases e β -1,4-endoglucanase usadas em aplicações industriais são derivadas de espécies fúngicas (GUNATA et al, 1999; KROGH et al, 2010; LI et al, 2010). No entanto, essas isoformas não-fúngicas expressas neste estudo, podem fornecer alvos enzimáticos para aplicações industriais. Até o momento, nenhuma pesquisa relatou o inseto-praga *T. licus licus* como uma fonte importante para a obtenção de novas celulases. Esse potencial destrutivo da praga é um diferencial, e a adaptação de suas enzimas digestivas ao degradar a biomassa da cana-de-açúcar. Este inseto passa a maior parte do seu ciclo de vida alimentando-se dentro da planta. Essa característica pode ter influenciado a expressão diferencial de suas enzimas digestivas, tornando-as mais eficientes para a degradação de dissacarídeos que propiciou sua evolução. Portanto, nós acreditamos que *T. licus licuse A. grandis* poderiam ser um recurso valioso para a obtenção de celulases úteis industrialmente. Outra β -glicosidase não-fúngica foi descrita. Liu e seus colaboradores (2011) expressaram uma β -glicosidase com domínio catalítico na família GH1, derivada de uma biblioteca metagenômica da microbiota marinha. Assim como formas fúngicas descrita por Bonfá (2016), que caracterizou uma β -glicosidase de 50 kDa (BG50) produzida pelo fungo termofílico *Myceliophthora thermophila* cultivado em meio sólido contendo uma mistura de bagaço de cana e farelo de trigo, e cuja enzima mostrou atividade ótima a 60 °C e pH 5,0; e por Chen et al. (2012), que relataram uma β -glicosidase expressa em *P. pastoris*TM do fungo *Neocallimastix patriciarum*, que apresentou domínios catalíticos na família GH3. Foi também descrito que besouros codificam suas próprias celulases, devido sua capacidade de degradar a celulose. No inseto *Gastrophysa viridula*, foram identificadas três supostas celulases da família glicosil hidrolase GH45 e GH9. Por meio da expressão heteróloga e ensaios funcionais, as enzimas endo- β -1,4-glucanase, se mostraram ativas na celulose amorfa (BUSCH et al, 2018).

6.4. Caracterização funcional das enzimas TlicBgluc1 e AgraGH45-1

A capacidade das enzimas recombinantes deste estudo, de digerirem celulose amorfa, foi observada no ensaio de difusão radial com a formação de um halo de degradação. A atividade da β -glicosidase TlicBgluc1 foi determinada usando os substratos carboximetilcelulose-CMC e celobiose. A formação do halo de degradação foi observada tanto para CMC como para o meio enriquecido com Celobiose, o que confirmou a atividade celulolítica entre 28 °C e 45 °C, uma maior atividade a 45 °C e

pH 4,5 (Figura 17). A atividade de β -1,4-endoglucanase AgraGH45-1, também foi detectada no ensaio de degradação de CMC, ativa a 50 ° C e pH 5,0 (Figura 18).

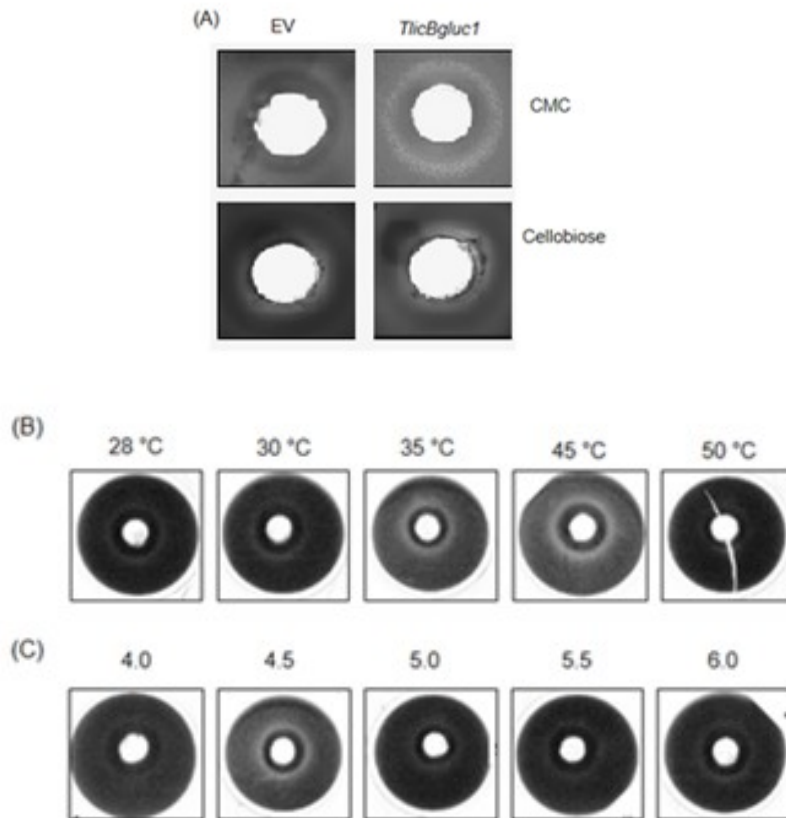


Figura 17: Atividade celulolítica qualitativa da enzima recombinante β -glicosidase TlicBgluc1 em substratos de carboximetilcelulose (CMC) e celobiose.

(A) β -glicosidase TlicBgluc1 exibe afinidade por CMC e celobiose em condições ideais de reação. Colônias transfectadas com EV não exibiram atividade hidrolítica, enquanto colônias positivas foram capazes de hidrolisar substratos, observados através de um halo de degradação.

(B) Efeito da temperatura na atividade β -glicosidase TlicBgluc1. A enzima pode atuar na degradação do substrato entre 28 ° C e 45 ° C, e uma atividade mais alta a 45 ° C. Temperaturas mais altas podem ser consideradas para limitar essa atividade enzimática específica.

(C) efeito do pH na atividade de β -glicosidase TlicBgluc1. A atividade enzimática foi testada em meio de reação variando de pH 4,0 a 6,0. Em todos os pHs testados, a enzima exibe atividade hidrolítica, mas a atividade mais alta foi observada em pH 4,5, sugerindo uma enzima com afinidade ácida.



Figura 18: Atividade celulolítica qualitativa da enzima recombinante β -1,4-endoglucanase AgraGH45-1 em substrato de carboximetilcelulose (CMC).
(I) Cultura de *P.pastoris*TM transformado com vetor vazio (50 μ L) (controle negativo).
(II) Cultura de *P. pastoris*TM transformada com pGAPZ α -B / AgraGH45-1 (50 μ L). Os halos brancos sobre fundo escuro indicam que a atividade em B amostra é superior ao controle.

A CMC é tradicionalmente utilizada para quantificar atividade de endoglucanases, e celobiose para β -glicosidases (ZHANG et al, 2004). Como foi realizado por Byeon e seus colaboradores (2005), na qual demonstraram a atividade enzimática de uma β -glicosidase digestiva de *Bombyx mori* (Lepidoptera), detectada em pH 6,0 a 35 °C, em placas contendo celobiose. Também foi observado na atividade dos genes clonados e expressos heterologicamente em *E.coli* BL21, do fungo filamentoso *Thermofilum adornatum*, cuja atividade enzimática das proteínas recombinantes foi caracterizada com CMC, revelando uma glucanase multifuncional capaz de hidrolisar beta-1,4-glicosídeos (ZAYULINA et al, 2020). Para confirmar a atividade celulolítica da β -glicosidase *TlicBgluc1*, observada no substrato com CMC e celobiose, foi analisado a quantidade de glicose liberada no ensaio com 3,5- dinitrosalicílico método ácido-DNS. Os resultados mostraram uma atividade enzimática duas vezes superiores (4,76 mM) ao nível basal (controle negativo) dos extratos protéicos da expressão em *P. pastoris*TM (Figura 19).

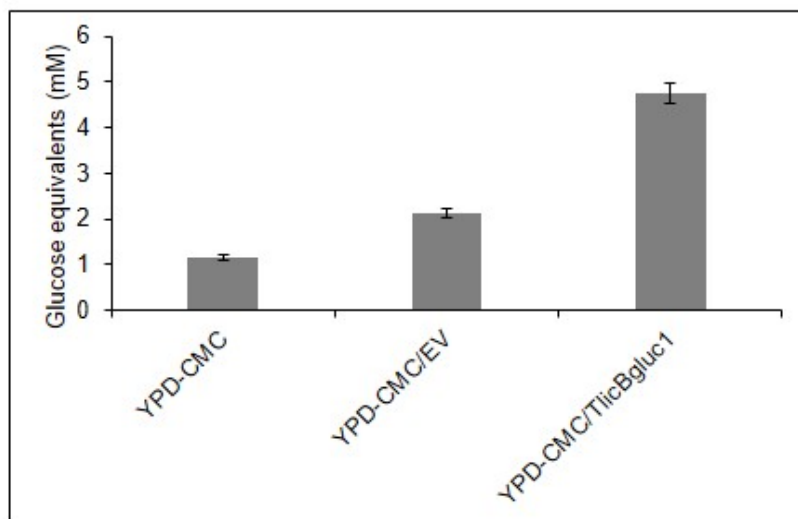


Figura 19. Atividade enzimática β -glicosidase TlicBgluc1 no ensaio de quantificação de açúcar redutor mediado por DNS. O sobrenadante enriquecido com enzima da expressão de *P. pastoris* foi incubado com solução CMC e a liberação de açúcares redutores pelo processo hidrolítico foi quantificada no ensaio de DNS. O sobrenadante foi incubado com o substrato celulósico por até 20 minutos. O meio não inoculado (YPD-CMC) e o meio inoculado transfectado com EV (YPD-CMC/EV) também foram detectados para descartar açúcares redutores residuais presentes no meio de cultura e na atividade basal de outras hidrolases de *P. pastoris*TM. A atividade de TlicBgluc1 (YPD-CMC/ *TlicBgluc1*) pode enriquecer o meio de reação em equivalentes de glicose aproximadamente o dobro do controle (glicose de 2,13 mM a glicose de 4,76 mM).

o

rendimento de etanol em função do substrato, em estudos que analisam a sacarificação enzimática do bagaço de cana-de-açúcar, assim como a bioconversão de açúcares simples para utilização em biorrefinarias (LAMOUNIER et al, 2020). Da mesma forma, para confirmar a atividade celulolítica de β -1,4-endoglucanase AgraGH45-1, foi utilizado outro substrato, hidroetilcelulose-HEC, onde a enzima recombinante mostrou-se ativa (Figura 20).

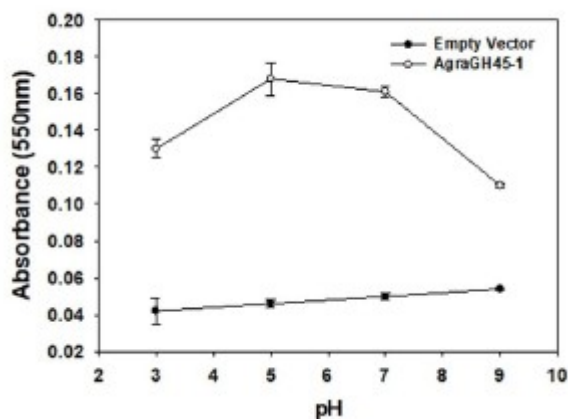


Figura 20: Avaliação da atividade enzimática β -1,4-endoglucanase AgraGH45-1. Atividade com pH variando de 3,0 a 9,0. A reação enzimática foi iniciada pela adição de 50 ml de OBR-HEC (5 mg / ml) como substrato. Cada ponto é a média de três medições. O aumento no Abs 550 nm é proporcional à degradação do substrato.

As β -1,4-endoglucanases derivadas de microorganismos termofílicos, geralmente apresentam atividade em substrato com hidroxietilcelulose (OBR- HEC) em pH 6,5 a 40 ° C (WIERZBICKA-WOŚ et al, 2019). Na ordem Coleoptera, algumas β -glicosidases já foram descritas na literatura. A β -glicosidase do intestino médio de *Leiptinotarsa decemlineata* (Coleoptero), por exemplo, apresentaram atividade na faixa de pH entre 4,0-5,5 a 50 ° C (SZILÁGYI et al, 2019). Yapi et al (2009) demonstraram a atividade ótima da β -glicosidase de *Rhynchophorus palmarum* (Coleoptero) em pH 5,0 a 55°C. No intestino médio de *Rhodnius prolixus* (Hemiptero) o valor ótimo de pH foi 4,5 (TERRA et al, 1988). Na ordem Lepidoptera, uma β -glicosidase digestiva foi purificada de *S. frugiperda*. Este inseto-praga é conhecido pela diversidade de culturas de plantas que ataca, e suas enzimas apresentaram dois sítios ativos que hidrolisam o p-nitrofenil-d-glucosídeo (NP β Glu) e celobiose (MARANA et al, 2000). Apesar da presença de muitas atividades carboidrolíticas nos insetos, seus mecanismos celulolíticos são pouco conhecidos. Enquanto genes da celulase estão ausentes no genoma da *Drosophila melanogaster*, por exemplo, outros insetos como cupins, besouros, produzem suas próprias celulasas (BUSCH et al, 2018). Estudos moleculares apresentam hipóteses sobre os mecanismos pelos quais, os insetos e seus simbioses microbianos, digerem celulose no intestino médio (HAHN et al, 2007; GABRISKO & JANECEK, 2011; WATANABE & TOKUDA, 2014). Há relatos de que as enzimas digestivas apresentem uma plasticidade adaptativa para promover uma absorção nutricional altamente eficiente (MOTOYUKI et al, 1994). Em 2015 uma nova hidrolase

de sacarose foi identificada em *Bombyx mori*, *Trilocha varians* e *Samia cynthia ricini* (WANG et al, 2015).

Os lepidópteros desenvolveram uma capacidade aprimorada de hidrolisar sacarose, sua principal fonte de energia. Já foram identificadas duplicações gênicas e embaralhamento de exons em diferentes regiões microssintênicas nos modelos *Bombyx mori* e *Amyelois transitella* (WANG et al, 2015). O trato digestivo das Lepidopteras é conhecido por sua alcalinidade. O seu sistema digestivo é bem diferente dos outros insetos e muito mais complexo. De fato, suas enzimas digestivas evoluíram para um pH ótimo ideal, no entanto, as modificações genéticas e mecanismos adaptativos dessas enzimas digestivas ainda não foram esclarecidos (LI et al, 2017).

Para demonstrar como a biologia digestiva dos insetos é passível de modificações, uma pesquisa em Lepidopteros teve como objetivo verificar como os nanomateriais podem comprometer a biologia digestiva desses insetos. Trataram folhas da mamona com diferentes concentrações de nanopartículas de prata (AgNPs) e alimentaram dois insetos-pragas, *Ricinus communis* e *Achaea janata*. Usando a microscopia eletrônica de transmissão, células intestinais foram analisadas e observadas um acúmulo de nanopartículas de prata nas organelas celulares. O efeito dos tratamentos provocaram diferenças nas atividades enzimáticas desintoxicantes, nas carboxilesterases (CarE), glicosidases (Glu) e glutathione S-transferases (GST) no intestino larval. Isso sugere que a exposição das larvas às nanopartículas induz o estresse oxidativo (YASUR et al, 2015).

Outros insetos-praga, como os insetos fitófagos, também pode ser uma fonte de prospecção de novas enzimas celulolíticas com aplicações industriais. Foram identificados, pela primeira vez, em *Zigentoma*, altos níveis relativos de atividade celulásica, quando comparados a grupos de insetos já conhecidos, como os cupins. O objetivo desse trabalho foi fornecer uma caracterização morfo-histológica da digestão da celulose e identificar celulasas altamente ativas presentes nos fluidos digestivos, por meio de nano cromatografia líquidos, além da análise do transcriptoma intestinal. Foram identificadas endoglucanases putativas, correspondentes a enzimas de artrópodes; e endo-1,6- β -glicosidases correspondente a enzimas bacterianas (HUANG et al, 2015; POTHULA et al, 2019).

Os padrões ácidos descritos neste estudo corroboram com a preferência industrial, onde as celulasas mais adequadas para este fim devem apresentar atividade estável, em pH 4,0-5,0 e temperaturas de 40 a 50 °C (CALIL et al, 2013), assim como

foi descrito a atividade da celulase comercial Novo 188 (Novozymes), com tolerância à acidez em pH 4 (SINGHANIA et al, 2010; JORDAN et al,2012). Portanto, uma possível transferência de genes da microbiota para o inseto e outras modalidades genéticas evolutivas podem justificar a atividade enzimática ácida observada no trabalho de Byeon (2005) e neste trabalho também. Como a maioria das celulases comerciais adequadas para conversão de etanol 2G demonstram melhor atividade em pH variando de 4 a 5 e em temperaturas próximas a 50 °C (FISCHER et al, 2013), as enzimas recombinantes aqui elucidadas, poderão ser alvo de pesquisas complementares, para, em um futuro próximo, contribuir com o processo de hidrólise enzimática na produção de biocombustíveis.

7. CONCLUSÃO

Este estudo teve como objetivo, estudar genes que estão envolvidos na degradação da biomassa vegetal, presentes nos genomas de dois insetos praga: *T.licus licus* e *A. grandis*. Foi verificado que o produto da expressão heteróloga dos genes clonados degradaram a celulose, apresentado assim, atividade enzimática. Estes dados reforçam a importância de identificar a bioquímica digestiva desses insetos praga com o objetivo de produzir ferramentas biotecnológicas, para as indústrias que utilizam a biomassa vegetal como matéria-prima. Além disso, este estudo apresentou indícios, já descritos na literatura, de que os insetos-praga carregam em seus genomas genes que expressam enzimas da família Glicosil Hidrolases, tornando-os excelentes predadores de materiais lignocelulósicos. Outro dado relevante é que as proteínas expressas apresentaram atividade ótima na faixa de pH 4 a 6, e temperaturas de 45 a 50 °C, perfil enzimático preferencial das indústrias.

Estudos futuros poderão elucidar a atividade ácida das proteínas expressas de *T.licus licus*, uma vez que o pH dos lepdópteros é alcalino. Bem como a identificação de mutações gênicas que tornaram este inseto o maior predador da cana-de-açúcar. Outro dado importante é que os genes clonados apresentaram sua expressão mais significativa nos estágios larvais e no intestino médio, local onde ocorre a maior degradação e absorção de carboidratos. Os transcritos apresentaram maior atividade no período em que o inseto estava se alimentando da cana-de-açúcar e do botão floral, respectivamente. Os dados gerados e discutidos neste estudo confirmam a clonagem e expressão de celulasas dos insetos-praga, *T.licus licus* e *A. grandis*, uma β -glicosidase e uma endo-1,4- β -glucanase, respectivamente. Essas celulasas expressas no sistema heterólogo *P.pastoris*TM, secretadas no sobrenadante, são funcionalmente ativas. Com esses dados iniciais, será possível contribuir com estudo sobre a bioquímica digestiva dessas pragas, que poderão ser aplicadas na produção de enzimas em processos biotecnológicos industriais.

REFERÊNCIAS

- AGÊNCIA NACIONAL DO PETRÓLEO. <http://www.anp.gov.br>. Dados estatísticos. Anuário Estatístico 2016. Consultado em 26 de junho de 2016.
- ALMEIDA, J. R. M. de; COLLARES, D. G.; BARBOSA, P. F. D. Bioprospecção microbiana. Repositório: BDPA. Ano de publicação: 2015.
- ALVIRA, P; TOMÁS-PEJÓ, E; BALLESTEROS, M; NEGRO, MJ. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. **Bioresour Technol.** 2010 Jul;101(13):4851-61. doi: 10.1016/j.biortech.2009.11.093. Epub 2009 Dec 29.
- ANDRADE, TCC; BITENCOURT, WC; BOMTEMPO, FVS; ALVES, FS; BARBOSA, RS; GUARDA, EA. Hidrólise Enzimática de Celulose para Obtenção de Glicose Utilizando Líquido Iônico como Meio Solvente. **Rev. Virtual Quim.**, 2019, 11 (1), 310-325.
- ANTUNES, FAF; CHANDEL, AK; SANTOS, JC; MILESSI, TSS; PERES, GFD; SILVA, SS. Hemicellulosic Ethanol Production in Fluidized Bed Reactor from Sugar Cane Bagasse Hydrolysate: Interplay among Carrier Concentration and Aeration Rate. **CS Sustainable Chem. Eng.** 2017, 5, 9, 8250-8259. <https://doi.org/10.1021/acssuschemeng.7b01916>
- ARAUJO, A. C. B.; CARVALHO, L. S.; TUPINAMBA, D. D.; ALMEIDA, J. M.; FAVARO, L. C. de L.; KRUGER, R. H.; QUIRINO, B. F. Isolamento e Caracterização de Microrganismos Isolados a Partir de Solo de Cerrado para a Produção de Etanol de Segunda Geração. Ano de publicação: 2012. Unidade: Embrapa Agroenergia.
- BAYER, EA; CHANZY, H; LAMED, R; SHOHAM Y. Cellulose, Cellulases and Cellulosomes. **Curr Opin Struc Biol** 8: 548–557, 1998.
- BELLOT, JL. O bicudo-do-algodoeiro (*Anthonomus grandis* BOH., 1843) nos cerrados brasileiros: biologia e medidas de controle. Instituto Mato-grossense do Algodão - IMAmt 2015. 254 p.: il.; 14,8 x 21 cm ISBN: 978-85-66457-07-0
- BENOLIEL, B; POÇAS-FONSECA, MJ; TORRES, FA; DE MORAES, LM. Expression of a Glucose-tolerant β -glucosidase from *Humicola grisea* var. *thermoidea* in *Saccharomyces cerevisiae*. **Applied Biochemistry and Biotechnology** v.160, p.2036-2044, 2010.
- BONFÁ, EC. Purificação parcial e caracterização bioquímica de uma isoforma de β -glicosidase do fungo termofílico *Myceliophthora thermophila* M.7.7. São José do Rio Preto, SP, 2016.

- BRADFORD, MA. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, v.72, n.1/2, p.248-254, 1976.
- BRETHAUER, S; STUDER, M H. Biochemical conversion processes of lignocellulosic biomass to fuels and chemicals - a review. **Chimia**, v. 69, n. 10, p. 572-581, 2015.
- BRIJWANI, K; OBEROI, HS; VADLANI, PV. Production of a cellulolytic enzyme system in mixed-culture solid-state fermentation of soybean hulls supplemented with wheat bran. **Process Biochemistry**, v. 45, n. 1, p. 120-128, 2010.
- BROWN, TR & BROWN, RC. A review of cellulosic biofuel commercial-scale projects in the United States. **Biofuels, Bioproducts and Biorefining**, 2013, 7(3), 235-245. doi: 10.1002/bbb.1387
- BUCKERIDGE, M.S.; DE SOUZA, A.P. Breaking the “Glycomic Code” of Cell Wall Polysaccharides May Improve Second-Generation Bioenergy Production from Biomass. **Bioenergy Research**, on line 24 april 2014.
- BUCKERIDGE, M.S. The evolution of the Glycomic Codes of Extracellular Matrices. **Biosystems** 164:112-120. 2018 doi: [10.1016/j.biosystems.2017.10.003](https://doi.org/10.1016/j.biosystems.2017.10.003).
- BUSCH A; KUNERT G; WIELSCH N; PAUCHET Y. Cellulose degradation in *Gastrophysa viridula* (Coleoptera: Chrysomelidae): functional characterization of two CAZymes belonging to glycoside hydrolase family 45 reveals a novel enzymatic activity. **Insect Mol Biol**. 2018 Oct;27(5):633-650. doi: 10.1111/imb.12500.
- BUSSAMRA, BC; FREITAS, SF; COSTA, AC. Improvement on sugar cane bagasse hydrolysis using enzymatic mixture designed cocktail. **Bioresource Technology**, v. 187, p. 173-181, 2015.
- BUSOLI, AC; SOARES, JJ; LARA, FM. O bicudo do algodoeiro e seu manejo. Jaboticabal: FUNEP. 32p. (Boletim, 5), 1994.
- BYEON, GM; LEE, KS; GUI, ZZ; KIM, I; KANG, PD; LEE, SM; SOHN, HD; JIN, BR. A digestive beta-glucosidase from the silkworm, *Bombyx mori*: cDNA cloning, expression and enzymatic characterization. **Comp Biochem Physiol B Biochem Mol Biol**. 2005 Aug;141(4):418-27.
- CALDERÓN-CORTÉS, N; WATANABE, H; CANO-CAMACHO, H; ZAVALA-PÁRAMO, G; QUESADA, M. cDNA cloning, homology modelling and evolutionary insights into novel endogenous cellulases of the borer beetle *Oncideres albomarginata* chamela (Cerambycidae). **Insect Mol Biol**. 19, 323–336, <https://doi.org/10.1111/j.1365-2583.2010.00991.x> (2010).

CALIL NETO, A; GUIMARÃES, MJOC; FREIRE, E. Panorama tecnológico da produção de bioetanol de 1ª e 2ª geração. **Revista Alcoolbrás**. Edição nº 142, pp. 39-41, 2013.

CAMARGO, PD. Força verde: um novo campo para a indústria química. **Revista Brasileira de Engenharia Química**, p. 18-21, 2007.

CAMPOS, C. G. Desenvolvimento de protocolos analíticos em metabolômica de leveduras fermentadoras de xilose. 2015. Dissertação (Mestrado em Química) - Universidade de Brasília, Brasília, DF. Orientadora CNPq: Patrícia Verardi Abdelnur.

CARDONA, CA; SÁNCHEZ, ÓJ. Fuel ethanol production: Process design trends **Dahod**. 2007.

CARVALHO, S. B.; FRANCO, P. F.; ALMEIDA, J. R. M. de. Construção de levedura *Saccharomyces cerevisiae* capaz de fermentar xilose para produção de etanol. **Repositório: BDPA**. Ano de publicação: 2015.

CAVALCANTI, CGC. Metabolômica de leveduras fermentadoras de xilose na identificação de alvos metabólicos para aumento da produção de etanol 2G. Goiânia, GO, 2019. Tese apresentada ao Instituto de Química da Universidade Federal de Goiás Orientador: Prof. Dra. Patrícia Verardi Abdelnur, pesquisadora da Embrapa Agroenergia.

CHA, M; CHUNG, D; ELKINS, JG; GUSS, AM. Westpheling J Metabolic engineering of *Caldicellulosiruptor bescii* yields increased hydrogen production from lignocellulosic biomass. **Biotechnology for Biofuels** volume 6, Article number: 85 (2013).

CHELLAPANDI, P; HIMANSHU, MJ. Production of endoglucanase by the native strains of *Streptomyces* isolates in submerged fermentation. **Braz. J. Microbiol.** v.39, n.1, p. 122-127, 2008.

CHEN, HL; CHEN, YC; LU, MYJ; CHANG, JJ; WANG, HTC; KE, HM; WANG, TY; RUAN, SK; WANG, TY; HUNG, KY; CHO, HY; LIN, WT; SHIH, MC; LI, WH. A highly efficient β -glucosidase from the buffalo rumen fungus *Neocallimastix patriciarum*. **Biotechnology for Biofuels** 2012, 5:24, 2012.

CHEN, M; SMITH, PM; WOLCOTT, MP. U.S. Biofuels Industry: A Critical Review of Opportunities and Challenges. **BioProducts Business** 1(4) 2016.

CHUNG, D; CHA, M; SNYDER, E; ELKINS, J; GUSS, A; WESTPHELING, J. Cellulosic ethanol production via consolidated bioprocessing at 75 °C by engineered *Caldicellulosiruptor bescii*. **Biotechnol Biofuels**. 2015, 8: 163. <https://doi.org/10.1186/s13068-015-0346-4>

CONAB. Companhia Nacional de Abastecimento. Acompanhamento da safra brasileira de cana-de-açúcar. – v. 1 – Brasília. Quadrimestral. ISSN 2318-7921. Disponível em: <http://www.conab.gov.br>. Acesso em: 05/03/2019.

COSTA, O. Y. A.; SOUTO, B. de M.; TUPINAMBA, D. D.; BERGMANN, J. C.; KRUGER, R. H.; KYAW, C. M.; BARRETO, C. C.; QUIRINO, B. F. Avaliação da diversidade microbiana do processo de produção de etanol de cana-de-açúcar por meio de abordagem independente de cultivo. **Repositório: BDPA**. Ano de publicação: 2015.

DAMASO, M. C. T.; FARINAS, C. S.; SALUM, T. F. C. Produção e imobilização de enzimas aplicadas à produção de etanol e biodiesel. **Repositório: BDPA**. Ano de publicação: 2012.

DANCHIN, EGJ; ROSSO, MN. Lateral gene transfers have polished animal genomes: lessons from nematodes. *Front in Cell. And Infect. Microbiol.* 2 (27) <https://doi.org/10.3389/fcimb.2012.00027> (2012).

DE SCHUTTER, K; LIN, Y; TIELS, P; VAN HECKE, A; GLINKA, S; WEBER-LEHMANN, J; ROUZE, P; VAN DE PEER, Y; CALLEWAERT, N. Genome sequence of the recombinant protein production host *Pichia pastoris*. **Nat. Biotechnol.** 27:561-566 (2009).

ENGDAHL, F.W. Behind Oil Price Rise: Peak Oil or Wall Street Speculation? Disponível em: <http://www.voltairenet.org>. 2012. Acesso em: 10 /05/2013. **Engineering**, v. 66, p. 36-56, 2014.

FERNANDES, MC; FERRO, MD; PAULINO, AFC; MENDES, JAS; GRAVITIS, J; EVTUGUIN, DV; XAVIER, AMRB. Enzymatic saccharification and bioethanol production from *Cynara cardunculus* pretreated by steam explosion. **Bioresource Technology**. v. 186, p. 309-315, 2015.

FIRMINO, AA; FONSECA, FC; DE MACEDO, LL; COELHO, RR ; ANTONINO DE SOUZA JD, JR; TOGAWA, RC; SILVA-JUNIOR, OB; PAPPAS GJ, JR; DA SILVA MC; ENGLER, G; GROSSI-DE-AS, MF. Transcriptome analysis in Cotton Boll Weevil (*anthonomus grandis*) and RNA interference in insect pests. **PLoS One**. <https://doi.org/10.1371/journal.pone.0085079> (2013).

FISCHER, R; OSTAFE, R; TWYMAN, RM. Cellulases from insects. **Advances in Biochemical Engineering/biotechnology**, 136, 51–64. doi:10.1007/10_2013_206, 2013.

FONSECA, FC; DE MACEDO, LL; COELHO, RR; ANTONINO DE SOUZA JD, JR; TOGAWA, RC; SILVA-JUNIOR, OB; PAPPAS GJ, JR; DE GÓIS, LA; DA SILVA, MC; GROSSI-DE-SÁ, MF. Sugarcane Giant Borer Transcriptome Analysis and Identification of Genes Related to Digestion. **PLoS One**. 2015; 10(2): e0118231.

GABRIŠKO M. & JANEČEK Š. Looking for the ancestry of the heavy-chain subunits of heteromeric amino acid transporters rBAT and 4F2hc within the GH13 α -amylase family. **FEBS journal** 276, 7265–7278, 2009.

GABRISKO M. & Janecek S. Characterization of maltase clusters in the genus *Drosophila*. **J Mol Evol** 72, 104–118, 2011.

GALBE, M; ZACCHI, G. Pretreatment: the key to efficient utilization of lignocellulosic materials. **Biomass and Bioenergy**, v. 46, p. 70-78, 2012.

GALLO, D; NAKANO, O; SILVEIRA NETO, S; CARVALHO, RPL; BATISTA, G C; BERTI FILHO, E; PARRA, JRP; ZUCCHI, RA; ALVES, SB; VENDRAMIN, JD. **Manual de entomologia agrícola**. 2 ed. São Paulo: Ceres, 486p, 1988.

GANDOLFI, CB; HARMANTAS, A; SOLER-LEONARTE, O. Abundant Clean Energy Fuels Brazil's Growth. **Power**. 156, p.46, 2012.

GIRARD, C; JOUANIN, L. Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae*. **Insect Biochem. Mol. Biol.** 29:1129–42, 1999.

GLASS, NL; SCHMOLL, M; CATE, J H; CORADETTI, S. Plant cell wall deconstruction by ascomycete fungi. **Ann. Rev. Microbiol.** 67(1), 2013. 477–98, //doi.org/10.1146/annurev-micro-092611-150044

GONDIM, DMC; JEAN-LOUIS, B; SILVIE, P; PEITI, N. Manual de identificação das pragas, doenças, deficiências minerais e injúrias do algodoeiro no Brasil. 3a ed. Cascavel: COODETEC/CIRAD-CA, 120p. (B.T. N° 33), 1999.

HAHN, MW; HAN, MV; HAN, SG. Gene family evolution across 12 *Drosophila* genomes. **PLoS Genetics** 3 : e197, 2007.

HARNPICHARNCHAI, P; CHAMPREDA, V; SORNLAKE, W; EURWILAICHITR, L. A thermotolerant beta-glucosidase isolated from an endophytic fungi, *Periconia* sp., with a possible use for biomass conversion to sugars. **Protein Expr. Purif.** 67, 61–69, 2009.

HASUNUMA, T; OKAZAKI, F; OKAI, N; HARA, KY; ISHII, J. A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. **Bioresource Technol.** 135((May)), 513–22, 2013. //doi.org/10.1016/J.BIORTECH.2012.10.047.

HUANG, S; DENG, G; YANG, Y; WU, Z; WU, L. Optimization of Endoglucanase Production from a Novel Bacterial Isolate, *Arthrobacter* sp. HPG166 and

characterization of Its Properties. **Braz. arch. biol. Technol.** v.58, n.5, p. 692-701, 2015.

IBARRA LN; ALVES AEOA; ANTONINO JD; PRADO GS; PINTO CEM; SOCCOL CR; DE VASCONCELOS ÉAR; GROSSI-DE-SA MF. Enzymatic activity of a recombinant β -1,4-endoglucanase from the Cotton Boll Weevil (*Anthonomus grandis*) aiming second generation ethanol production.**Sci Rep.** 2019 Mar 19;10(1):5367. doi: 10.1038/s41598-020-62217-2.

INVITROGEN™ pPICZ A, B, & C *Pichia* Vectors. *Pichia* expression vectors for selection on Zeocin™ and purification of recombinant proteins. Catalog no. V190-20 Rev. Date: 7 July 2010 Manual part no. 25-0148.

JAISWAL, D; DE SOUZA, AP; LARSEN, S; LEBAUER, D; MIGUEZ, FE; SPAROVEK, G; BOLLERO, G; BUCKERIDGE, M.S; Long, SP. Brazilian sugarcane ethanol as an expandable green alternative to crude oil use. **Nature Climate Change.** 7, 788–792. 2017

JONES, RT; SANCHEZ, LG; FIERER, NA. A cross-taxon analysis of insect-associated bacterial diversity. **PLoS One.** San Francisco, v. 8, n.4, p. e61218, 2013.

JORDAN, DB; BOWMAN, MJ; BRAKER, JD; DIEN, BS; HECTOR, RE; LEE, CC; MERTENS, JA; WAGSCHAL K. Plant cell walls to ethanol. **Biochem. J.** 442, 241–252, 2012.

KARBOUNE S., GERAERT P.A., KERMASHA S. Characterization of selected cellulolytic activities of multi-enzymatic complex system from *Penicillium funiculosum*. **Journal of Agricultural and Food Chemistry.** 56: 903-909, 2008.

KIM SK, CHUNG D, HIMMEL ME, BOMBLE YJ, WESTPHELING J. Engineering the N-terminal end of CelA results in improved performance and growth of *Caldicellulosiruptor bescii* on crystalline cellulose. **Biotechnol Bioeng.** 2017 May;114(5):945-950. doi: 10.1002/bit.26242.

KIRSCH, R.; KUNERT, G; VOGEL, H; PAUCHET, Y. Pectin digestion in herbivorous beetles: impact of pseudoenzymes exceeds that of their active counterparts. **Front. Physiol.** 10, 685, 2019.

KOHLHEPP, G. Análise da situação da produção de etanol e biodiesel no Brasil. **Estudos Avançados.** São Paulo, v. 24, n. 68, 2010.

KROGH, KB; HARRIS, PV; OLSEN, CL; JOHANSEN, KS; HOJER-PEDERSEN, J; BORJESSON, J; OLSSON, L. Characterization and kinetic analysis of a thermostable GH3 B-glucosidase from *Penicillium brasilianum*. **Applied. Microbiology and Biotechnology,** v.86, n.1, p.143-154, 2010.

KUMAR, R., SINGH, S. & SINGH, O. V. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. **J. of Ind. Microbiol. & Biotech.** **35** (5), 377–91, 2008.
<https://doi.org/10.1007/s10295-008-0327-8>.

KURTZMAN, CP; FELL, JW; BOEKHOUT, T. The yeasts, a taxonomic study. **Elsevier Science**, v.2, 1062p., 2011.

LEE, LL; CROSBY, JR; RUBINSTEIN, GM; LAEMTHONG, T; BING, RG; STRAUB, CT; ADAMS, MW; KELLY, RM. The biology and biotechnology of the genus *Caldicellulosiruptor*: recent developments in 'Caldi World'. **Extremophiles**. 2020 Jan;24(1):1-15. doi: 10.1007/s00792-019-01116-5. Epub 2019 Jul 29.

LI, XH; YANG, HJ; ROY, B; PARK, EY; JIANG, LJ; WANG, D; MIAO, YG. Enhanced cellulase production of the *Trichoderma viride* mutated by microwave and ultraviolet. **Microbiol Res**, 165 (3). pp. 190 – 198, 2010.

LI, XH; SHI, L; YANYAN, Z; HONGQING, X; XIANGPING, D.; RONGQIAO L; YUYIN, C; HUABING, W. Molecular evolutionary mechanisms driving functional diversification of α -glucosidase in Lepidoptera. **Scientific Reports**. 7: 45787. 2017. doi: [10.1038/srep45787](https://doi.org/10.1038/srep45787)

LIU, J; ZHANG, X; FANG, Z; FANG, W; PENG, H; XIAO, Y. The 184th residue of beta-glucosidase Bgl1B plays an important role in glucose tolerance. **J. Biosci. Bioeng.** 112, 447–450. 2011.

LIU, D; ZHANG, R; YANG, X; ZHANG, Z; SONG, S; MIAO, Y; SHEN, Q. Characterization of a thermostable beta-glucosidase from *Aspergillus fumigatus* Z5, and its functional expression in *Pichia pastoris* X33. **Microb. Cell Fact.** 11, 25, 2012.

LORENZI, BR; ANDRADE, THN. Review: Continuous hydrolysis and fermentation for cellulosic ethanol production. **Rev. bras. Ci. Soc.** vol.34 no.100 São Paulo, 2019. Epub Sep 23, 2019. <https://doi.org/10.1590/3410014/2019>

LYND, LR; WEIMER, PJ; VAN ZYL, WH. PRETORIOUS, I.S. Microbial cellulase utilization: fundamentals and biotechnology. **Microbiol. Mol. Biol. Rev.** 66, 506–577, 2002.

MACAULEY-PATRICK, S; FAZENDA M.L; MCNEIL, B; HARVEY L.M. Heterologous protein production using the *Pichia pastoris* expression system. **Yeast** 22:249-270, 2005.

MACHADO, CMM; DAMASO, MCT; SOUZA, DT de; BRAGA, M; ONOYAMA, M M; GARCIA, LC; QUIRINO, BF; RODRIGUES, D de S; CARVALHO, F B de

- P; ALMEIDA, J R M.de; VALADARES, L F; GONCALVES, SB; MENDES, T D; PACHECO, TF. Contextualização, mercado e gargalos de P&D do etanol lignocelulósico. **Repositório: BDPA**. Ano de publicação: 2016.
- MARANA, SR; TERRA, WR; FERREIRA, C. Purification and properties of a β -glycosidase purified from midgut cells of *Spodoptera frugiperda* (Lepidoptera) larvae **Insect Biochemistry and Molecular Biology**.30 (2000) 1139–1146
- MARTIN, D F; BARBOSA, S; CAMPANHOLA, C. Observações preliminares e comentários sobre o bicudo do algodoeiro, no Estado de São Paulo. Jaguariúna, EMBRAPA-CNPDA. 21p. (EMBRAPA-CNPDA, Circular Técnica, 1), 1987.
- MARTINS, FA; MARTIM, T; CORRÊA, AM; OLIVEIRA, FF. A produção de etanol de segunda geração a partir do bagaço da cana-de-açúcar. **Rev. Lat.-Am. Inov. Eng. Prod. [ReLAIInEP]**, Curitiba (PR), Brasil. ISSN-e: 2317-6792, 2018.
- MENDONÇA, AF; BALDANI, JI; GROSSI- DE- SÁ, MF; SILVA, MCM; MACEDO, LLP; FONSECA, FCA; BARBOSA, CRC; GUZZO, NEC. A broca gigante da cana-de-açúcar, *T. licus licus* Drury, (Lepidoptera: Castniidae), p. 133–167. In A.F. Mendonça (ed.), *Pragas da cana-de-açúcar*. Maceió, Insetos & Cia, 239 p. 1996.
- MOTOYUKI S; YUAN X. L. & FUJIYOSHI M. Sucrase activity and its kinetic properties in peritrophic membrane, and in membrane-bound and soluble fractions of midgut in the silkworm, *Bombyx mori*. *Comparative Biochemistry and Physiology Part A: Physiology* 108, 255–264 (1994).
- NANDA, S; MOHAMMAD, J; REDDY, SN. Pathways of lignocellulosic biomass conversion to renewable fuels. **Biomass Conversion and Biorefinery**, v. 4, n. 2, p. 157-191, 2014.
- NEGRISOLI JUNIOR, AS; BALDANI, JI; GROSSI- DE- SÁ, MF; SILVA, MCM; MACEDO, LLP; FONSECA, FCA; BARBOSA, CRC; GUZZO, NEC. Manejo da broca-gigante da cana-de-açúcar (*Telchin licus*) (Drury) (Lepidoptera: Castniidae) no nordeste do Brasil. Aracaju: Embrapa Tabuleiros Costeiros, 2015. 50 p. (Documentos / Embrapa Tabuleiros Costeiros, ISSN 1678-1953; 198).
- NYKO, D; GARCIA, JLF; MILANEZ, AY; DUNHAM, FB. A corrida tecnológica pelos biocombustíveis de segunda geração: uma perspectiva comparada. **BNDES Setorial**, n. 32, p. 5-48. BNDES, Rio de Janeiro, 2018.
- PACHECO, TF. Produção de Etanol: Primeira ou Segunda Geração? Circular Técnica 04: **Embrapa Agroenergia Abril**, ISSN 2177-4420, 2011.
- PAUCHET, Y; WILKINSON, P; VAN MUNSTER, M; AUGUSTIN, S; PAURON, D; FFRENCH-CONSTANT, RH. Pyrosequencing of the midgut transcriptome of the

poplar leaf beetle *Chrysomela tremulae* reveals new gene families in Coleoptera. **Insect Biochemistry and Molecular Biology** 39: 403–413, 2009.

PAUCHET, Y; WILKINSON, P; VOGEL, H; NELSON, DR; REYNOLDS, SE; HECKEL, DG; FFRENCH-CONSTANT, RH. Pyrosequencing the *Manduca sexta* larval midgut transcriptome: messages for digestion, detoxification and defence. **Insect Molecular Biology** 19: 61–75, 2010.

PAUCHET, Y; WILKINSON, P; CHAUHAN, R. & FFRENCH-CONSTANT, R. H. Diversity of beetle genes encoding novel plant cell wall degrading enzymes. **PLOS ONE**. 5. <https://doi.org/10.1371/journal.pone.0015635>, 2010.

PETROBRAS, 2017. “RenovaBio: Diretrizes Estratégicas para Biocombustíveis”. Petróleo Brasileiro S.A. Disponível em ww.novacana.com/pdf/12042017140446_PE_TROBRAS__Relatorio_Tecnico_-_Consul ta__Publica__Renovabio.pdf, consultado em 11/5/2019.

PINTO, AS; GARCIA, J F; OLIVEIRA, HN. Manejo das principais pragas da cana-de-açúcar. In: SEGATO, S. V.; PINTO, A. S.; JENDIROBA, E.; NÓBREGA, J. C. M. de. Atualização em produção de cana-de-açúcar. **Piracicaba: Prol**, p. 107-120, 2006.

POTHULA, R; SHIRLEY, D; PERERA, OP; KLINGEMAN, WE; OPPERT,C; ABDELGAFFAR, HMY; JOHNSON, BR; JURAT-FUENTES, JL. The digestive system in *Zygentoma* as an insect model for high cellulase activity. **PLoS One**. 2019 Feb 28;14(2):e0212505. doi: 10.1371/journal.pone.0212505

PROMEGA. Technical Manual PowerQuant® System Instructions for Use of Products PQ5002 and PQ5008.

PROTASIO, TP; BUFALINO, L; MENDES, LM. Brazilian lignocellulosic wastes for bioenergy production: characterization and comparison with fossil fuels. **Bioresources**, 8, 1166-1185, 2013.

PULZ, O.; GROSS, W. Valuable products from biotechnology of microalgae. **Applied Microbiology Biotechnology**, v.65, p.635-648, 2004.

RABELLO, G. C. PIROTA, R. D. P. B. BARROS, G.; FARINAS, C. S. Simultaneous biosynthesis of biomass-degrading enzymes using co-cultivation of *Aspergillus niger* and *Trichoderma reesei*. **Biocatalysis and Biotransformation** [S. l.], v. 32, p. 236-243, 2014.

RIOU, C; SALMON, JM; VALLIER,MJ; GÜNATA, Z; BARRE, P. Purification, characterization, and substrate specificity of a highly glucose-tolerant B-glucosidase from *Aspergillus oryzae*. **Applied. Microbiology and Biotechnology**, v.64, n.10, p. 3607-3614, 1998.

SAADEDIN A The complexities of hydrolytic enzymes from the termite digestive system. **Crit Rev Biotechnol.** 2014 Jun;34(2):115-22. doi: 10.3109/07388551.2012.727379. Epub 2012 Oct 5.

SANTANA, H. Isolamento, seleção e caracterização de microalgas com alta produtividade de biomassa em meio de cultivo a base de vinhaça e CO₂. 2016. Dissertação (Mestrado em Biociências) - Universidade Federal da Bahia, Instituto Multidisciplinar em Saúde, Vitória da Conquista, BA. Orientador CNPq: Dr. Félix Gonçalves de Siqueira. Co-orientador CNPq: Dr. Bruno dos Santos Alves Brasil.

SANTOS, MFRF dos; BORSCHIVER, S; COUTO, MAPG. Iniciativas para o uso da biomassa lignocelulósica em biorrefinarias: a plataforma sucroquímica no mundo e no Brasil. **Economia & Energia**, n. 82, p. 14-32, 2011.

SANTOS, FP dos. Expressão e produção de monoxigenases bacterianas em *Komagataella phaffii* (*Pichia pastoris*) para utilização como enzimas acessórias para desconstrução de biomassa. 2017. Dissertação (Mestrado em Biotecnologia) - Universidade Federal do Tocantins, Gurupi, TO. Orientador CNPq: Bruno dos Santos Alves Figueiredo Brasil. Coorientadora CNPq: Thaís Fabiana Chan Salum.

SAQIB, AA; FAROOQ, A; IQBAL, M; HASSAN, J.U; HAYAT, U; BAIG, S. A. Thermostable crude endoglucanase produced by *aspergillus fumigatus* in a novel solid state fermentation process using isolated free water. **Enzyme Res.**v. 2012, n. 2012, p. 6, 2012.

SEGATO, SV; MATTIUZ, CFM; MOZAMBANI, AE. Atualização em produção de cana-de-açúcar. **Piracicaba: Livroceres**, p.19-36, 2006.

SILVÉRIO, RM; POPPE, MK; MORESI, EAD; ROCHA, BB; SILVA, ECC; MAIA, JMF; SILVA, TRC. Relatório prospecção etanol de segunda geração - E2G 2030. Panorama analítico prospectivo dos biocombustíveis e bioprodutos. Projeto Agenda Positiva da Mudança do Clima e do Desenvolvimento Sustentável. Brasília: **Centro de Gestão e Estudos Estratégicos**, CGEE, 2018.

SILVIE, P; LEROY, T; BELOT, J-L; MICHEL, B. Manual de identificação das pragas, e seus danos no algodoeiro. 1ª ed. Cascavel: COODETEC/CIRAD-CA, 2001.100p. (B.T. Nº 34), 2001.

SINGHANIA, RR; PATEL, AK; SUKUMARAN, RK; LARROCHE, C; PANDEY, A. Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production. **Bioresource Technology.** 127, 500–507, 2013.

SINGH, R.; KUMAR, M.; MITTAL, A.; MEHTA, P. K. Microbial enzymes: industrial progress in 21st century. **3 Biotech**, Índia, v. 6, n. 2, p. 1–15, 2016.

SOUZA, LTA; VERÍSSIMO, LAA; PESSELA, BC; SANTORO, JMM; RESENDE, RR;

MENDES, AA. Imobilização enzimática: princípios fundamentais e tipos de suporte. In book: **Biotecnologia Aplicada à Agro&Indústria** - Vol. 4, pp.529-568. 2017. DOI: 10.5151/9788521211150-15

SOUZA WR, MARTINS PK, FREEMAN J, PELLNY TK, MICHAELSON LV, SAMPAIO BL, VINECKY F, RIBEIRO AP, DA CUNHA BADB, KOBAYASHI AK, DE OLIVEIRA PA, CAMPANHA RB, PACHECO TF, MARTARELLO DC, MARCHIOSI R, FERRARESE-FILHO O, DOS SANTOS WD, TRAMONTINA R, SQUINA FM, CENTENO DC, GASPAR M, BRAGA MR, TINÉ MAS, RALPH J, MITCHELL RAC, MOLINARI HBC. Suppression of a single BAHD gene in *Setaria viridis* causes large, stable decreases in cell wall feruloylation and increases biomass digestibility. **New Phytol.** 2018 Apr;218(1):81-93. doi: 10.1111/nph.14970.

SRIVASTAVA, N; SRIVASTAVAC, M; MISHRAB, PK; GUPTAD, VK; MOLINAF, G; RODRIGUEZ-COUTOG, S; MANIKANTAB, A; RAMTEKEA, PW. Applications of fungal cellulases in biofuel production: Advances and limitations. **Renew. Sust. Energ. Rev.** 82,2379–2386, 2018. <https://doi.org/10.1016/j.rser.2017.08.074>

SUN, JZ; SCHARF, ME. Exploring and integrating cellulolytic systems of insects to advance biofuel technology. **Insect Science** 17: 163–165, 2010.

SZILÁGYI, E; HÁMORI, C; BÍRÓ-MOLNÁR, P; KANDRA, L; REMENYIK, J; GYÉMÁNT, G. Cooperation of enzymes involved in carbohydrate digestion of Colorado potato beetle (*Leptinotarsa decemlineata*, Say). **Bull Entomol Res.** 2019 Oct;109(5):695-700. doi: 10.1017/S0007485319000099. Epub 2019 Apr 11.

TERRA, WR; FERREIRA, C. The physiological role the peritrophic membrane and trehalase: digestive enzymes in the midgut and excreta of starved larvae of *Rhynchosciara Americana*. **J.Insect Physiol.** 1981. 27, 325-331, 1981.

TERRA, WR; FERREIRA, C. Insect digestive enzymes: properties, compartmentalization and function. **Comparative Biochemistry and Physiology** 109B: 1–62, 1994.

TERRA, WR; FERREIRA, C. Biochemistry of digestion. In: GILBERT, L. I., IATROV, K., GILL, S. (Org.). **Comprehensive Molecular Insect Science**, v. 4, p. 171-224, 2005.

TERRA, WF; FERREIRA, C. Biochemistry and Molecular Biology of Digestion. In: GILBERT, L. L (Ed.). **Insect Molecular Biology and Biochemistry.** p. 366-406, 2012.

TOLEDO, DM; BRODSKY, HA; PARDO, GE; CONCI, OC; BRAGA SOBRINHO, R. Monitoreo del Picudo del Algodonero (*Anthonomus grandis* Bh.) em el Noreste Argentino. Universidad Nacional Del Nordeste, **Comunicaciones Cientificas y Tecnológicas**, 2000.

VOLYNETS, B; EIN-MOZAFFARI, F; DAHMAN, Y. Biomass processing into ethanol: pretreatment, enzymatic hydrolysis, fermentation, rheology, and mixing. **Green Processing and Synthesis**, v. 6, p. 1-22, 2017.

YAPI, AY; GNAKRI, D; LAMINE, NS; PATRICE, KL. Purification and biochemical characterization of a specific beta-glucosidase from the digestive fluid of larvae of the palm weevil, *Rhynchophorus palmarum*. **J Insect Sci.** 2009;9:4. doi: 10.1673/031.009.0401.

YASUR, J; USHA, R; PATHIPA, TI. Lepidopteran insect susceptibility to silver nanoparticles and measurement of changes in their growth, development and physiology. **Chemosphere**. Volume 124, Pages 92-102. 2015. <https://doi.org/10.1016/j.chemosphere.2014.11.029>

YEOMAN, CJ; HAN, Y; DODD, D; SCHROEDER, CM; MACKIE, RI; CANN, IK. Thermostable enzymes as biocatalysts in the biofuel industry. **Advances in Applied Microbiology**. v. 70, p. 1–55, 2010.

YOON, LW; ANG, TN; CHUA, ASM. Fungal solid-state fermentation and various methods of enhancement in cellulase production. **Biomass and Bioenergy**, v. 67, p. 319-338, 2014.

YUE, D; YOU, F; SNYDER, SW. Biomass-to-bioenergy and biofuel supply chain optimization: overview, key issues and challenges. **Computers and Chemical Engineering**, v. 66, p. 36-56, 2014.

ZHANG, YHP; LYND, LR. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. **Biotechnology and Bioengineering**, v. 88, n. 7, p. 797-824, 2004.

ZHOU, X; SMITH, JA; OI, FM; KOHLER, PG; BENNETT, GW; SCHARF, ME. Correlation of cellulose genes expression and cellulolytic activity throughout the gut of the termite *Reticulitermes flavipens*. **Gene** 395:29–39, 2007.

ZILBERMAN, D; TIMILSINA, GR. The impacts of biofuels on the economy, environment, and poverty. A global perspective. 2014. **Springer**, New York.

ZÚÑIGA, U.F.R. Desenvolvimento de um bioprocesso para a produção de celulasas específicas na cadeia produtiva do etanol de segunda geração. Tese de doutorado. Escola de Engenharia de São Carlos, Universidade de São Paulo, 2010.

WANG, H; KIUCHI, T; KATSUMA, S; SHIMADA, T. A novel sucrose hydrolase from the bombycoid silkworms *Bombyx mori*, *Trilocho varians*, and *Samia cynthia*

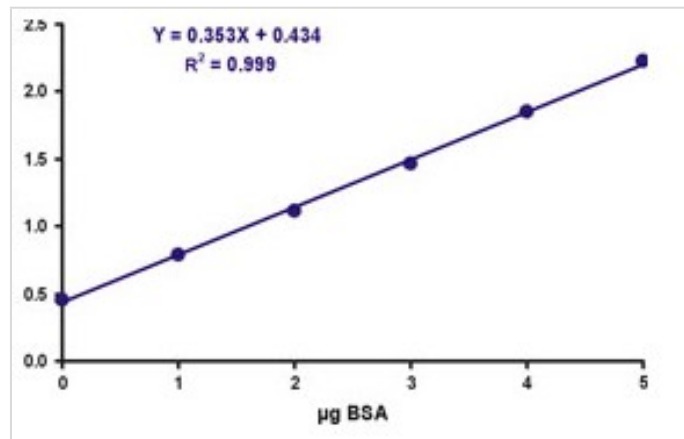
ricini with a substrate specificity for sucrose. **Insect Biochem Mol Biol** 61, 46–52, 2015.

WATANABE, H; TOKUDA, G. Cellulolytic systems in insects. **Annu Rev Entomol.** 2010;55:609-32. doi: 10.1146/annurev-ento-112408-085319.

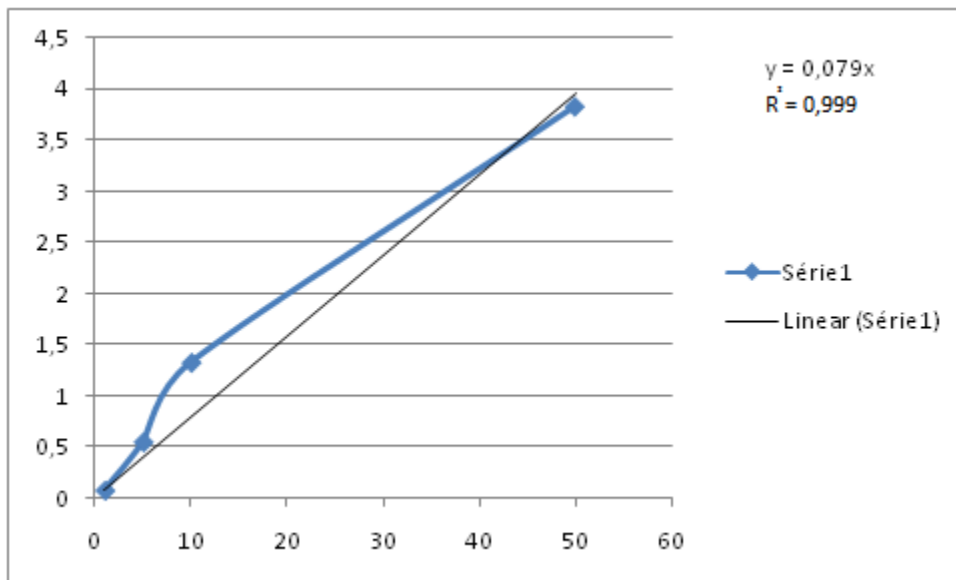
WILLIS, J D; OPPERT, C; JURAT-FUENTES, J L. Methods for discovery and characterization of cellulolytic enzymes from insects. **Insect Science**, 17(3), 184–198. 2010. doi:10.1111/j.1744-7917.2010.01322.

WYSOCKA-KAPCINSK, M; SANDOVAL, JAC; PAL, A; FINDLAY, JB. Expression and characterization of recombinant human retinol-binding protein in *Pichia pastoris*. **Protein Expression and Purification**. Doi:10.1016/j.pep.2010.01.015.

WOLFENBERGER, DA; HAMED, AA; LUTTRELL, RG. Toxicity of *Bacillus thuringiensis* against the boll weevil *Anthonomus grandis* Boh. (Coleoptera, Curculionidae). **Beltsville Cotton Conferences**, p. 1296-1300, 1997.



ANEXO 2- Curva de calibração do ensaio com açúcares redutores.



ANEXO 3: Artigo publicado.

SCIENTIFIC REPORTS

nature research

Enzymatic activity of a recombinant β -1,4-endoglucanase from the Cotton Boll Weevil (*Anthonomus grandis*) aiming second generation ethanol production

Liz Nathalia Ibarra^{1,2,7}, Ana Elizabeth Oliveira de Araújo Alves^{2,3,7}, José Dijair Antonino^{2,4}, Guilherme Souza Prado^{2,5}, Clídia Eduarda Moreira Pinto^{2,3}, Carlos Ricardo Soccol¹, Érico Augusto Rosas de Vasconcelos^{6,7*} & Maria fátima Grossi-de-Sá^{2,5}

In the last years the production of ethanol fuel has started to change with the introduction of second-generation ethanol (2G Ethanol) in the energy sector. However, in Brazil, the process of obtaining 2 G ethanol did not reach a basic standard to achieve relevant and economically viable results. Several studies have currently been addressed to solve these issues. A critical stage in the ethanol production is the deployment of efficient and stable enzymes to catalyze the saccharification step into the process of biomass conversion. The present study comprises a screening for genes coding for plant biomass degradation enzymes, followed by cloning a selected gene, addressing its heterologous expression, and characterizing enzymatic activity towards cellulose derived substrates, with a view to second-generation ethanol production. A cDNA database of the Cotton Boll Weevil, *Anthonomus grandis* (Coleoptera: Curculionidae), an insect that feeds on cotton plant biomass, was used as source of plant biomass degradation enzyme genes. A larva and adult midgut-specific β -1,4-Endoglucanase-coding gene (*AgraGH45-1*) was cloned and expressed in the yeast *Pichia pastoris*. Its amino acid sequence, including the two catalytic domains, shares high identity with other Coleoptera Glycosyl Hydrolases from family 45 (GH45). *AgraGH45-1* activity was detected in a Carboxymethylcellulose (CMC) and Hydroxyethylcellulose (HEC) degradation assay and the optimal conditions for enzymatic activity was pH 5.0 at 50 °C. When compared to commercial cellulase from *Aspergillus niger*, *AgraGH45-1* was 1.3-fold more efficient to degrade HEC substrate. Together, these results show that *AgraGH45-1* is a valid candidate to be engineered and tested for 2G ethanol production.

In the last 44 years, Brazilian fuel industry makes progress in the production of ethanol fuel. Sugar cane is the main raw material used in the country for ethanol production, which has been very successful in recent years^{1,2}.

Until not that long ago, the country's production was limited to the fermentation of sugarcane molasses. However, since 2007 several 2G ethanol research funding projects have been established, with the founding of new laboratories and research institutions. These projects came to fruition with the start of large-scale 2 G ethanol production.

duction, when in 2014 the first three commercial 2G ethanol plants started operating³. Besides all scientific and technological foment, Brazilian 2G ethanol industry faces some problems, such as enzyme cost, to keep sustainable activities and supply the market fuel⁴.

¹Universidade Federal do Paraná - UFPR, Curitiba, PR, 81530-980, Brazil. ²Embrapa Recursos Genéticos e Biotecnologia, Brasília, DF, 70770-917, Brazil. ³Universidade de Brasília - UnB, Biology Institute, Brasília, DF, 70910-900, Brazil. ⁴Universidade Federal Rural de Pernambuco - UFRPE, Recife-PE, 52171-900, Brazil. ⁵Universidade Católica de Brasília-UCB, Brasília, DF, 70790-160, Brazil. ⁶Centro Universitário Euroamericano-UNIEURO, Brasília, DF, 70200-001, Brazil. ⁷These author contributed equally: Liz Nathalia Barran, Ana Elizabeth Oliveira de Araújo Alves and Érico Augusto Rosas de Vasconcelos.*email:ericoarv@gmail.com

The conversion of plant cell wall polysaccharides into second-generation ethanol has been studied by different approaches, ranging from the use of genetically modified plants, which are developed for obtaining plant cell walls that are more sensitive to enzyme degradation^{5,6}, up to engineering more efficient enzymes that are applied in chemical techniques for biofuel production^{7,8}. Specific enzymes acting on the hydrolysis of β -glycosidic bonds in cellulosic biomass are still a limiting factor in this respect. While the production of cellulases has increased

over the years, there has been a decrease in the costs of saccharification processes for the production of fermentable sugar^{7,9}. Studies about enzymes production and stabilization methods are in progress aiming to drop even more the cost for bioethanol production and other industrial applications, such as feed manufacture, laundry, and textile processing¹⁰⁻¹². This bottleneck in the development of an efficient system for the production of second-generation ethanol underscores the need for new enzymes that show high efficiency and stability, being consequently used in the conversion of biomass into fermentable products^{13,14}.

Although microorganisms, such as bacteria and fungi, are the primary source of cellulolytic enzymes¹⁵⁻¹⁷, recently it has been found that invertebrates like insects and nematodes possess a whole arsenal of enzymes that degrade plant cell wall, such as glycosyl hydrolases (Endoglucanases, polygalacturonases and xylanases) and other pectin-modifying enzymes¹⁸⁻²⁰. Remarkably, it has been proposed that insects and nematodes have acquired these enzymes from bacteria and fungi through horizontal gene transfer with many independent events occurring along time²¹. It has been proposed that the acquisition of these enzymes have permitted insects, especially from the Coleoptera order, to gain advantage efficiently exploring plant biomass²².

Regarding insects, many of these enzymes have been validated *in vivo* at different systems such as insect cell lines²³ and yeast, especially *Pichia pastoris* and *Saccharomyces cerevisiae*^{24,25}. Indeed, as far as we know, no microorganism had been engineered with any insect-derived enzyme other than an alpha-amylase from *Sitophilus oryzae* was used for simultaneous saccharification and fermentation (SSF) from raw starch aiming bioethanol production²⁶. Moreover, insect genomes

contain considerable amounts of genes coding for plant cell wall-degrading enzymes²⁷, especially coleopteran insects from Curculionoidea and Chrysomeloidea superfamilies²³. Several insects from the same family possess a number of these enzymes such *Sitophilus oryzae*²⁸, *Dendroctonus ponderosae* and *Hypotenemus hampei*²⁹. Specifically, the Cotton Boll Weevil, *Anthonomus grandis*, (Coleoptera: Curculionidae), feeds on cotton plant biomass from floral bud and cotton boll during its life cycle, so its digestive plant cell wall-degrading enzymes are essential for its nutrition. We have published a transcriptome of *A. grandis*³⁰ and it could represent an important source of genes coding to plant biomass degradation enzymes.

In this work, we report the cloning and heterologous expression of a β -1,4-endoglucanase gene specifically expressed in *A. grandis* larva and adult midgut (*AgraGH45-1*). The gene was expressed in *Pichia pastoris* heterologous system and the recombinant enzyme was shown to be efficient to degrade Carboxymethylcellulose (CMC) and Hydroxyethylcellulose (HEC) at pH 5.0 and 50 °C. Here we have showed that *AgraGH45-1* activity was 1,3-fold higher and more efficient to degrade HEC than commercial cellulase from *Aspergillus niger* at 50 °C. In addition, with temperatures ranging from 40 °C to 60 °C, the efficiency of *AgraGH45-1* is not as affected as the one from commercial *A. niger* cellulase.

Results and Discussion

In the Brazilian fuel industry, the process of obtaining 2 G ethanol did not reach a basic standard to conduct the production steps in order to achieve relevant and economically viable results. The lack of an agricultural and industrial system designed to make full use of sugarcane, biomass pretreatment, capital cost, pentose fermentation and enzyme cost are bottle necks that need to be overcome in order to keep sustainable industrial activity and supply Brazilian fuel market^{4,31}. Currently, several researches are being developed to solve these problems. Research focused on the 2 G ethanol production

efficiency, economical processing of raw materials³², use of all fermentable fractions³³, production of efficient and profitable enzymes^{34,35}, among others. Here we shown, in a first step, the efforts to clone a β -1,4-endoglucanase gene from Cotton Boll Weevil (*A. grandis*). After that, a functional recombinant enzyme (AgraGH45-1) was produced and evaluated about its efficiency to catalyze cellulose derived substrate.

Gene selection and AgraGH45-1 clone procedures for heterologous expression.

A review of the main enzymes used in the plant biomass fermentation industry was conducted by mining through data on current protocols of plant biomass fermentation for the generation of bioenergy. The glucose production from plant biomass requires a serial action of at least three main enzymes from the Glycosyl Hydrolyse (GH) family: β -1,4-endoglucanases, cellobiohydrolases and β -glucosidases. A search for conserved nucleotide sequences coding for each one of these enzymes was performed under the Arthropoda taxon in NCBI-GenBank database. Conserved sequences were used as query in a search for similar sequences into *A. grandis* cDNA database. Subsequent analyses, considering the enzyme potential for cellulose degradation, applicability to plant biomass fermentation and BlastX hits e-value ($< 1e-30$), pointed to β -1,4-endoglucanase as the best target for cloning procedures. β -1,4-endoglucanases are crucial to break glycoside linkages that joins glucose residues in cellulose polymers, the first step to enzymatic cellulose degradation for obtaining the fermentable product, glucose³⁶. However, to an efficient bioprocess, β -1,4-endoglucanases activity should be in synergism with other plant biomass degradation enzymes, encompassing an enzyme consortium for which substantial pH and heat stability are fundamental. The *A. grandis* β -1,4-endoglucanases gene cloned in this work was named AgraGH45-1 and its enzyme product was assayed for pH and temperature efficiency. The nucleotide sequence from A_grandis_454_rep_c946 contig (Accession number: GABY01019746.1), comprising a β -1,4-endoglucanase putative gene (AgraGH45-1), was first used to design primers for cloning the whole sequence of AgraGH45-1, including its 3' UTR by running a 3' RACE-PCR assay (Table 1). Further, we subcloned the entire mature enzyme-coding sequence, without the signal peptide, (named AgraGH45-1 Δ SP, which comprises 654 bp that are translated to 217 amino acids residues). Then, the amplicon AgraGH45-1 Δ SP was subcloned with restriction sites inserted on its 3' and 5' ends in order to allow its insertion into the

pGAPZ α -B vector in frame with the α -factor secretion signal and under control of the GAP constitutive promoter for heterologous expression in *P. pastoris* (Fig. 1). The subcloned expression vector was sequenced and confirmed as coding to a β -1,4-endoglucanase by BLAST based on GenBank-NCBI, confirming the sequence from *A. grandis* transcriptome.

Insect transcription profile of AgraGH45-1.

Initially, we have evaluated the transcript profile of AgraGH45-1 in *A. grandis* adult and larval tissues, including midgut and carcass. The carcass comprises all other insect tissues without the intestine. In both stages, AgraGH45-1 is far more expressed in the midgut than in carcass (Fig. 2A,B). These results, together with the presence of a predicted signal peptide, suggest that it can Gene Primer name Primer Sequence (5'-3')

AgraGH45-1	AgraGH45-1_qPCR_Fw	ATGTGACGAGCTCCCTACAGA	AgraGH45-1_qPCR_Rv	TTCAGTTGGGCAGGTAATTTG	GAPDH	AgraGAPDH_qPCR_F
						AgraGAPDH_qPCR_R
						AgraBtub_qPCR_F
						AgraBtub_qPCR_R

GCACCACCGAGTAAGTGTTT Table 1. Primers used in RT-qPCR experiments.

Figure 1. Schematic representation of the insertion of *A. grandis* β -1,4-endoglucanase gene (AgraGH45-1) into pGAPZ α -B expression vector. AgraGH45-1 was subcloned between EcoRI and Sall restriction sites in frame with the α -factor secretion signal and under control of the GAP constitutive promoter. Figure 2. Relative quantification of AgraGH45-1 in different *A. grandis* tissues. Quantification in larva (A) and adult (B) carcass and whole gut. Scientific Reports | (2019) 9:19580 |

<https://doi.org/10.1038/s41598-019-56070-1> 4 www.nature.com/scientificreports/ www.nature.com/scientificreports be a functional enzyme¹⁹. Indeed, coleopterans from families Curculionidae and Chrysomelidae have variable number of Glycosyl hydrolases in their genomes, including endoglucanases^{19,23,37} from families GH5, GH9, GH45 and GH48. However, functional validation in vivo is still necessary due to the special feature of non-functional enzymes may act as a decoy to compete with plant-derived enzyme inhibitors, as suggested by Kirsch and colleagues²⁰, or even new function acquisition being specific to a different substrate³⁸.

AgraGH45-1 amino acids alignment with other endoglucanases.

Endoglucanases are organized into 13 Glycosyl Hydrolases (GH) families based on primary and tertiary structures similarities, substrate specificities, as well as catalytic mechanisms³⁹. Such properties need to be appreciated when a recombinant enzyme is considered for industrial purposes. Plant biomass conversion process requires temperature and pH patterns in which both primary and tertiary enzyme structures need to keep its stability to improve enzyme activity. The predicted amino acid sequence of AgraGH45-1 shares a high identity level on the active site with other functional family 45 glycosyl hydrolases from beetles (Fig. 3), such as Rf-GH45, a *Rhynchophorus ferrugineus* β -1,4-Endoglucanase (74% whole amino acid identity) and Oa-EGase II, a β -1,4-Endoglucanase from *Oncideres albomarginata*, which structure was modeled by homology and suggests a polyphyletic origin for animal cellulases²⁷. The conservation of tertiary structure in a recombinant β -1,4-Endoglucanase is also fundamental for enzymatic efficiency. Gao and colleagues have shown that a loop movement is needed to the catalytic reaction in a *Tielavia terrestris* β -1,4-Endoglucanase⁴⁰, which shares the same catalytic mechanism than AgraGH45-1, where an aspartic acid (D), close to N-terminus (Fig. 3), serve as catalytic acid. The catalytic aspartic acid is conserved among AgraGH45-1, Rf-GH45 and Oa-EGase beetle β -1,4-Endoglucanases and shown in Fig. 3. Oa-EGase II shares 56% whole amino acid identity with AgraGH45-1. Besides all amino acids are conserved in the catalytic domains of the three beetle β -1,4-Endoglucanase, the sequences show slight differences between amino acids from glycosylation sites (Fig. 3). BLASTx results also reported a weak match against an *Aspergillus niger* Glycosyl Hydrolase (GH61), showing only 22% identity with AgraGH45-1.

AgraGH45-1 heterologous expression in *P. pastoris*.

Around 10 μ g of pGAPZ α -B/AgraGH45-1 were used to transform X-33 *P. pastoris* electrocompetent cells according to procedures described in Pichia Expression Kit (Termo Fisher Scientific, USA). *Pichia pastoris* system is the method of choice for heterologous expression of functional eukaryotic soluble proteins since this yeast is able to perform post-translational modifications, like N-glycosylation, required for protein functional folding^{41,42}. β -1,4-Endoglucanase is known to present glycosylation sites (Fig. 3), and such post-translational modification is crucial for its activity⁴³. Recently, several β -1,4-Endoglucanase enzymes have been successfully expressed in *P. pastoris*^{40,43,44}. The heterologous AgraGH45-1 was purified by affinity chromatography

and detected by Western blotting with Anti-6xHis antibody (Fig. 4). The recombinant enzyme was stable for five days under expression conditions (28 °C and 225 rpm). However, even though its predicted molecular weight is around 23 kDa, through our Western blotting detection, it appeared to be ~35 kDa (Fig. 4). This may be explained by mutation on the Kex2 cleavage site⁴⁵, which could prevent the signal peptide excision after protein secretion. The retention of the α -factor signal peptide could increase the molecular weight of the recombinant protein in 10 kDa. Hyperglycosylation is also commonly known to increase the molecular weight of recombinant proteins expressed in *P. pastoris*⁴⁶. In addition, according to the parameters for heterologous production of a recombinant enzyme, other properties, like optimal temperature and pH for enzymatic activity, can differ from the native enzyme. Chahed and colleagues produced a GH45 endoglucanase from *Sclerotinia sclerotiorum* that shows maximum activity at pH 7.0 and 60 °C, against pH 5.0 and 60 °C for the native enzyme⁴⁶. In another study, Akbarzadeh and colleagues presented a recombinant endoglucanase II from *Trichoderma reesei* expressed in *P. pastoris* with reduced disulfide bonds. The recombinant enzyme showed higher activity and thermostability than the native enzyme⁴⁷. Such variations could even better adjust the recombinant enzyme for industrial physicochemical conditions for plant biomass conversion. Figure 3. Alignment of the amino acid sequence of beetle β -1,4-endoglucanases. Signal peptide is underscored at N-terminus end. The conserved catalytic site, including catalytic aspartic acid (D), is shown in a solid box. Potential N-glycosylation sites are underlined below the alignment. Arrows indicate conserved cysteine residues. Scientific Reports | (2019) 9:19580 | <https://doi.org/10.1038/s41598-019-56070-1> 5 www.nature.com/scientificreports/ www.nature.com/scientificreports/

Evaluation of pH influence and substrate type on AgraGH45-1 activity.

Despite efforts for making cellulosic biofuels via enzymatic hydrolysis have begun since the 1990s, industry is still looking for new enzyme classes displaying superior hydrolysis kinetics and good stability for commercial purposes^{8,9}. The high cost of cellulase enzyme is one of the major barriers for commercialization of bioethanol production from lignocellulose biomass¹⁰. The pH and temperature are key parameters to improve enzymatic saccharification step in bioconversion processes⁴⁸. Moreover, the stability of enzymes involved in lignocellulosic degradation, when exposed to such parameters, can contribute to a cost-effective bioethanol production. The major part of

commercial cellulases suitable for bioethanol conversion has shown the best activity at pH ranging from 4 to 5 and at temperatures close to 50°C⁴⁹. More recently, two recombinant β -1,4-Endoglucanases were expressed in *P. pastoris* and evaluated for optimal pH, temperature and stability (AnCel5A – from *A. niger*, and TtCel45A – from thermophilic fungus *Thermotoga terrestris*)^{40,43}. Both enzymes showed optimal pH and temperature for its activity according to industrial requirements, as also achieved by AgraGH45-1. Nevertheless, AgraGH45-1 seems to keep its higher activity in a wider pH range. The enzymatic activity of AgraGH45-1 over Hydroxyethylcellulose (HEC) substrate was higher at pH 5.0 when analyzed in a range of pH from 3 to 9. Figure 5A shows that AgraGH45-1 efficiency over HEC is substantially stable at pH ranging from 5.0 to 7.0, dropping only 5.5%. On the other hand, the pH range for higher activity of AnCel5A and TtCel45A was from 4.0 to 5.0. AgraGH45-1 appears to keep its activity in a wider pH range, notably the range needed for bioethanol bioprocess. Such property could improve its applicability for plant biomass conversion. AgraGH45-1 capability to digest another substrate analogous to cellulose, Carboxymethylcellulose (CMC), was also evaluated in a radial diffusion assay. The supernatant of the culture inoculated with *P. pastoris* transformed with pGAPZ α -B/AgraGH45-1 was able to digest CMC more than 2-fold higher than the control at pH 5.0 (Fig. 5B). Previous reports also indicated β -1,4-endoglucanase activity in pH ranging from 4.5 to 5.5 for other insects, such as *Diabrotica virgifera virgifera*, *Apriona germari*^{24,50}, and even for the mollusc *Ampullaria crosseana*⁵¹.

Evaluation of temperature on AgraGH45-1 activity.

The efficiency of AgraGH45-1 at pH 5 and temperature ranging from 40 to 60 °C was evaluated and compared to *Aspergillus niger* commercial cellulase over Hydroxyethylcellulose substrate. *A. niger* is the most important source of commercial cellulases and has been used in several industrial processes for cellulose bioconversion. Its genome contains around 170 genes involved in polysaccharide degradation⁵². AgraGH45-1 maximum activity was found at 50 °C, and it was 1.3-fold higher than the one from *A. niger* commercial cellulase (Fig. 6). It was also observed that at 40 and 60 °C AgraGH45-1 activity falls 38% and 39%, respectively, compared to its maximum activity, whereas commercial *A. niger* cellulase dropped down 61% and 45%, respectively, compared to its maximum activity at 50 °C. Therefore, these results suggest that AgraGH45-1 is more efficient than *A. niger* commercial cellulase, and its activity

can slightly resist to temperature variation. Such AgraGH45-1 property could improve its applicability to an efficient bioprocess for biomass conversion. The activity of AgraGH45-1 over CMC and HEC demonstrates the capability of that recombinant enzyme to break glycoside bounds in such cellulose analogous substrate. However, it is known that the efficiency of some enzymes over CMC and HEC sometimes is not reproduced over complex biomass substrates⁵³. We have now Figure 4. SDS-PAGE (A) and Western Blotting (B) of AgraGH45-1 expression and purification. (1) Negative control (supernatant of a 5 days culture of *P. pastoris* transformed with empty pGAPZ α -B); (2) Supernatant of a 5 days culture of *P. pastoris* transformed with pGAPZ α -B_AgraGH45-1; (3) 4 μ g of purified AgraGH45-1. Gel and Blot were cropped to improve the clarity and conciseness of the presentation. Full-length blots/gels are presented in Supplementary Fig. 1. Scientific Reports | (2019) 9:19580 | <https://doi.org/10.1038/s41598-019-56070-1> 6 www.nature.com/scientificreports/ www.nature.com/scientificreports pointed our aims towards further assessing the ability of AgraGH45-1 to degrade complex biomass-derived substrates when added to an enzyme consortium. In order to improve biofuel production worldwide, advances in agricultural, industrial, and technological areas should contribute to making second generation ethanol production less expensive. Our study attempts to shed a light over such necessity of the renewable fuel industry.

Conclusion

In this study we performed the cloning of AgraGH45-1, an *A. grandis* midgut-specific β -1,4-endoglucanase, as well as assessed both heterologous expression and activity of that recombinant enzyme over substrates analogous to cellulose. Our results strongly indicate that AgraGH45-1 is more efficient than commercial *A. niger* for Hydroxyethylcellulose degradation at 50 °C. In addition, its activity presents smaller variation when assayed at 40 °C and 60 °C than *A. niger* commercial cellulase. AgraGH45-1 can potentially be tested as an enzyme consortium component for plant biomass digestion aiming second-generation ethanol production.

Methods

Gene selection.

A search for genes coding for plant biomass degradation enzymes from arthropods was performed at the NCBI (National Center for Biotechnology Information -

<http://www.ncbi.nlm.nih.gov>) GenBank. Conserved sequences were used as query in a search for similar sequences within *Anthonomus grandis* transcriptome30. An *A. grandis* transcriptome contig harbouring a β -1,4-endoglucanase coding sequence (A_grandis_454_rep_c954) was selected by similarity search using the BLASTx tool54, based on an e-value threshold of $1e^{-30}$.

Insect rearing.

Insects were obtained from a colony maintained at the Insect Breeding Laboratory at Embrapa Genetic Resources and Biotechnology in Brasília, DF, Brazil. Eggs, larvae, pupae and adults were kept under Figure 5. AgraGH45-1 enzymatic activity assessment. (A) AgraGH45-1 activity at pH ranging from 3.0 to 9.0. The enzymatic reaction was initiated by adding 50ml of OBR-HEC (5mg/ml) as substrate. Each point is the average of three measurements. Increase in Abs 550nm is proportional to substrate degradation. (B) AgraGH45-1 activity over carboxymethylcellulose in a radial diffusion assay. (I) Culture of *P. pastoris* transformed with empty vector (50 μ L) (negative control). (II) Culture of *P. pastoris* transformed with pGAPZ α -B/AgraGH45-1 (50 μ L). The white haloes on the dark blue background indicate that activity on B sample is more than 2 folds higher than control. Figure 6. Efficiency of AgraGH45-1 compared to commercial *A. Niger* cellulase. AgraGH45-1 was 1.3-fold more efficient than *A. niger* commercial cellulase to hydrolyze OBR-HEC 5mg/mL at 50 °C. When analyzed at 40 and 60 °C, the falling ratio of AgraGH45-1 activity was smaller than *A. niger* commercial cellulase. Scientific Reports | (2019) 9:19580 | <https://doi.org/10.1038/s41598-019-56070-1> 7 www.nature.com/scientificreports/ www.nature.com/scientificreports/ controlled temperature conditions ($26\pm 2^\circ\text{C}$), relative humidity ($70\pm 10\%$) and 12h:12h light:dark photo period. Insects were fed daily with artificial diet55.

A. grandis.

β -1,4-endoglucanase (AgraGH45-1) relative transcripts quantification by RT-qPCR in larva and adult tissues To verify the expression profile of AgraGH45-1, the *A. grandis* carcass and gut of both larvae and adult insects were dissected in 0.15M NaCl solution. Total RNA was isolated using TRIzol reagent (Termo Fisher Scientific, USA) according to the manufacturer's instructions. The integrity of RNA samples was confirmed by 1.0% agarose gel electrophoresis, and a Nanodrop spectrophotometer (Termo Fisher

Scientific, USA) was used to determine RNA quantity. For each sample, 2µg of total RNA treated with Ambion DNase I (RNase-free) (Termo Fisher Scientific, USA) was used to synthesize the first-strand cDNA using M-MLV Reverse Transcriptase Kit (Invitrogen, USA) according to the manufacturer's guidelines. OligoPerfect Designer (Termo Fisher Scientific, USA) was used to design primers for the reverse transcription quantitative PCR (RT-qPCR). AgraGH45-1 specific primers and those from the two reference genes (GAPDH and β-tubulin) used in RT-qPCR are displayed in Table 1. Each RT-qPCR reaction included 5.0µL of GoTaq qPCR Master Mix 2×(Promega, USA), 2.6µL of nuclease-free water, 0.2µL (0.2µM) of forward and reverse primers and 2.0µL of cDNA (diluted 1:20), for a total volume of 10µL. RT-qPCR assays were performed in a CFX96 thermocycler (Bio-Rad, USA) under the following conditions: incubation at 95°C for 20s, followed by 40 cycles at 95°C for 3s and 60°C for 30s. Reactions were set-up in 96-well Microseal PCR plates (Bio-Rad) in triplicate. At the end of each RT-qPCR experiment a dissociation curve for each amplicon was created to verify the possible formation of primer dimers or sample contamination. The efficiency of each primer for each reaction and the Cq value were calculated individually using the qPCR miner software (www.miner.ewindup.info) 56. Analyses of AgraGH45-1 expression were performed by using qBASE+software (Biogazelle, Belgium)57.

AgraGH45-1 cloning and insertion into expression vector.

Nucleotide sequence from the selected contig was used to design primers to clone, firstly the whole AgraGH45-1 ORF sequence. The AgraGH45-1 was amplified from total RNA extracts of *A. grandis* by 3' RACE-PCR and cloned in pGEM-T Easy Vector (Promega, USA). Then, a new primer set was designed to subclone AgraGH45-1 lacking the signal peptide correspondent sequence. The signal peptide sequence was identified using Signal-P tool58. Forward primer was designed to insert an EcoRI site (underlined) at the 5' end of AgraGH45-1 (AgraGH451GAPfwd: 5' TAAGAATTCCACTTAGCGGTTCTGGAACTACT 3'). Reverse primer was designed to insert a Sall site (underlined) at the 3' end of AgraGH45-1 (AgraGH451GAPrev: 5' TTAGTCGACCATATTGTCATTTACACAACC 3'). The amplified material, with new primer set, from pGEM-T-AgraGH45-1 clone was subcloned again in pGEM-T Easy, and named pGEM-T-AgraGH45-1ΔSP. This clone was digested with EcoRI and Sall, the insert was purified, and then inserted into pGAPZα-B expression vector previously digested with the same restriction enzymes, resulting in pGAPZα-B-AgraGH45-1ΔSP

clone. After all cloning procedures a de novo nucleotide sequencing was performed to confirm the correct insertion in frame with vector start codon.

Alignment of the amino acid sequence of AgraGH45-1 with others beetle cellulases.

Nucleotide sequence of AgraGH45-1 was translated into its amino acid sequence using Expasy Translate tool ([https:// web.expasy.org/translate/](https://web.expasy.org/translate/)) and aligned with Oa-EGase II (a β -1,4-endoglucanase from *O. albomarginata* [GenBank:GU001942]), and β -1,4-endoglucanase from *R. ferrugineus* (Rf- GH45) [GenBank: ANU06045.1]. Sequences alignment was performed with Clustal Omega tool for the multiple alignment of sequences⁵⁹ using default parameters and manually adjusted accordingly.

Heterologous expression of AgraGH45-1.

The expression vector harbouring the AgraGH45-1 (pGAPZ α -B-AgraGH45-1 Δ SP) was linearized with AvrII restriction enzyme and used to transform X-33 *P. pastoris* electrocompetent cells according to the Pichia Expression Kit (Termo Fisher Scientific, USA). Transformed cells were streaked in YPG (1% Yeats Extract, 2% Peptone and 2% Glucose) medium containing Zeocine 100 μ g/mL and kept at 28 °C for four days. Emerging colonies were screened by PCR using specific primers. PCR positive colonies were selected for small scale expression assay in liquid YPG/Zeocine medium (50mL of medium, 28 °C at 225 rpm for three days). A colony transformed with pGAPZ α -B empty vector was used as negative control during expression assays.

Evaluation of pH influence on AgraGH45-1 activity.

AgraGH45-1 activity over Hydroxyethylcellulose conjugated to Ostazin Brilliant-Red (OBR-HEC) chromophore⁶⁰ was evaluated on a range of pH. Therefore, 25 μ L of the AgraGH45-1 expression supernatant was incubated for 15min at 45°C with different buffers (50mM Sodium Citrate pH 3.0, 50mM Sodium Acetate pH 5.0, 50mM Sodium Phosphate pH 7.0 and 50mM Tris/HCl pH 9.0). The reaction was started by addition of 50 μ L of OBR-HEC 5mg/mL in a final volume of 150 μ L and kept for 30min at 45 °C. Reaction was stopped by adding 900 μ l of ethanol and then centrifuged at 10,000 \times g. The supernatant absorbance was read at 550nm. As a control for each reaction was used 25 μ L of the supernatant from an expression inoculated with *P. pastoris* transformed with an empty pGAPZ α -B vector. Each reaction was carried out in triplicate.

Radial diffusion assay.

Activity of the recombinant enzyme was detected as described by Jimenez and colleagues²⁴. Plates were prepared with agarose (1%), Congo red (0,02%) and Carboxymethylcellulose (CMC) (0,5%) dissolved in 50mM sodium acetate pH 5.0. After solidification into Petri plates, circular wells were punched into the agar with a 5mm diameter cork borer to a depth of 10mm. Activity assay was performed with 50 μ L of cell culture supernatant inoculated with *P. pastoris* transformed with pGAPZ α -B-AgraGH45-1 Δ SP and incubated at 45 °C overnight. After 24 hours the residual Congo red dye was removed by rinsing the plate with distilled water and then fixed by flooding the plate with acetic acid (10%) for 1 hour at room temperature. β -1,4-endoglucanase activity zones appeared as white haloes on a dark blue background. As a negative control it was used 50 μ L of cell culture supernatant inoculated with *P. pastoris* transformed with empty pGAPZ α -B.

Purification of AgraGH45-1.

A X-33 *P. pastoris* colony transformed with pGAPZ α -B-AgraGH45-1 Δ SP was inoculated in 100mL of YPG/Zeocine 100 μ g/mL and kept at 28 °C and 225 rpm for 3 days. After that, 10mL of the culture was used to inoculate 1 L of YPG/Zeocine 100 μ g/mL. The culture was kept at 28 °C and 225 rpm for 4 days. The cells were harvested by centrifugation, the supernatant filtered through 0.2 μ m and then diluted in sodium phosphate buffer pH 7.4 (1:1) to a final concentration of 20mM. Sodium chloride was added to a final concentration of 0.5M. The final volume was kept circulating through a HisTrap FF crude 5mL (G.E. Healthcare) column previously equilibrated with binding buffer (20mM sodium phosphate pH 7.4, 0.5M NaCl). The column was washed with 6 volumes of binding buffer containing 20 mM imidazole, and the recombinant AgraGH45-1 was eluted with binding buffer containing 0.5 M imidazole. Eluted fractions were dialyzed against 0.25 mM sodium bicarbonate and freeze-dried. Protein samples were quantified and used in electrophoretic assays to molecular weight determination. AgraH45-1 was detected by Western Blotting with 6 \times -His Tag Monoclonal Antibody (Termo Fisher Scientific, USA).

Comparison of AgraGH45-1 activity and thermo stability with *Aspergillus niger* commercial cellulase.

AgraGH45-1 and commercial *A. niger* cellulase (Sigma Aldrich, USA) activity on Hydroxyethylcellulose conjugated to Ostazin Brilliant-Red (OBR-HEC) was evaluated at pH 5.0 and temperature ranging from 40 to 60°C. Therefore, 10µg of both enzymes were dissolved in 25µL of 50mM sodium citrate pH 5.0 and incubated for 15min at 45 °C. The reaction was started by addition of 50 µL of OBR-HEC 5mg/mL in a final volume of 150 µL and kept for 30 min at 40 °C. The same procedure was repeated to incubation at 50 and 60 °C. Reactions were stopped by adding 900 µl of ethanol and then centrifuged at 10,000×g. The supernatant absorbance was read at 550nm. Each reaction was carried out in triplicate.

Received: 2 July 2019; Accepted: 20 November 2019; Published: xx xx xxxx

References

1. Franck, A., Trevisan, L., Oliveira, G., Ziani, F. & Coronel, D. Empirical analysis of Brazilian ethanol competition (1999–2016). *Rev. Adm. Neg. Amaz.* 10, (3) <https://doi.org/10.18361/2176-8366/rara.v10n3p53-72> (2018).
2. NovaCana; Senai starts project with Chinese state-owned company to produce ethanol from corn waste. Available at <https://www.novacana.com/n/etanol/2-geracao-celulose/senai-projeto-estatal-chinesa-produzir-etanol-residuos-milho-070819> (2019).
3. Soutinho, L., Rosario, F. & Lima, A. Mapping of technological routes of the Brazilian energetic suction sector technological change in ethanol production derived from sugar cane. *Rev. Econ. Pol. Desen.* 6(18), 80–89, <https://doi.org/10.28998/repd.v6i18.8660> (2014).
4. Marques, F. Bioenergia: Obstáculos no caminho. *Pesq. FAPESP.* 268, 58–63. Preprint at <https://revistapesquisa.fapesp.br/2018/06/18/obstaculos-no-caminho/> (2018).
5. Doblin, M. S., Johnson, K. L., Humphries, J., Newbiggin, E. J. & Bacic, A. Are designer plant cell walls a realistic aspiration or will the plasticity of the plant's metabolism win out? *Curr. Opin. Biotechnol.* 26, 108–114 (2014).
6. Souza, W. R. et al. Suppression of a single BAHD gene in *Setaria viridis* causes

- large, stable decreases in cell wall feruloylation and increases biomass digestibility. *New Phytol.* 218, 81–93, <https://doi.org/10.1111/nph.14970> (2018).
7. Johnson, E. Integrated enzyme production lowers the cost of cellulosic ethanol. *Biofuel Bioprod. Bior.* 10, 164–174, <https://doi.org/10.1002/bbb.1634> (2016).
 8. Brunecky, R. et al. High activity CAZyme cassette for improving biomass degradation in thermophiles. *Biotechnol Biofuels.* 11, 22, <https://doi.org/10.1186/s13068-018-1014-2> (2018).
 9. Jordan, D. B. et al. Plant cell walls to ethanol. *Biochem. J.* 442, 241–252, <https://doi.org/10.1042/BJ20111922> (2012).
 10. Liu, G., Zhang, J. & Bao, J. Cost evaluation of cellulase enzyme for industrial-scale cellulosic ethanol production based on rigorous Aspen Plus modelling. *Bioprocess Biosyst. Eng.* 39, 133–140, <https://doi.org/10.1007/s00449-015-1497-1> (2016).
 11. Olofsson, J., Barta, Z., Borjesson, P. & Wallberg, O. Integrating enzyme fermentation in lignocellulosic ethanol production: life-cycle assessment and techno-economic analysis. *Biotechnol. Biofuels.* 10, 51, <https://doi.org/10.1186/s13068-017-0733-0> (2017).
 12. Kuhad, R. C., Gupta, R. & Singh, A. Microbial cellulases and their industrial applications. *Enzyme Res.* 2, 280696, <https://doi.org/10.4061/2011/280696> (2011).
 13. Singhania, R. R., Sukumaran, R. K., Patel, A. K., Larroche, C. & Pandey, A. Advancement and comparative profiles in the production technologies using solid-state and submerged fermentation for microbial cellulases. *Enzyme Microb. Technol.* 46, 541–549 (2010).
 14. Srivastava, N. et al. Applications of fungal cellulases in biofuel production: Advances and limitations. *Renew. Sust. Energ. Rev.* 82, 2379–2386, <https://doi.org/10.1016/j.rser.2017.08.074> (2018).
 15. Glass, N. L., Schmoll, M., Cate, J. H. D. & Coradetti, S. Plant cell wall deconstruction by ascomycete fungi. *Ann. Rev. Microbiol.* 67(1), 477–98, <https://doi.org/10.1146/annurev-micro-092611-150044> (2013).
 16. Kumar, R., Singh, S. & Singh, O. V. Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives. *J. of Ind. Microbiol. & Biotech.* 35(5), 377–91, <https://doi.org/10.1007/s10295-008-0327-8> (2008).
 17. Hasunuma, T., Okazaki, F., Okai, N., Hara, K. Y. & Ishii, J. A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. *Bioresour. Technol.* 135((May)), 513–22,

<https://doi.org/10.1016/J.BIORTECH.2012.10.047> (2013).

18. Pauchet, Y. & Heckel, D. G. The genome of the mustard leaf beetle encodes two active xylanases originally acquired from bacteria through horizontal gene transfer. *Proc. R. Soc. London B Biol. Sci.* 280 (2013).

19. Pauchet, Y. et al. Studying the organization of genes encoding plant cell wall degrading enzymes in *Chrysomela tremula* provides insights into a leaf beetle genome. *Insect Mol. Biol.* 23(3), 286–300, <https://doi.org/10.1111/imb.12081> (2014).

20. Kirsch, R., Kunert, G., Vogel, H. & Pauchet, Y. Pectin digestion in herbivorous beetles: impact of pseudoenzymes exceeds that of their active counterparts. *Front. Physiol.* 10, 685 (2019).

21. Danchin, E.G.J. & Rosso, M.N. Lateral gene transfers have polished animal genomes: lessons from nematodes. *Front. Cell. and Infect. Microbiol.* 2 (27), <https://doi.org/10.3389/fcimb.2012.00027> (2012).

22. Wybouw, N., Pauchet, Y., Heckel, D. G. & Leeuwen, T. V. Horizontal gene transfer contributes to the evolution of arthropod herbivory. *Gen. Biol. and Evol.* 8(6), 1785–1801, <https://doi.org/10.1093/gbe/evw119> (2016).

23. Kirsch, R. et al. Horizontal gene transfer and functional diversification of plant cell wall degrading polygalacturonases: Key events in the evolution of herbivory in beetles. *Insect Biochem. Mol. Biol.* 52, 33–50 (2014). *Scientific Reports* | (2019) 9:19580 | <https://doi.org/10.1038/s41598-019-56070-1> 9 www.nature.com/scientificreports/ www.nature.com/scientificreports

24. Jiménez, A. V., Wang, H. & Siegfried, B. D. Expression and characterization of a recombinant endoglucanase from western corn rootworm, in *Pichia pastoris*. *J. Insect Sci.* 14, 242, <https://doi.org/10.1093/jisesa/ieu104> (2014).

25. Shirley, D. et al. Expression of an endoglucanase from *Tribolium castaneum* (TcEG1) in *Saccharomyces cerevisiae*. *Insect Sci.* 21, 609–618, <https://doi.org/10.1111/1744-7917.12069> (2014).

26. Celińska, E., Borkowska, M. & Białas, W. Evaluation of a recombinant insect-derived amylase performance in simultaneous saccharification and fermentation process with industrial yeasts. *Appl. Microbiol. Biotechnol.* 100, 2693–707 (2016).

27. Calderón-Cortés, N., Watanabe, H., Cano-Camacho, H., Zavala-Páramo, G. & Quesada, M. cDNA cloning, homology modelling and evolutionary insights into novel endogenous cellulases of the borer beetle *Oncideres albomarginata* (Cerambycidae). *Insect Mol Biol.* 19, 323–336, <https://doi.org/10.1111/j.1365->

2583.2010.00991.x (2010).

28. Kirsch, R., Heckel, D. G. & Pauchet, Y. How the rice weevil breaks down the pectin network: Enzymatic synergism and subfunctionalization. *Insect Biochem. Mol. Biol.* 71, 72–82 (2016).

29. Pauchet, Y., Wilkinson, P., Chauhan, R. & Ffrench-Constant, R. H. Diversity of beetle genes encoding novel plant cell wall degrading enzymes. *PLOS ONE*. 5. <https://doi.org/10.1371/journal.pone.0015635> (2010).

30. Firmino, A. A. P. et al. Transcriptome analysis in Cotton Boll Weevil (*Anthonomus grandis*) and RNA interference in insect pests. *PLOS ONE*. 8, <https://doi.org/10.1371/journal.pone.0085079> (2013).

31. Milanez, A. et al. From promise to reality: how cellulosic ethanol can revolutionize the sugarcane industry - an assessment of competitive potential and public policy suggestions. *Biofuels BNDES Sector* 41, 237–294 (2017).

32. Lopes, M. L. Ethanol production in Brazil: a bridge between science and industry. *Braz. J. Microbiol.* 47, 64–76, <https://doi.org/10.1016/j.bjm.2016.10.003> (2016).

33. Silva, G., Galeano, C. & Cavalcanti, I. Adaptation of *saccharomyces cerevisiae* metabolic model IMM904 for in silico determination of ethanol/xylitol production fows using xylulose. *Braz. J. Dev.* 5(9), 15236–15241 (2019).

34. Borelli, G. et al. Positive selection evidence in xylose-related genes suggests methylglyoxal reductase as a target for the improvement of yeasts' fermentation in industry. *Genome Biol. Evol.* 11(7), 1923–1938, <https://doi.org/10.1093/gbe/evz036> (2019).

35. Santos, C. A., Morais, M. A. B., Terrett, O. M. & Lyczakowski, J. J. An engineered GH1 β -glucosidase displays enhanced glucose tolerance and increased sugar release from lignocellulosic materials. *Sci. Rep.* 9(4903), 1–10, <https://doi.org/10.1038/s41598-019-41300-3> (2019).

36. Santa-Rosa, P. S. et al. Production of thermostable β -glucosidase and CMCase by *Penicillium* sp. LMI01 isolated from the Amazon region. *Electron. J. Biotechnol.* 31, 84–92, <https://doi.org/10.1016/j.ejbt.2017.11.005> (2018).

37. Eyun, S. et al. Molecular evolution of glycoside hydrolase genes in the Western Corn Rootworm (*Diabrotica virgifera virgifera*). *PLOS ONE*. 9, <https://doi.org/10.1371/journal.pone.0094052> (2014).

38. Busch, A., Danchin, E. G. J. & Pauchet, Y. Functional diversification of horizontally acquired glycoside hydrolase family 45 (GH45) proteins in Phytophaga beetles. *BMC*

Evol. Biol. 19, 100 (2019).

39. Cantarel, B. L. et al. The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 37, D233–D238 (2009).

40. Gao, J. et al. Characterization and crystal structure of a thermostable glycoside hydrolase family 45 1,4- β -endoglucanase from *Tielavia terrestris*. *Enzyme Microb. Technol.* 99, 32–37, <https://doi.org/10.1016/j.enzmictec.2017.01.005> (2017).

41. Bahadori, Z., Kalhor, H. R. & Mowla, S. J. Producing functional recombinant human keratinocyte growth factor in *Pichia pastoris* and investigating its protective role against irradiation. *Enzyme Microb. Technol.* 111, 12–20, <https://doi.org/10.1016/j.enzmictec.2018.01.003> (2018).

42. Cregg, J. M., Cereghino, L., Shi, J. & Higgins, D. R. Recombinant protein expression in *Pichia pastoris*. *Mol. Biotechnol.* 16, 23–52 (2000).

43. Yan, J. et al. Functional and structural analysis of *Pichia pastoris*-expressed *Aspergillus niger* 1,4- β -endoglucanase. *Biochem. Biophys. Res. Commun.* 475(1), 8–12, <https://doi.org/10.1016/j.bbrc.2016.05.012> (2016).

44. Liu, G. et al. Functional and structural analyses of a 1,4-endoglucanase from *Ganoderma lucidum*. *Enzyme Microb. Technol.* 86, 67–74, <https://doi.org/10.1016/j.enzmictec.2016.01.013> (2016).

45. Yang, S. et al. Enhanced production of recombinant secretory proteins in *Pichia pastoris* by optimizing Kex2 P1' site. *PLOS ONE.* 8 (2013).

46. Chahed, H., Boumaiza, M., Ezzine, A. & Marzouki, M. N. Heterologous expression and biochemical characterization of a novel thermostable *Sclerotinia sclerotiorum* GH45 endoglucanase in *Pichia pastoris*. *Int. J. Biol. Macromol.* 106, 629–635, <https://doi.org/10.1016/j.ijbiomac.2017.08.062> (2018).

47. Akbarzadeh, A. et al. Disulfide bonds elimination of endoglucanase II from *Trichoderma reesei* by site-directed mutagenesis to improve enzyme activity and thermal stability: An experimental and theoretical approach. *Int. J. Biol. Macromol.* 120, 1572–1580 (2018).

48. Saini, J. K., Saini, R. & Tewari, L. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *3 Biotech.* 5, 337–353, <https://doi.org/10.1007/s13205-014-0246-5> (2015).

49. Nieves, R. A., Ehrman, C. I., Adney, W. S., Elander, R. T. & Himmel, M. E. Technical Communication: survey and analysis of commercial cellulose preparations

suitable for biomass conversion to ethanol. *World J. Microbiol. Biotechnol.* 14, 301–304 (1998).

50. Lee, S. J. et al. cDNA cloning, expression, and enzyme activity of a cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comp. Biochem. Physiol.* 139, 107–116 (2004).

51. Guo, R., Ding, M., Zhang, S., Xu, G. & Zhao, F. Expression and characterization of two secreted His6-tagged endo-b-1,4-glucanases from the mollusc *Ampullaria crosseana* in *Pichia pastoris*. *Acta Biochim. Biophys. Sinica.* 40, 419–425 (2008).

52. Pel, H. J. et al. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* CBS 513.88. *Nat. Biotechnol.* 25, 221–231 (2007).

53. Su, L. J. et al. Evaluation of cellulolytic activity in insect digestive fluids. *Gen and Mol Res* 12, 2432–2441 (2013).

54. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402 (1997).

55. Monnerat, R. G. et al. Criação massal do bicudo do algodoeiro *Anthonomus grandis* em laboratório. *Embrapa Cenargen Technical Communication*, 46 (2000).

56. Zhao, S. & Fernald, R. D. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J. Comput. Biol.* 12, 1047–1064 (2005).

57. Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., & Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* 8 (2007).

58. Nielsen, H., Engelbrecht, J., Brunak, S. & Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10, 1–6 (1997).

59. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539, <https://doi.org/10.1038/msb.2011.75> (2011).

60. Biely, P., Mislovicova, D. & Toman, R. Soluble chromogenic substrates for the assay of endo-1,4-beta-xylanases and endo-1,4-beta-glucanases. *Anal. Biochem.* 144, 142–146 (1985). *Scientific Reports* | (2019) 9:19580 | <https://doi.org/10.1038/s41598-019-56070-1>

www.nature.com/scientificreports/ www.nature.com/scientificreports

Acknowledgements

We thank Dr. Elton Jose Rosas de Vasconcelos (Leeds Omics, University of Leeds,

UK) for reading and commenting on the manuscript, and Dr. Lidia Maria Pepe Moraes for contribute to the *Pichia pastoris* expression experiments. We also thanks to CNPq and CAPES for funding this research.

Author contributions

All the authors contributed to the final manuscript as described. Liz Nathalia Ibarra Duarte (PhD student): Perform experiments involving cloning, expression, purification and analysis of the enzymatic activity. Ana Elizabeth Oliveira de Araújo Alves (PhD student): Perform optimization of gene expression in *Pichia pastoris*. José Dijair Antonino (Researcher): Supervision and execution of experiments involving gene cloning. Contribution as manuscript writing advisor. Guilherme Souza Prado (Post-doc): Perform protein quantification, SDS-PAGE and western blot. Contribution as manuscript writing advisor. Clidia Eduarda Moreira Pinto (PhD student): Perform qPCR assays. Carlos Ricardo Soccol (Researcher): Financial support. Supervision of experiments and manuscript writing. Érico Augusto Rosas de Vasconcelos (Researcher): Bench experiments advisor to gene cloning, heterologous expression, heterologous protein purification and enzymatic activity assays. Manuscript writing and edition. Corresponding author. Maria Fátima Grossi-de-Sá (Researcher/Team leader): Financial support. Supervision of experiments and manuscript writing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-56070-1>. Correspondence and requests for materials should be addressed to É.V.R.d.V. Reprints and permissions information is available at www.nature.com/reprints. Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain

permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>. © Te Author(s) 2019