

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

Maria Augusta Sabadine

**EXPRESSÃO HETERÓLOGA EM
SACCHAROMYCES CEREVISIAE E
ENSAIO DE ATIVIDADE DE UMA β -1,4-
ENDOGLUCANASE ISOLADA DO
CUPIM VELOCITERMES
HETEROPTERUS (ISOPTERA,
TERMITIDAE)**

São Carlos
2013

MARIA AUGUSTA SABADINE

**Expressão Heteróloga em Saccharomyces
Cerevisiae e Ensaio de Atividade de uma β -
1,4- Endoglucanase Isolada do Cupim
Velocitermes heteropterus (Isoptera,
Termitidae)**

Dissertação de mestrado apresentado como requisito para a obtenção
do título Mestre em Genética Evolutiva e Biologia Molecular
pelo Programa de Pós-Graduação em Genética
Molecular e Biologia Evolutiva através
Universidade Federal de São Carlos.

São Carlos
2013

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

S113eh Sabadine, Maria Augusta.
Expressão heteróloga em *Saccharomyces cerevisiae* e ensaio de atividade de uma β -1,4- endoglucanase isolada do cupim *Velocitermes heteropterus* (Isoptera, Termitidae) / Maria Augusta Sabadine. -- São Carlos : UFSCar, 2014.
139 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2013.

1. Genética. 2. *Velocitermes heteropterus*. 3. Celulase. 4. β -1,4-endoglucanase. 5. Cupim. 6. Bioetanol. I. Título.

CDD: 575.1 (20^a)

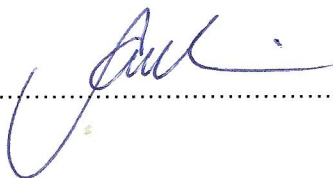
UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA EVOLUTIVA E BIOLOGIA MOLECULAR

Dissertação de Mestrado de

MARIA AUGUSTA SABADINE

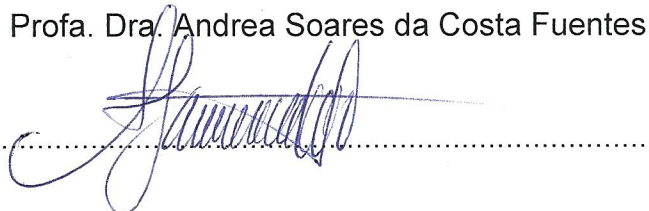
Banca Examinadora

Prof. Dr. Anderson Ferreira da Cunha



.....

Profa. Dra. Andrea Soares da Costa Fuentes



.....

Profa. Dra. Ana Maria Costa Leonardo



.....

Dedico este trabalho a todos aqueles
que sabem as dificuldades que tive,
principalmente a meus pais e amigos mais íntimos.

Agradecimentos

Agradeço à Deus, primeiramente.

Agradeço aos meus pais, por tanto carinho e amor. E por me apoiarem nessa jornada.

Agradeço ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular.

Agradeço a CNPq e a Capes.

Aos membros da banca.

Agradeço meus (minhas) amigos (as) que estiveram comigo durante toda a graduação e pós-graduação que me proporcionaram momentos inesquecíveis, ajudando-me sempre a seguir em frente e a nunca desistir dos meus objetivos e sonhos; por mais duro que tenham sido alguns obstáculos. Obrigada pela amizade e por estarem sempre comigo.

Aos meus amigos de trabalho, pela confiança, colaboração, companheirismo e amizade; que, de maneira direta e indireta, me ajudaram a superar obstáculos e permanecer em minhas metas. Principalmente a Karen, Karina, Carla e Krissia; e ao Breyer pela enorme ajuda.

Aos amigos do Departamento de Genética e Evolução.

Aos professores da graduação e pós-graduação pelos ensinamentos.

À Prof^ª. Dr^ª. Andrea Soares da Costa Fuentes pelo tempo disponível, por me ouvir e ajudar sempre que necessário.

À Prof^ª. Dr^ª. Ana Maria Costa-Leonardo por ceder gentilmente os cupins (objeto de estudo deste trabalho) e por compartilhar conhecimentos da área.

Ao Prof. Dr. Flávio Henrique Silva, pelo tempo disponível, conselhos e ajuda prestada.

Ao Prof. Dr. Iran Malavazi, pela colaboração, confiança, ajuda na solução de problemas e companheirismo.

Ao Leandro Seiji, porque sem ele boa parte deste estudo não teria ocorrido. Agradeço a enorme ajuda prestada e conselhos e longas conversas de onde saíram inúmeras ideias.

Ao Prof. Dr. Reinaldo Alves de Brito e a Prof^ª. Dr^ª. Maria Tereza Marques Novo pela colaboração.

Ao Prof. Dr. Francis Nunes, pela ajuda.

Ao Prof. Dr. Gonçalo Amarante Guimarães Pereira e ao Laboratório de Genômica e Expressão (LGE- Unicamp, sp), ao Marcelo e Melline pela colaboração quanto ao transcriptoma.

Ao Prof. Dr. Luiz Lehmann Coutinho, pela construção da biblioteca RNASeq (ESALQ-USP de Piracicaba, Departamento de Zootecnia).

Ao pós-doc Gustavo Gasparin e Sonia C. S. Andrade, pela construção da biblioteca RNASeq.

Ao laboratório LBM (Laboratório de Biologia Molecular) pela ajuda e auxílio, principalmente a Taise e Delita.

Ao Laboratório de Genética de Populações e Evolução, pela ajuda; principalmente ao André Luís Andrade Lima.

Ao Laboratório LBBMM (Laboratório de Bioquímica e Biologia Molecular de Microrganismos) pela ajuda quando eu precisei. Pelos equipamentos, kits, reagentes emprestados, dentre outros.

Aos presentes, pelo tempo disponível, por ouvirem-me.

Ao Prof. Dr. Anderson Ferreira da Cunha pela orientação.

The struggles I'm facing
The chances I'm taking
Sometimes might knock me down
But no, I'm not breaking
I may not know it
But these are the moment that I'm gonna remember most and
Just gotta keep going
And I, I got to be strong
Just keep pushing on, cause

There's always gonna be another mountain
I'm always gonna want to make it move
Always gonna be an uphill battle
Sometimes you're going to have to lose
Ain't about how fast I get there
Ain't about what's waiting on the other side
It's the climb

(The Climb- Miley Cyrus; compositores Jessi Alexander e Jon Mabe)

Resumo

A visão atual de que as mudanças ambientais que o planeta vem sofrendo se deve, em grande parte, ao uso de combustíveis fósseis. Esse panorama, lançou as necessidades para a substituição de combustíveis fósseis para combustíveis renováveis. Neste contexto, as novas tecnologias estão sendo desenvolvidas para a produção de novas fontes de bioenergia. A biomassa lignocelulósica representa um dos recursos renováveis mais abundantes na terra, e certamente um dos mais baratos, motivo pelo qual a mesma é extensivamente estudada como fonte de energia sustentável para uso industrial. Diversos métodos para obtenção de bioetanol lignocelulósico estão em experimentação e é imprescindível que o Brasil mantenha uma liderança natural neste campo de pesquisa. Assim, o uso de celulases presentes em organismos eucariotos vem tendo destaque no repertório comercial de enzimas com finalidade de degradar materiais lignocelulósicos em açúcares redutores (glicose) como alvo para o processo fermentativo. Os cupins, insetos eusociais, são organismos potenciais para tal investimento; pois se alimentam de materiais lignocelulósicos e possuem as enzimas necessárias para a degradação da celulose e obtenção de glicose utilizada no processo industrial de produção de etanol de segunda geração. Dessa forma, identificar, isolar e expressar celulases presentes neste tipo de inseto é um objetivo essencial para a busca de uma enzima compatível com o processo de produção de etanol implementado no Brasil atualmente. A espécie Velocitermes heteropterus possivelmente pode expressar uma celulase de interesse, presumivelmente ativas nos tanques de fermentação. Nesse estudo foi possível a identificação, sequenciamento e clonagem de uma das 3 principais enzimas do processo de degradação da celulose, a enzima β -1,4-endoglucanase de Velocitermes heteropterus. Experimentos iniciais demonstraram uma forte expressão heteróloga desta enzima na levedura Saccharomyces cerevisiae (linhagem FGSC9721). Testes visando purificação, atividade enzimática e biológica desta proteína, encontram-se em andamento.

Palavra-chaves: Velocitermes heteropterus. Celulase. β -1,4-endoglucanase. Cupim. Bioetanol.

Abstract

The current world scenery points out that the environment changes in the Earth planet is due to the intense use of fossil fuels along decades. This panorama launched the needs for the replacement of fossil fuels to sustained renewable fuels. In this context new technologies are being developed for the production of new bioenergy sources. The lignocellulosic biomass represents an abundant renewable source of cellulose on Earth. This is certainly one of the cheapest materials and therefore has been extensively studied as source of sustainable energy for industrial purposes. Several methods for obtaining bioethanol from lignocellulosic biomass are being sought worldwide and in that regard it is essential that the Brazil could keep its natural leadership in this field of research. In this way, the use of cellulases from eukaryotic organisms has achieved high importance in the commercially-available enzymes repertoire with purpose to degrade lignocellulosic materials in reduction sugars (glucose) for the fermentation process. The termites, eusocial insects, are potential organisms for this research investment, because they feed lignocellulosic materials and their digestory system presents the enzymes for cellulose hydrolysis generating glucose for the industrial process for production second generation ethanol. Thus, the identification, isolation and expression of termites cellulases can be considered are a key-step to search for enzymes suitable for the use of cellulose-derived glucose in current process of ethanol production implemented in Brazil. Velocitermes heteropterus can possibly express a cellulase of interest presumably active in the fermentation tanks. In this study, we identified, sequenced and cloned one of three major enzymes of the cellulose degradation system, the endo-1,4- β -glucanase from Velocitermes heteropterus. Our results showed a high expression of its enzyme in the yeast Saccharomyces cerevisiae (FGSC9721 strain). Experiments for purification and enzymatic activity of this enzyme are in development.

Key-Words: Velocitermes heteropterus. Cellulase. β -1,4-endoglucanase. Termite. Bioethanol.

Lista de Figuras

Figura 1: Diagrama de fluxo da produção de açúcar e bioetanol de cana.	20
Figura 2: Palha ou Palhada.	28
Figura 3: Composição da lignocelulose.	30
Figura 4: Estrutura do Fragmento de Celulose.	32
Figura 5: Estrutura química da celulose.	32
Figura 6: Microfibrila de Celulose.	33
Figura 7: Estrutura química polimérica de hemicelulose e locais de hidrólise enzimática envolvidos na degradação do polímero de hemicelulose.	34
Figura 8: Componentes da fração hemicelulose.	35
Figura 9: Estrutura geral da Lignina.	37
Figura 10: Precursores primários da lignina.	37
Figura 11: Estrutura da parede celular vegetal.	39
Figura 12: Fibrilas.	39
Figura 13: Esquema do efeito do pré-tratamento em materiais lignocelulósicos.	43
Figura 14: Visão esquemática da biodegradação da celulose.	46
Figura 15: Sinergismo entre celulases.	47
Figura 16: Diferenciação morfológica das castas em cupins – Referência : Nasutitermitinae.	53
Figura 17: <i>Velocitermes heteropterus</i> - Termitidae: Nasutitermitinae.	58
Figura 18: Integridade do RNA total extraído de Operários de <i>V. heteropterus</i> sp.	75
Figura 19: Verificação da síntese do DNA complementar com o uso de primer β -actina.	76
Figura 20: Amplificação do cDNA de β -1,4-endoglucanase de operários de <i>V. heteropterus</i>	77
Figura 21: Esquematização da Construção do pGemT-Easy+EG.	78
Figura 22: PCR da colônia positiva para inserção do plasmídeo pEG.	79
Figura 23: Análise da qualidade de extração plasmidial da colônia positiva.	80
Figura 24: Extração plasmidial do pYADE4.	80
Figura 25: Sequência Consenso da β -1,4- endoglucanase de <i>V. heteropterus</i> encontrada a partir do alinhamento de trechos sequenciados a partir da construção pEG.	81
Figura 26: Sequência consenso comparada contra bando de dados utilizando a ferramenta BlastN.	82
Figura 27: Sequência consenso alinhada com a β -1,4-endoglucanase de <i>Nasutitermes takasagoensis</i>	82
Figura 28: Esquema do alinhamento entre a sequência consenso de β -1,4-endoglucanase de <i>V. heteropterus</i> sp com sequência de <i>Nasutitermes takasagoensis</i>	83
Figura 29: Sequência consenso obtida pelo sequenciamento do plasmídeo pEG comparada contra Sequência do transcriptoma de <i>V. heteropterus</i>	84
Figura 30: Comparação da sequência consenso identificada no sequenciamento e transcriptoma com sequências das espécies de cupins <i>Nasutitermes walkeri</i> , <i>Coptotermes</i>	

formosanus, <i>Coptotermes acinaciformis</i> , <i>Sinocapritermes mushae</i> , <i>Reticulitermes speratus</i> , <i>Odontotermes formosanus</i> , e <i>Reticulitermes flavipes</i>).....	86
Figura 31: Alinhamento dos sítios conservados encontradas nas diferentes espécies de cupins (<i>Nasutitermes walkeri</i> , <i>Coptotermes formosanus</i> , <i>Coptotermes acinaciformis</i> , <i>Sinocapritermes mushae</i> , <i>Reticulitermes speratus</i> , <i>Odontotermes formosanus</i> , e <i>Reticulitermes flavipes</i> e <i>V. heteropterus</i> sp).....	86
Figura 32: Análise filogenética utilizando o Método Neighbor-Joining.	87
Figura 33: Análise filogenética utilizando o Método de Máxima Verossimilhança.	89
Figura 34: Esquematização da clivagem do pGemT-Easy+EG e vetor de expressão pYADE4.	90
Figura 35: Digestão do pGemT-Easy+EG usando BamHI e XhoI.	91
Figura 36: Purificação do pYADE4 e do inserto (iEG) após a digestão.	91
Figura 37: Esquematização da ligação entre pYADE4 e o iEG.....	92
Figura 38: PCR de Colônia para as colônias selecionadas de bactérias para a verificação da construção do plasmídeo yEG (pYADE4+EG).	93
Figura 39: Midipreparação plasmidial do yEG.	94
Figura 40: PCR para a verificação da construção do plasmídeo yEG.	94
Figura 41: Análise de expressão de β -1,4-endoglucanase em <i>S. cerevisiae</i>	95
Figura 42: SDS-PAGE da linhagem recombinante e controle do lisado e do meio.....	96
Figura 43: Teste CMC com linhagens transformantes selecionadas.	97
Figura 44: Comparação da análise Fermentativa da Linhagem Selvagem PE-2 com a linhagem recombinante.....	99
Figura 45: Sequência Consenso da β -1,4- endoglucanase de <i>V. heteropterus</i> encontrada a partir do alinhamento de trechos sequenciados a partir da construção yEG.	101
Figura 46: Sequência consenso de aminoácidos alinhada com a β -1,4-endoglucanase de <i>Nasutitermes takasagoensis</i>	103
Figura 47: Predição da estrutura terciária para β -1,4-endoglucanase.....	104
Figura 48: Macromolécula de β -1,4-endoglucanase (template d1Ks8a).....	105
Figura 49: Diferenças entre estrutura secundária e terciária para β -1,4-endoglucanase para <i>V. heteropterus</i> , <i>N. takasagoensis</i> e estrutura 1Ks8.	106
Figura 52: Ligação de adaptadores na Construção de sequência sem PCR.....	128
Figura 53: Geração de <i>Clusters</i>	129
Figura 54: Sequenciamento.	131
Figura 55: Sequência da β -1,4-endoglucanase de <i>V. heteropterus</i> a partir de seu transcriptoma.	133
Figura 56: Estrutura Secundária de β -1,4-endoglucanase de <i>N. takasagoensis</i>	135
Figura 57: Alinhamento da β -1,4-endoglucanase de <i>Nasutitermes takasagoensis</i> com <i>template</i> d1Ks8a.	136
Figura 58: Estrutura Secundária de β -1,4-endoglucanase de <i>V. heteropterus</i>	138
Figura 59: Alinhamento da β -1,4-endoglucanase de <i>V. heteropterus</i> com <i>template</i> d1Ks8a.	139

Lista de Tabelas

<i>Tabela 1: Comparação das diferentes matérias-primas para a produção de etanol e bioetanol</i>	22
<i>Tabela 2: Composição média do bagaço e palha da cana-de-açúcar</i>	26
<i>Tabela 3: Composição química parcial de alguns materiais lignocelulósicos</i>	30
<i>Tabela 4: Sequências de primers utilizados e sua finalidade</i>	63

Abreviações e siglas

ADH2: Álcool desidrogenase 2
Amp^r: Resistência a Ampicilina
BG: Beta-glucanase
bp: pares de base
CBH: Celobiohidrolase
cDNA: DNA complementar
CMC: Carboximetilcelulose
DNA: ácido desoxirribonucleico
EDTA: ácido etilenodiamino tetra-acético
EG: Endoglucanase
EGTA: ácido tetracético de etileno-glicol
°C: graus Celsius
g: grama
HCl: Ácido clorídrico
IPTG: β-D-1-tiogalactopiranosídeo isopropílico
kDa: quiloDaltons
LB: Luria Bertani
Leu: Leucina
mg: micrograma
M: Molar
μL: microlitro
μM: micromolar
mL: mililitro
mM: micromolar
MOPS: ácido 3-(N- morfolino)-propanossulfônico
mRNA: ácido ribonucleico mensageiro
NaCl: Cloreto de Sódio
NCBI: *Nacional Center for Biotechnology Information*
OD: Densidade Óptica
PCR: Reação de polimerase em Cadeia
PDB: *Protein Data Base* (Banco de dados proteicos)
pmSF: fluoreto de fenilmetilsulfonilo
RNA: ácido ribonucleico
rRNA: ácido ribonucleico ribossômico
rpm: rotação por minuto
SDS: Dodecil sulfato de Sódio
TAE: Tris-Acetato-EDTA
TE: Tris-EDTA
Tris: trisaminometano
tRNA: ácido ribonucleico transportador
Trp: triptofano
U: Unidade
Ura: Uracila
YNB: Nitrogênio Levedura Base
YPD: Extrato de Levedura Peptona Dextrose

Sumário

Resumo	7
Abstract	8
1 Introdução	16
2 Revisão Bibliográfica	18
2.1 Recursos energéticos e produção de etanol de primeira geração	18
2.2 Biomassa e sua relação energética	21
2.2 Bioetanol (Etanol de Segunda Geração)	21
2.3 Bagaço de cana-de-açúcar	24
2.4 Palha da cana-de-açúcar	25
2.5 Materiais Lignocelulósicos: entendendo sua estrutura	28
2.5.1 Celulose	31
2.5.2 Hemicelulose	34
2.5.3 Lignina	36
2.6 Produção de Etanol a partir de Materiais Lignocelulósicos	40
2.7 Pré-tratamento de Materiais Lignocelulósicos	42
2.8 Hidrolise Enzimática da Celulose	45
2.9 Celulases Animais	49
2.10 Cupins e celulases endógenas	52
3 Objetivo	58
3.1 Objetivo Geral	58
3.2. Objetivos Específicos	58
4 Material e Métodos	59
4.1 Material biológico	59
4.1.1 Cupins	59
4.1.2 Linhagens para clonagem e expressão	60
4.2 Meios de cultura	60
4.3 Soluções	61
4.4 Plasmídeos	62

4.5 Primers	63
4.6 Extração de RNA total.....	63
4.7 Síntese de DNA complementar (cDNA).....	64
4.8 Verificação da síntese de DNA complementar.....	65
4.9 Construção dos primers para amplificação do fragmento que codifica a β -1,4-endoglucanase	65
4.10 Amplificação do cDNA que codifica a β -1,4-endoglucanase de <i>V. heteropterus</i>	66
4.11 Clonagem do fragmento que codifica β -1,4-endoglucanase de <i>V. heteropterus</i> em vetor pGemT-Easy	67
4.12 Minipreparação de plasmídeo pGemT-Easy+EG e pYADE4+EG	67
4.13 Sequenciamento dos clone recombinantes.....	68
4.14 Bioinformática e análise dos dados	68
4.15 Expressão heteróloga da β -1,4-endoglucanase em levedura <i>S. cerevisiae</i> FGSC 9721	69
4.15.1 Midipreparação de plasmídeos pGemT-Easy+EG e pYADE4+EG	70
4.15.2 Transformação em linhagem de levedura <i>S. cerevisiae</i> PGSC 9721	71
4.16 Produção de proteína recombinante	72
4.17 Teste de atividade em meio CMC.....	73
4.18 Teste comparativo e nível proteico.....	73
4.19 Extração de DNA plasmidial de levedura para verificação da positividade da transformação.....	74
5 Resultados e Discussões.....	75
5.1 Extração de RNA.....	75
5.2 Síntese de cDNA e Verificação da eficiência da transcrição	76
5.3 Amplificação do fragmento que amplifica a celulase β -1,4-endoglucanase	76
5.4 Clonagem de β -1,4-endoglucanase de <i>V. heteropterus</i> no plasmídeo pGemT-Easy.....	77
5.5 PCR de colônia para verificação de clones positivos e Minipreparação de plasmídeos ..	79
5.6 Sequenciamento dos clones	80
5.7 Filogenética da β -1,4-endoglucanase	87
5.8 Construção do plasmídeo recombinante β -1,4-endoglucanase e Expressão heteróloga da em levedura <i>S. cerevisiae</i> FGSC 9721	89
5.9 Expressão da proteína recombinante β -1,4-endoglucanase	95
5.10 Teste de atividade em meio CMC.....	96
5.11 Testes de crescimento e viabilidade celular e nível proteico	98
5.12 Sequenciamento do pYADE4+EG.....	100
5.14 Predição da estrutura terciária.....	101
Considerações Finais.....	107
Referências	110

APÊNDICE	125
ANEXO I	134
ANEXO II	137

1 Introdução

A escassez ou ausência de combustíveis fósseis, além das alterações climáticas geradas devido a queima destes combustíveis, vêm estimulando pesquisas relacionadas a combustíveis de cunho renovável e que sejam ambientalmente favoráveis¹⁻⁴. Atualmente, o álcool se insere como principal substituto do petróleo; de modo que, sua produção é crescente⁵, principalmente nos Estados Unidos e Brasil. Todavia, a demanda deste biocombustível gera expansão desmedida dos plantios de milho e cana-de-açúcar sobre as vegetações nativas e florestas, causando danos ambientais e conflitos com a produção de subsistência (alimentar).

Somente no Brasil, cerca de 30 bilhões de etanol em m³ são produzidos anualmente a partir da cana-de-açúcar, perdendo somente para os Estados Unidos; de modo que esta quantidade tende a aumentar nos próximos anos, (podendo chegar à 65 bilhões de m³ em 2020^{6,7}, devido a demanda nacional e internacional do combustível^{1,2,8,9}. E, assim sendo, para se evitar a expansão desmesurada das áreas cultivadas, tem-se desenvolvido processos biotecnológicos^{2,3,10} a fim de aproveitar a biomassa residual de composição lignocelulósica, como bagaço e palha da cana, geradas durante o processo de produção alcoólica nas usinas^{11,12}.

Tais resíduos, como citado por Leite e Leal (2007)¹³ mostram-se de suma importância na utilização para a produção de bioetanol de segunda geração; haja visto que, devido ao fim das queimadas, redução do consumo energético pelas usinas e início do recolhimento da palha, a produção de etanol a partir da cana processada poderá crescer em até 50% sobre os valores de produção atuais.

O potencial para o uso do bagaço e da palha de cana para produzir o chamado etanol de segunda geração é enorme, especialmente por causa da grande disponibilidade

desta matéria-prima, cuja qual vem crescendo ano a ano no Brasil, principalmente devido a ascendente produção de açúcar e etanol tanto para consumo interno como para exportação; além de servir como alternativa fundamental na produção de combustíveis renováveis sem o prejuízo da produção de alimentos ^{11,14,15}.

É, portanto, neste contexto, que a busca por processos biotecnológicos que visem a transformação da celulose em glicose faz-se necessário ^{16,17}. Isto porque essa etapa é fundamental para a geração do biocombustível ao fazer uso desta matéria-prima ^{4,18}. A partir daí, segue-se o processo de fermentação desses açúcares até a formação do etanol.

O processo de quebra da celulose é realizado de forma enzimáticas, envolvendo 3 tipos diferentes de celulasas (β -1,4-endoglucanase, exoglucanase e β -glucosidase). O processo de conversão do açúcar extraído da celulose em etanol é realizado por leveduras, da mesma forma do processo atual de fermentação que é inteiramente dominado pelas usinas ⁶.

O problema não é produzir o etanol celulósico, levando em conta a enorme gama de enzimas celulolíticas presentes no mercado internacional; mas refere-se ao custo competitivo com o do petróleo e o do etanol produzido hoje de maneira tradicional, a partir da fermentação do açúcar presente no caldo de cana ^{17,19}. Além disso, as enzimas existentes atualmente não são estáveis em condições de fermentação atuais; fazendo com que sua sinergia com o processo de fermentação, o qual se denomina Sacarificação e Fermentação Simultânea (*Simultaneous Saccharification and Fermentation – SSF*), seja inviável ²⁰. A hidrólise e Fermentação em separado (*Separated Hydrolysis and Fermentation – SHF*) é a concepção mais antiga, na qual a hidrólise da celulose ocorre em duas etapas: a primeira ocorre em altas temperaturas, após o pré-tratamento da

biomassa, hidrolisando e solubilizando a celulose em glicose, a segunda é o processo fermentativo em si. Tais etapas ocorrem separadamente²¹.

Dessa maneira, identificar celulasas que, além de possuírem menor custo¹², sejam eficazes também durante o processo fermentativo, de modo a serem competentes sob condição de pH e temperatura ótima²² referentes a levedura quando em fermentação nas dornas tem sido primordial entre as pesquisas. Isso porque, visa-se a adição destas enzimas às dornas juntamente com as leveduras *S. cerevisiae* que converteriam a glicose gerada da quebra da celulose de palha e bagaço em etanol^{4,8,18}; da mesma forma que convertem glicose e frutose proveniente da quebra da sacarose presente no melaço ou caldo de cana; eliminando etapas intermediárias e separação de processos⁸. Isso resultaria em menor custo e tempo, aumentando a produção e afetando positivamente setores agrários, econômicos, energéticos e ambientais.

2 Revisão Bibliográfica

Neste capítulo são apresentados os diferentes aspectos teóricos que tangem a temática abordada: a relação das celulasas advindas de cupins na produção de etanol de segunda geração através da quebra da celulose de palha e bagaço de cana-de-açúcar. Para isso, necessitamos entender os seguintes tópicos:

2.1 Recursos energéticos e produção de etanol de primeira geração

Segundo Masiero e Lopes (2008)²³, os recursos energéticos vêm tendo destaque global, já que é notório que o suprimento energético contínuo propicia um maior crescimento econômico e melhoria na qualidade de vida; uma vez que, a energia

encontra-se claramente relacionada à segurança nacional e a estabilidade econômica global^{3,4,8}. Assim, devido a crescente demanda por energia, pesquisas são realizadas no tocante ao desenvolvimento de tecnologias e sistemas eficazes, além da diversificação de fonte de suprimento, visando especialmente, energias limpas e renováveis^{2,3,9}. Fontes energéticas de carvão, petróleo e gás são, nesse sentido, gradativamente somadas e/ou substituídas com/por outras, como nuclear, eólica, elétrica e dos biocombustíveis³. O último com ênfase mundial na área econômica, sobretudo como alternativa ao uso do petróleo; nas formas de bioetanol e biodiesel (bioenergia) – este a partir de algodão, babaçu, resíduos, palma, canola, girassol, nabo forrageiro, mamona, soja e gordura animal^{13,23} e aquele, de milho e cana-de-açúcar, principalmente.

Neste tocante, o etanol produzido a partir da cana-de-açúcar, vem se mostrando como uma das alternativas mais eficientes de substituição da matriz energética. Para sua produção, a cana-de-açúcar passa por um processamento (moagem) nas usinas²⁴; formando o bagaço (fibra) que será utilizado para a produção de energia através de queima nas caldeiras e o caldo contendo os açúcares da cana, que será, então, destinado à produção de açúcar ou bioetanol. (Figura 1)

O caldo, por sua vez, é inicialmente peneirado e tratado quimicamente, para coagulação, floculação e precipitação de impurezas, que são eliminadas por decantação. Em usinas sucroalcooleiras, o caldo tratado é, posteriormente, concentrado em vaporizadores de múltiplo efeito e cozedores para cristalização da sacarose. A parte não cristalizada (mel ou melaço) retorna ao processo para recuperação de mais açúcar ou então, é utilizado como matéria-prima na produção de etanol mediante fermentação devido ao seu alto teor de açúcares redutores²³.

A produção de etanol de primeira geração, em usinas produtoras apenas de etanol, é realizada através do uso de caldo direto. As primeiras etapas do processo de

fabricação, da recepção da cana ao tratamento inicial do caldo, são semelhantes ao processo de fabricação do açúcar. Em um tratamento mais completo, o caldo passa pela calagem, aquecimento e decantação, assim como no processo do açúcar. Uma vez tratado, o caldo é evaporado para ajustar sua concentração de açúcares e, eventualmente, é misturado com o melaço, dando origem ao mosto, uma solução açucarada e pronta para ser fermentada. Este (mosto) é fermentado nas dornas por leveduras da espécie *Saccharomyces cerevisiae*, formando o vinho que é, posteriormente, destilado. O etanol é, então, recuperado inicialmente na forma hidratada (96° GL); que pode ser desidratado, formando o etanol anidro (99,7° GL)^{23,25,26}.

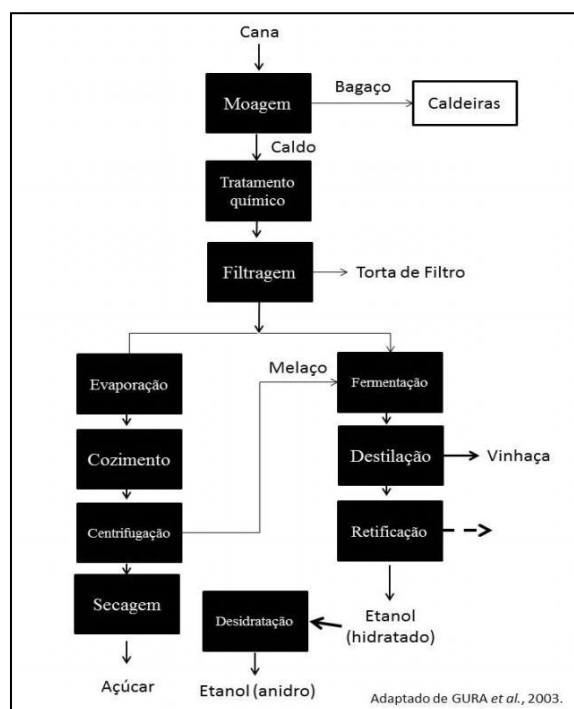


Figura 1: Diagrama de fluxo da produção de açúcar e bioetanol de cana.

A cana é processada, formando o caldo e bagaço. O caldo, tratado e filtrado segue dois caminhos: evaporação, cozimento, centrifugação e secagem para produção de açúcar; ou fermentação, destilação e retificação para produção de bioetanol hidratado. O melaço (parte da sacarose não cristalizada) pode ser utilizado junto ao caldo (mosto) para também produção de etanol. O etanol hidratado pode sofrer desidratação em colunas de desidratação para a formação do etanol anidro.

2.2 Biomassa e sua relação energética

Do ponto de vista ecológico, biomassa é denominada como a quantidade total de matéria viva existente num ecossistema ou numa população animal/ vegetal. Todavia, do ponto de vista da geração de energia, denomina-se biomassa qualquer matéria de origem vegetal que dispõe de bioenergia e que pode ser processada para fornecer formas bioenergéticas mais elaboradas e adequadas para o uso final; sendo considerada um recurso natural renovável²⁷. Isso porque a planta de origem consome o CO₂ advindo da combustão desta fonte; permitindo um equilíbrio na emissão de CO₂, sendo, portanto, menos poluente. No contexto bioenergético, os tradicionais combustíveis fósseis são exclusivos, haja visto que, apesar destes serem derivados de vida vegetal/ animal (carvão mineral, petróleo e gás natural); são resultados de várias transformações que acontecem por milhões de anos; não sendo renovados a curto prazo.

2.2 Bioetanol (*Etanol de Segunda Geração*)

Com exceção da cana-de-açúcar, as tecnologias comercialmente disponíveis na atualidade para a produção de etanol de primeira geração a partir de biomassas sacaríneas e amiláceas envolvem ganhos energéticos e ambientais bastante estreitos. Já para a produção de etanol de segunda geração a partir de materiais lignocelulósicos, parece ser uma alternativa bastante viável em ambos os aspectos (Tabela 1). Além disso, essas matérias-primas apresentam uma limitada vantagem econômica e encontram, em geral, mercados alternativos mais remuneradores, como alimentos ou insumos para outros fins²³.

Na seleção de um cultivo como fonte de matéria-prima para a fabricação de bioetanol, é imperativo considerar os pressupostos de eficiência, em um sentido amplo. Assim, cabe priorizar os cultivos que minimizem os requerimentos de terra, água e aportes externos de agroquímicos, entre outros aspectos. [...], a viabilidade econômica deve ser levada em conta, havendo pouco sentido em propor o uso de cultivos nobres e de bom valor de mercado como fonte de bioenergia. A matéria-prima representa entre 60% e 70% do custo final do bioetanol e a busca de alternativas de baixo custo é fundamental (MASIERO e LOPES, 2008)²³.

Tabela 1: Comparação das diferentes matérias-primas para a produção de etanol e bioetanol

Matéria-prima	Relação de energia	Emissões evitadas
<i>Cana</i>	9,3	89%
<i>Milho</i>	0,6 – 2,0	-30% a 38%
<i>Trigo</i>	0,97 – 1,11	19% a 47%
<i>Beterraba</i>	1,2 – 1,8	35% a 56%
<i>Mandioca</i>	1,6 – 1,7	63%
<i>Resíduos lignocelulósicos*</i>	8,3 – 8,4	66% a 73%

*Estimativa teórica, processo em desenvolvimento. Nesta tabela este é a única fonte utilizada na produção de bioetanol.

Fonte: Elaborado com base em DAI *et al.*, 2006²⁸; EBAMM, 2005²⁹; IEA, 2004³⁰; NGUYEN *et al.*, 2007³¹ apud BNDES, 2008²⁷.

Mesmo assim, apesar de suas vantagens, a cana-de-açúcar não é uma opção viável para todas as regiões do planeta; e por esse motivo, a busca incessantemente por rotas tecnológicas se torna necessária para que haja uma produção de biocombustível eficiente, tanto do ponto de vista ambiental quanto do ponto de vista econômico²³.

Nesse tocante, estudos atuais buscam o desenvolvimento de processos biotecnológicos que permitam a utilização de biomassas residuais de composição lignocelulósica, como palha de milho e arroz, bagaço de cana-de-açúcar e resíduos da indústria de celulose, abundantemente geradas nos setores agrícolas e florestais, para a produção de etanol de segunda geração³². Isso porque além de possuírem um custo mais acessível, são abundantes e podem ser produzidos nas mais variadas condições de solo e clima¹³.

O bioetanol advindo de biomassa lignocelulósica tem sido produzido desde o final do século XIX por meio da utilização da hidrólise e fermentação destes materiais lignocelulósicos; todavia, somente nos últimos 20 anos essa tecnologia tem sido proposta para atender o mercado de combustíveis. Os principais programas de pesquisa e desenvolvimento são conduzidos nos Estados Unidos e na Europa, basicamente em escalas experimentais de produção. Seu sucesso poderia transformar o bioetanol em um biocombustível passível de ser produzido em quase todas as regiões do mundo, aproveitando a alta disponibilidade de resíduos orgânicos de diversas fontes. Praticamente todos os resíduos de biomassa, produzidos nas atividades agrícolas e industriais, e mesmo o lixo urbano, apresentam elevados teores de materiais lignocelulósicos^{18,27}.

A população mundial produz, anualmente, milhões de toneladas de resíduos lignocelulósicos que poderiam ser utilizados como matéria-prima para a produção de bioetanol. Somente no Brasil, atualmente, são produzidos cerca de 150 milhões de toneladas de material lignocelulósico vindo apenas da palha e bagaço de cana-de-açúcar²³. Nesse sentido, se for assumido que cerca de 30% do bagaço produzido esteja disponível como excedente, i.e, sem destino adequado, (o restante corresponderia ao consumo da agroindústria, com alguma geração de excedentes de eletricidade, dependendo da tecnologia adotada) e recuperando-se em torno de 50% da palha da cana, poderiam ser disponibilizados em torno de 60 milhões de toneladas de material lignocelulósico (em base seca), já coletado e com certo preparo²³. A utilização desta biomassa excedente como matéria base poderia aumentar produção de etanol em até 50%^{13,33}; sendo, portanto, de grande interesse para o mercado alcooleiro.

Mesmo assim, a produção de bioetanol ainda não é eficiente, devido ao alto custo e a baixa atividade, nas condições fermentativas, das enzimas atualmente

disponíveis no mercado, além de tecnologias não adequadas para a sua produção em larga escala, dificultando o processo de obtenção¹².

2.3 *Bagaço de cana-de-açúcar*

O bagaço de cana-de-açúcar, é constituído por conjunto de partículas de diferentes tamanhos, que oscilam entre 1 e 25 mm³⁴, apresentam um tamanho médio de 20 mm; sendo o subproduto (assim como a palha e a sacarose) remanescente da moagem dos colmos da cana-de-açúcar^{35,36}.

De modo geral, a morfologia do bagaço se encontra relacionada com a estrutura da cana, onde as fibras exteriores do colmo ou da casca, feixes fibrovasculares e outras formas fibrosas dão resistência ao colmo vegetal. Na estrutura fibrosa do bagaço encontram-se as frações polissacarídicas constituídas basicamente de hemicelulose (30-35%) e celulose (40-45%), infiltradas por macromoléculas, de álcoois aromáticos, denominados ligninas (25%)^{12,35}. Em menores proporções podem ser encontradas resinas, taninos, ácidos graxos, fenóis, compostos nitrogenados e sais minerais. A composição química do bagaço depende das condições climáticas a que esteve exposta a cana como também do tipo de microrganismos que afetam o canavial.

Assim como a palha, o bagaço compõe, em media, de 28% do peso da cana de açúcar, ou seja, a cada tonelada de cana processada, 280 kg de bagaço de cana com 50% de umidade são gerados³⁷. Destes, 50% é destinado a geração de energia³⁸, de modo que parte dessa energia, muitas vezes, é vendida para diferentes cidades, num processo chamado de Cogeração. Assim, numa primeira fase, o bagaço substituiu a lenha para a geração de calor nas usinas.

Atualmente, o vapor produzido através da queima do bagaço é utilizado para gerar três tipos de energia: térmica, transferindo calor para os processos industriais; mecânica, na movimentação de máquinas; e elétrica, através da movimentação de turbinas, suprimindo o consumo energético do parque industrial²³ e, conseqüentemente, permite a redução da emissão de carbono para a atmosfera.

Além da geração de energia, o bagaço pode ser utilizado ainda para a produção de papel e papelão, aglomerados, ração animal e adubo³⁵. Por um mecanismo de hidrólise, a fração hemicelulósica é separada e os açúcares, como xilose, glicose e arabinose são liberados^{35,39,40}, podendo ser utilizados em diversos processos biotecnológicos. Por exemplo, a utilização em processos fermentativos para obtenção de proteína microbiana⁴¹; butanodiol⁴²; xilitol⁴³; etanol⁴⁴ dentre outros.

Apesar dessas aplicações, entre 10% e 15% do bagaço é gerado como excedente no processo, não tendo uma destinação apropriada, tornando-se poluente^{45,46}. Esse excedente poderia ser utilizado para a conversão em bioetanol⁴⁷; todavia, uma análise das características biológicas e morfológicas das fontes e da matéria lignocelulósica remanescente do processamento da cana faz-se necessária, já que o desenvolvimento de tecnologias de hidrólise deve ser específico para as características ou condições da lavoura brasileira e compatíveis com as necessidades do setor sucroalcooleiro⁴⁸.

2.4 Palha da cana-de-açúcar

A palha de cana-de-açúcar é gerada a partir do desfolhamento da cana durante o processo de colheita, podendo ser obtido até 140 kg de palha (massa seca) por tonelada de cana moída^{13,23}. Possui composição semelhante ao bagaço no que se refere à celulose e hemicelulose, contudo a quantidade de lignina é cerca de 30% menor⁴⁹.

Na tabela 2, verifica-se a composição do bagaço e da palha da cana-de-açúcar. Nota-se que a composição de celulose tanto em bagaço quanto em palha são semelhantes, mas que a palha apresenta menores valores de umidade em detrimento do aumento de cinzas presentes em sua composição que é uma característica importante para sua hidrólise⁵⁰.

Tabela 2: Composição média do bagaço e palha da cana-de-açúcar

Composição (Base seca)	Bagaço	Palha
<i>Celulose</i>	46,6 %	45,1 %
<i>Hemicelulose</i>	25,2 %	25,6 %
<i>Lignina</i>	20,7 %	14,1 %
<i>Organosolúveis</i>	2-3 %	3,5 %
<i>Aquosolúveis</i>	2-3 %	-
<i>Cinzas</i>	2-3 %	8 %
<i>Umidade</i>	48-52 %	9,7 %

Fonte: ROSSELL, s.d.⁵⁰

Em muitos Estados, as usinas empregavam e ainda empregam (mesmo que em menor escala) o método da queima da palha para a remoção do colmo. A queima facilita a colheita manual, diminuindo, de forma considerável, riscos de acidentes por picadas de animais peçonhentos. Todavia, a queimada do canavial leva a inúmeros problemas de ordem ambiental e de saúde pública; como exemplo a liberação de gases que contribuem para o efeito estufa; morte de espécies presentes no local e geração de sérios problemas respiratórios para a população exposta devido à grande produção de fuligem e diversos compostos tóxicos que comprometem a qualidade do ar nas regiões de plantio, além de provocar acidentes rodoviários^{12,19,51}.

Em adição, a sustentabilidade da agricultura também fica comprometida, levando em consideração que o fogo altera as propriedades químicas, físicas e biológicas do solo, o que prejudica sua reciclagem nutricional e reduz o percentual de açúcares redutores totais (ART) contidos na cana⁵¹.

Por estes motivos, em 2002/2003 foi regulamentada a Lei N. 11.241/02 no estado de São Paulo, que proíbe a queima da palha nos canaviais e estima uma redução progressiva da queima até a eliminação total. Isto faz com que a palha da cana torne-se um dos principais resíduos gerados, devido à mecanização da colheita⁵¹. Essa lei deve ser empregada nos demais Estados produtores, o que levará a um aumento na quantidade de palha nos campos⁷.

Mesmo assim, quando em excesso, a palhada pode prejudicar as plantas em desenvolvimento, afetando a umidade, luminosidade e temperatura do solo; dificultando as operações de cultivo e adubação da cana soca (resultante da rebrota)⁵² e reduzindo a taxa líquida de mineralização de nitrogênio no período de um ano agrícola⁵³. Além disso, pode favorecer a colonização por parasitas de sementes, que se abrigam sob a palha e dificultam a seleção de plantas^{27,49}. Além disso, ocorre também um aumento significativo na proliferação de insetos e roedores nestas regiões. O grande volume de palha sobre a cana soca dificulta a sua emergência, levando a uma falha na rebrota³⁴. Por outro lado, a retirada completa da palha de cana poderia comprometer a fertilidade do solo, desencadear processos erosivos e alterar a retenção de água no solo.

Assim sendo a palhada (Figura 2), quando deixada na superfície do solo em uma quantidade favorável permite uma melhoria em sua fertilidade, através do retorno nutricional por meio de processos de mineralização, controle de processos erosivos e maior retenção de água, além da propensão ao aumento na microbiota do solo; reduzindo o uso de fertilizantes químicos nos canaviais⁵¹.



Figura 2:Palha ou Palhada.

Material que fica no campo após a colheita da cana. É composto por folhas verdes, pontas do vegetal, palha e restos do caule.

Assim sendo, o excesso desses resíduos vem impulsionando inúmeras pesquisas que visam desenvolver tecnologias de aproveitamento mais racional. Apesar disso, devido à grande quantidade de resíduo gerada, juntamente com o bagaço de cana, ocorre um desperdício de uma importante fonte lignocelulósica para a geração de energia elétrica, biocombustíveis e fabricação de produtos como bioplásticos, carvão para siderúrgicas e até cimento⁵⁴.

2.5 Materiais Lignocelulósicos: entendendo sua estrutura

A utilização de resíduos da cana (bagaço e palha) tem sido cogitada como fonte para a produção de álcool de modo a aumentar sua quantidade, em volume final¹⁸. Para que isto ocorra sem a necessidade de ampliar as áreas cultivadas, é necessária a compreensão da composição estrutural das fontes de materiais lignocelulósicos, i.e., compreender a disposição espacial das fibras celulósicas do bagaço e palha da cana e,

dessa maneira, desenvolver metodologias capazes de desconstruí-lo, mediante acesso enzimático, para a geração de glicoses fermentáveis para a produção de etanol de segunda geração: um combustível totalmente renovável³.

Essa ideia é bastante razoável para uma demanda crescente de álcool combustível, já que maior custo no processo fermentativo está diretamente relacionado ao substrato. A expansão do processo requer, dessa maneira, a utilização de substratos que detenham menor custo como é o caso da lignocelulose que oferece um grande potencial na redução de custeio e um aumento na produção de etanol¹⁸.

Por ser de origem vegetal, a lignocelulose pode ser generalizada como a matéria-prima de maior disponibilidade na Terra, sendo uma fonte abundante para a produção de biocombustíveis. No entanto, uma grande limitação na utilização desta biomassa vegetal como fonte de energia renovável é a eficiência ineficaz da despolimerização industrial da lignocelulose⁵⁵⁻⁵⁸.

Isso porque, os materiais lignocelulósicos são formados por estruturas duras e fibrosas, compostas majoritariamente pelos polissacarídeos celulose e hemicelulose (cerca de 70% da massa seca), entremeados por outra macromolécula formada por álcoois aromáticos, a lignina, aos quais se encontram unidos por ligações covalentes e de hidrogênio (Figura 3)^{16,59,60}. Em menores proporções, e dependendo da origem do vegetal, também podem ser encontrados resinas, ácidos graxos, fenóis, taninos, compostos nitrogenados e sais minerais, principalmente, de cálcio, potássio e magnésio⁶¹⁻⁶⁴. É essa estrutura que age como barreira protetora natural à degradação enzimática e/ou microbiana⁶⁵.

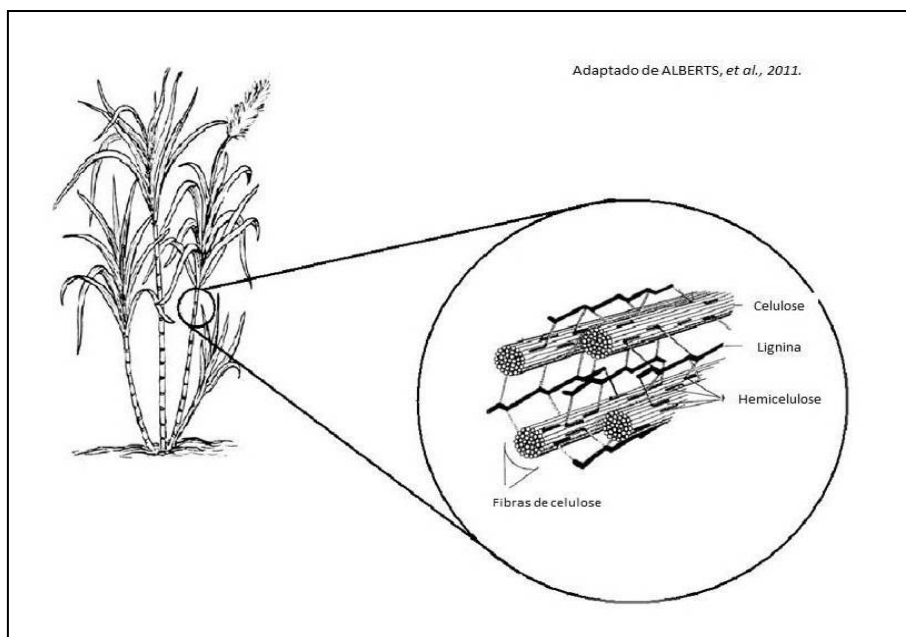


Figura 3: Composição da lignocelulose.

Os materiais lignocelulósicos ou biomassa vegetais, como o bagaço, são constituídos por três principais componentes macromoleculares: celulose, hemicelulose e lignina e apresentam uma estrutura lamelar, na qual se distribuem seus componentes. A celulose e a hemicelulose predominam na região da parede celular e a lignina se distribui por toda a estrutura⁶⁶.

A composição e a estrutura da biomassa tem forte influência na natureza e nos rendimentos dos processos de hidrólise e fermentação. Sendo que a composição básica da biomassa lignocelulósica depende do vegetal de origem (Tabela 3). Celuloses e Hemiceluloses podem ser fermentadas em etanol; lignina não⁶⁷.

Tabela 3: Composição química parcial de alguns materiais lignocelulósicos

Material lignocelulósico	Celulose (%)	Hemicelulose (%)	Lignina (%)
Farelo de cevada	23,0	32,7	24,4
Sabugo de milho	31,7	34,7	20,3
Folhas de milho	37,6	34,5	12,6
Bagaço de cana	40,2	26,4	25,2
Palha de arroz	43,5	22,0	17,2
Palha de trigo	33,8	31,8	20,1
Palha de sorgo	34,0	44,0	20,0

Casca de aveia	30,5	28,6	23,1
<i>Eucalyptus grandis</i>	40,2	15,7	26,9
<i>Eucalyptus globulus</i>	46,3	17,1	22,9

Fonte: CANETTI, SILVA e FELIPE, 2001⁶⁸.

Desta maneira, faz-se necessário conhecer a estrutura e as principais características dos principais componentes da biomassa lignocelulósica: celulose, hemicelulose e lignina.

2.5.1 Celulose

A celulose (23% - 50% da matéria seca da biomassa lignocelulósica) é um polímero linear que contém entre 2000 a 25000 unidades de β -D-glicoses (β -D-glucopiranosose) organizadas em dímeros de celobiose (Figura 4); sendo unidas por ligações glicosídicas do tipo β -1,4 carbono-carbono pela eliminação de uma molécula de água entre seus grupos hidroxilas do carbono 1 e carbono 4 (ligação β 1 \rightarrow 4), de fórmula geral $(C_6H_{10}O_5)_n$ ⁶⁶; e por ligações de hidrogênio intramoleculares (ligações entre unidades de glicose da mesma molécula) e intermoleculares (entre unidades de glicose de moléculas adjacentes) (Figura 5)^{69,70}. A celobiose é definida como a unidade conformacional mínima da celulose, enquanto que a glucose apresente somente a unidade funcional das cadeias do homopolímero^{65,66}.

A ocorrência de uma ligação hemiacetal na molécula de glicose, devido à liberdade rotacional da cadeia linear, faz com que um grupo aldeído esteja presente no carbono-1, devido ao ataque da hidroxila do carbono 5⁶⁵. Esta ocorrência faz com que uma das extremidades da cadeia celulósica apresente um terminal redutor⁶³.

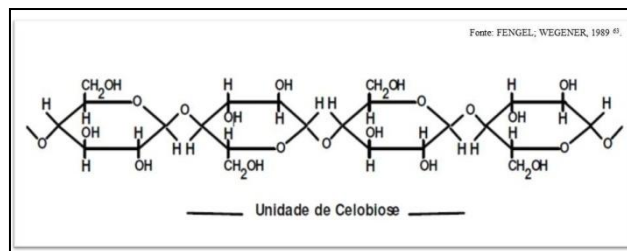


Figura 4: Estrutura do Fragmento de Celulose.

Estritamente, a celulose é uma cadeia polimérica onde a unidade repetitiva é a celobiose, que é formada a partir de duas β -D-glicose unidas por ligações do tipo β -1,4, cuja qual permite a linearidade da cadeia pela ligação dos carbonos 1 e 4 e eliminação de água. Trata-se, pois, da unidade conformacional mínima da celulose.

As ligações intermoleculares são responsáveis pela rigidez; e as ligações intramoleculares são responsáveis pela formação de fibrilas, estruturas altamente ordenadas que se encontram associadas às fibras de celulose (microfibrilas)^{63,69,70}. As microfibrilas consistem, portanto, em agregados cristalinos altamente organizados e com micrometros de extensão; possuindo uma força de tensão semelhante ao aço (Figura 6)⁷¹.

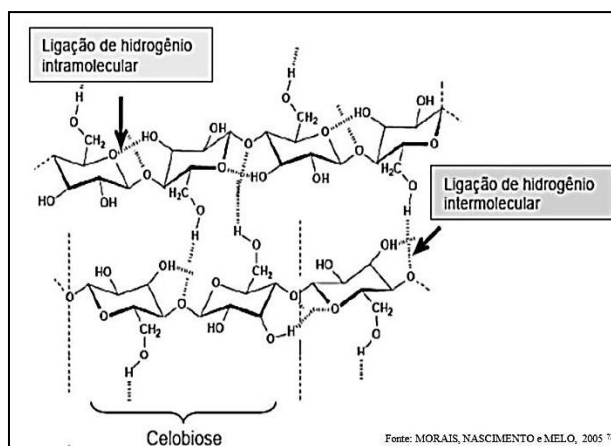


Figura 5: Estrutura química da celulose.

Ligações β -1,4; de hidrogênio intra e intermoleculares podem ser vistas na figura acima. As ligações intramoleculares formam as fibrilas, ligando as glicoses da mesma molécula enquanto que as ligações intermoleculares conferem rigidez a estrutura, ligando glicoses entre moléculas adjacentes. Ligações β -1,4 permite a ligação entre glicoses pelos carbonos 1 e 4 para a formação da celobiose, permitindo uma conformação de dímero, onde as glicoses encontram-se invertidas um em relação a outra.

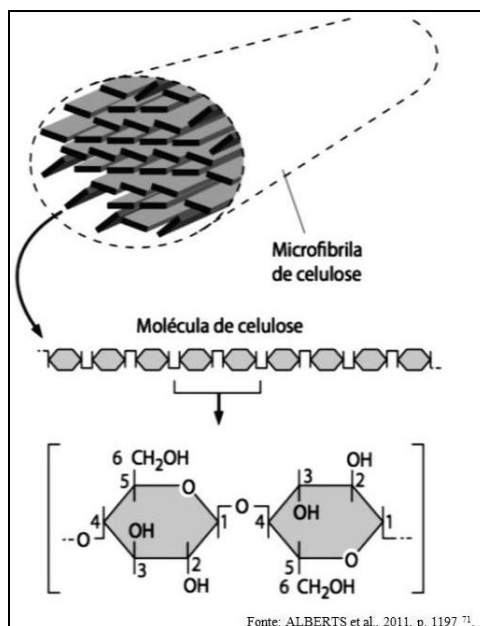


Figura 6:Microfibrila de Celulose.

As moléculas de celulose são longas, não ramificadas e constituídas por unidades de glicose com ligações do tipo β -1,4. Cada resíduo de glicose é invertido com relação ao seu adjacente (como mostrado na parte inferior da imagem), e as repetições de celobioses resultantes ocorrem centenas de vezes em uma única molécula. Cerca de 40 moléculas de celulose reunidas formam as microfibrilas, onde estão ligadas por ligações de hidrogênio.

O índice de cristalinidade e o grau de polimerização são propriedades importantes para a classificação dos polímeros celulósicos; de modo que o grau de polimerização confirma a frequência relativa de ligações glicosídicas internas e terminais, disponíveis para atuação de celulases; sendo definido com base no número médio de monômeros e no peso médio do polímero, ou ainda a partir de sua viscosidade. Já o índice de cristalinidade está associado à reatividade do substrato e pode ser quantificado pelo método de difração de raios X ^{60,62,64}.

Estas características, juntamente com o envoltório de lignina, conferem à macromolécula celulose grande resistência à hidrólise, o que representa um grande desafio para a utilização dos materiais lignocelulósicos em aplicações biotecnológicas, como a produção de etanol de segunda geração ⁶⁹.

2.5.2 Hemicelulose

Hemicelulose (Figura 7) é um polímero heterogêneo de baixa massa molecular – 100 a 200 unidades glicosídicas –, (formado por pentoses, hexoses e açúcares ácidos), possuidores de cadeias de açúcares ramificadas (glicanos), cujas unidades incluem principalmente aldopentoses, como xilose e arabinose, e aldohexoses, como glicose, manose e galactose⁷³, cujas quais as classificam. Esta macromolécula contém ainda, ácidos hexurônicos, como os ácidos β -D-glucurônico, D-4-O-metilglucurônico e β -D-galacturônico, e deoxiexoses (Figura 8). A variedade de ligações e de ramificações, assim como a presença de diferentes unidades monoméricas, contribui para a complexidade da estrutura hemicelulósica e suas diferentes conformações⁷⁴.

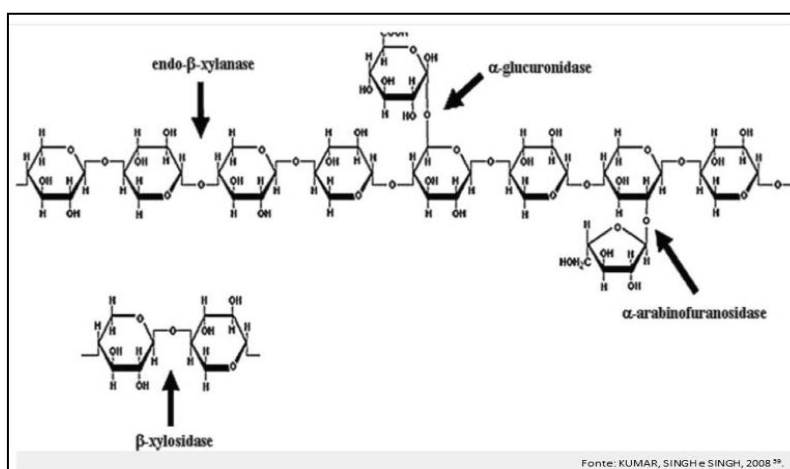


Figura 7: Estrutura química polimérica de hemicelulose e locais de hidrólise enzimática envolvidos na degradação do polímero de hemicelulose.

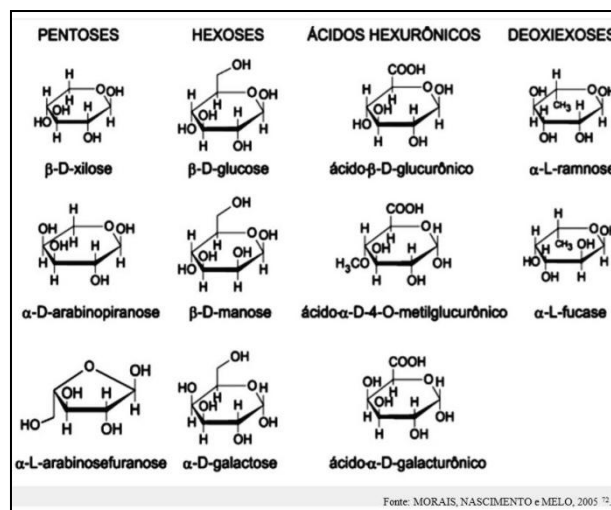


Figura 8: Componentes da fração hemicelulose.

Aldopentoses, como xilose e arabinose, e aldohexoses, como glicose, manose e galactose; ou ainda, ácidos hexurônicos, como os ácidos β-D-glucurônico, D-4-O-metilglucurônico e β-D-galacturônico, e deoxiexoses são alguns dos componentes que podem estar presentes nas estruturas das hemiceluloses.

Isso porque, as cadeias podem ser constituídas por apenas uma unidade monossacarídica, como exemplo as xilanas, ou por duas ou mais, como o caso das xiloglucanas ou arabinoxilanas ^{63,65}. Essa natureza química é variável entre os vegetais tanto em relação ao tecido vegetal quando a espécie. Em geral, compõe cerca de 20 a 30% da biomassa seca, podendo chegar até 40% ^{65,67,70,75}.

Trata-se, pois, de uma estrutura semelhante a celulose e são depositadas na parede celular antes da lignina. As ramificações e cadeias laterais interagem facilmente com a celulose, conferindo estabilidade e flexibilidade ao agregado ⁷⁶, formando uma rede complexa de função semelhante ao colágeno ⁷¹.

A hemicelulose é essencialmente não cristalina (caráter relativamente amorfo) e geralmente mais reativa do que a celulose, i.e, mais suscetível à hidrólise química sob condições mais brandas ⁷⁷, uma vez que oferecem maior acessibilidade aos ácidos minerais comumente usados como catalisadores ⁶⁵. Todavia, a fermentação dos açúcares

de cinco carbonos (pentoses) ainda não é tão desenvolvida quanto os processos envolvendo a glicose¹⁶.

2.5.3 Lignina

A lignina, depois da celulose, é um dos polímeros mais abundantes e importantes da natureza. Sua estrutura bioquímica (composta por 7 a 30% de massa seca) não se baseia em uma molécula simples de açúcar⁷⁵, e, portanto, não é utilizada para a produção de bioetanol por rotas fermentativas (Figura 9).

Por deter estrutura amorfa e de alta massa molecular⁷³ e fortemente aromática, essa fração tridimensional formada por unidades de *p*-propilfenol, desempenha um papel fundamental para o sucesso da tecnologia de hidrólise, uma vez que dificulta o acesso à celulose. Isso porque, a lignina é formada pela polimerização de três diferentes monômeros: álcool cumárico, álcool coniferílico e álcool sinapílico (Figura 10)⁷⁸ – uma das etapas finais da diferenciação celular do xilema⁶⁵ -; geralmente associada à celulose e às hemiceluloses através de ligações éter e carbono-carbono (ligações covalentes com as unidades monossacarídeos). Assim, tal polímero ocorre como uma “rede” que circunda as microfibrilas ricas em carboidratos das paredes celulares de plantas, além de estarem presentes, também, nos espaços intersticiais entre as células de plantas maduras⁶⁷.

Mesmo assim, existe uma grande diversidade na estrutura da lignina e quantidade quando se passa de uma espécie vegetal para outra, ou até mesmo dentro da mesma espécie, quando são analisadas partes diferentes do vegetal⁷⁹. Como resultado dessa diferença, sua estrutura química acaba sendo bastante complexa e não é bem conhecida ainda⁶³.

De mais a mais, sendo um dos principais componentes dos tecidos de gimnospermas e angiospermas, possui importância no transporte de água, nutrientes, metabólitos, além de conferir resistência mecânica e proteger contra o ataque de microorganismos⁶³.

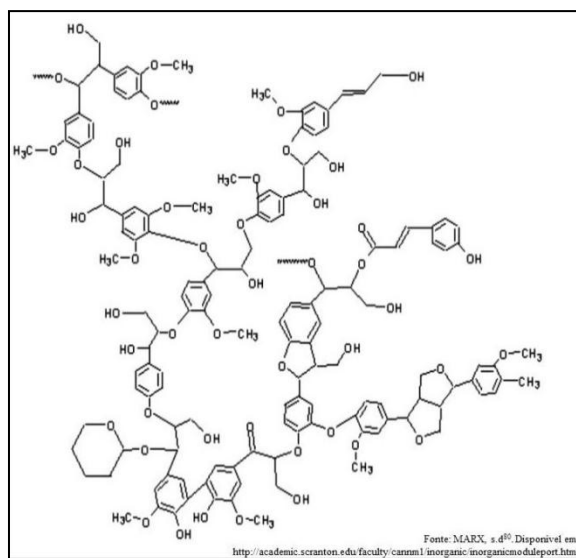


Figura 9: Estrutura geral da Lignina.

A lignina é um polímero amorfo presente nas paredes celulares de plantas, conferindo resistência e proteção, além de transporte de nutrientes, metabólitos e água.

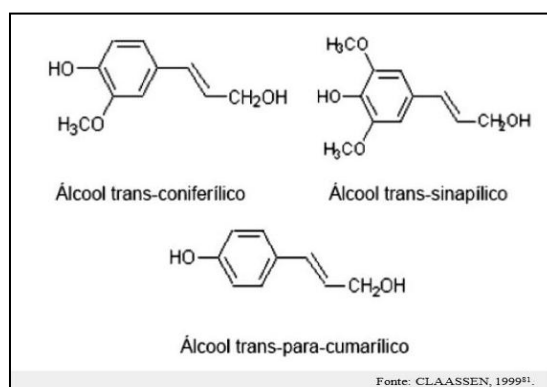


Figura 10: Precursores primários da lignina.

A lignina é formada via polimerização dos alcoóis trans-coniferílico, trans-sinapílico e trans-para-cumarílico; visando a formação de uma cadeia amorfa que se associará com as celulosas e hemicelulosas, formando uma complexa rede entre as microfibrilas e espaços intersticiais.

A lignina representa um dos maiores estoques de carbono/energia da natureza e é o maior depósito de estruturas químicas aromáticas, constituindo-se em uma fonte potencial de valiosos insumos para a indústria química^{67,78,79,82} como vanilina, ácido gálico, óleos, fenol, ácido acético, entre outros^{68,83}; devido a sua natureza aromática e muito complexa. Também é possível remover metais pesados de efluentes usando ligninas oxidadas^{68,82}, bem como usá-la na fabricação de filmes, poliuretanos, macro monômeros de poliésteres, fibras de carbono, fertilizantes entre outros⁸⁴.

Dessa maneira, a parede celular das células vegetais é formada por uma matriz rígida constituída principalmente por celulose, hemiceluloses e lignina, além de outros componentes que dificultam, como previamente dito, o acesso enzimático para uma possível hidrólise. Isso porque, sua composição forma uma “rede” de fibrilas emaranhadas (Figura 11), cuja qual pode ser mais ou menos flexível^{29,62,63,73}. A lignina, neste caso forma uma camada que confere maior resistência à tração, ao alongamento e à solvatação (absorção de solventes, como a água), formando, assim, regiões amorfas e regiões mais cristalinas⁷³ (Figura 12); de acordo com o grau de ordenação dessa rede. Regiões amorfas tendem a serem mais flexíveis, macias e podem absorver água. Essa rede confere a forma final da célula e da planta como um todo⁷¹.

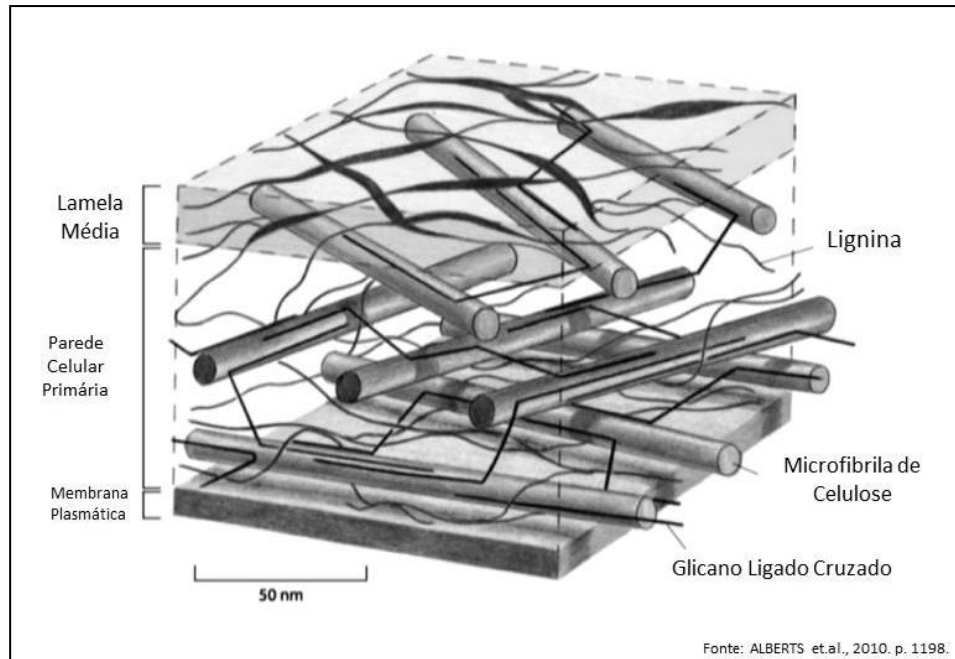


Figura 11: Estrutura da parede celular vegetal.

A parede vegetal primária é principalmente constituída por longas e resistentes microfibrilas de celulose que estão em associação com as hemiceluloses (polissacarídeos ramificados) através de ligações de hidrogênio; e lignina através de ligações éter e carbono-carbono; que juntos conferem uma “rede” rígida e protetora.

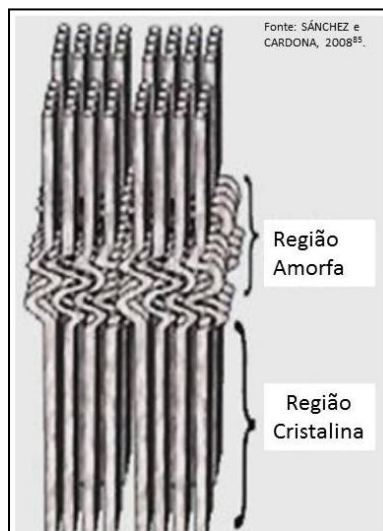


Figura 12: Fibrilas.

As fibrilas são compostas por regiões cristalinas, que são ordenadas e garantem a rigidez; e por regiões amorfas, que permitem a flexibilidade do feixe.

2.6 *Produção de Etanol a partir de Materiais Lignocelulósicos*

A produção de etanol a partir da cana-de-açúcar ocorre, atualmente, através da fermentação alcoólica da sacarose, permitindo uma fonte de energia renovável na matriz de combustível líquido principalmente para o transporte; contudo, a quantidade de etanol necessária para atingir as metas planejadas exigirá outras fontes de matéria-prima em detrimento da cana e milho, uma vez que existe uma competição direta entre a produção de combustível e a produção de alimentos, causando um impacto na sua disponibilidade para usos tradicionais e conseqüentemente nos preços dos alimentos derivados⁸⁶.

Nesse tocante, o etanol produzido a partir biomassa celulósica tem recebido atenção como um potencial combustível^{18,67}. Isso porque, nos últimos cinco anos ou mais, tem-se aumentando o interesse comercial em relação ao etanol, causando uma intensa competitividade de custos em tecnologias que convertam a lignocelulose em bioetanol⁶⁷.

Os resíduos da agroindústria como o bagaço e palha da cana, palha de trigo, entre outros materiais lignocelulósicos estão sendo visados para a obtenção de etanol a partir da celulose e, assim, complementar a produção do etanol combustível. Todavia, desafios como (i) a implantação de um sistema logístico para transportar a palha de cana, por exemplo, até a usina e (ii) processos de conversão eficientes e viáveis economicamente^{86,87} ainda necessitam ser superados.

Justifica-se, pois, que no processo de obtenção do etanol celulósico, o objetivo principal é “quebrar” e “desmontar” a parede celular, liberando os polissacarídeos, que nada mais são do que fonte de açúcares fermentáveis. Contudo, já se salientou aqui o quão complexa é a estrutura da parede celular vegetal e o quão esta é variável tanto

entre espécies, na própria espécie, ou até mesmo, entre tecidos. Atualmente, esse processo é feito por meio de uma hidrólise ácida que, embora seja funcional, não é eficiente para uma produção comercial do bioetanol ⁸⁸, uma vez que é necessária a utilização de temperaturas elevadas e de ácidos mineiras fortes para satisfazer a hidrólise ⁸⁹.

Dessa forma, para superar esses obstáculos, diversos pesquisadores estão desenvolvendo meios para, em curto prazo, os materiais lignocelulósicos possam ser utilizados na produção de combustível. Com isso, em longo prazo, a tecnologia do etanol celulósico servirá de base para a produção sustentável de *commodities*, fazendo uso de diversos produtos químicos e materiais obtidos a partir do fracionamento da biomassa, surgindo assim as biorrefinarias ^{3,19,67}. Estas, por sua vez, poderão utilizar a biomassa e seus resíduos, i.e., matérias-primas de fontes renováveis; de forma integrada e diversificada para a produção de combustíveis, produtos químicos, energia e outros materiais de interesse no mercado industrial, com a geração mínima de resíduos e emissões de gases nocivos ^{19,67}.

Acredita-se que essa combinação de processos seja altamente eficiente. Entretanto, a utilização desses materiais lignocelulósicos nos chamados bioprocessos visando à produção de etanol e outros produtos de interesse, necessitam da hidrólise de polissacarídeos, a fim de disponibilizar seus monômeros para os processos posteriores⁸⁸.

Embora as enzimas do complexo celulolítico não penetrem de forma hábil e rápida na barreira de lignina, e que seu processo de degradação seja lento comparado com a hidrólise ácida, podem converter a celulose em glicose de maneira mais seletiva, formando poucos subprodutos residuais. Nesse tocante, faz-se necessário um pré-tratamento a fim de se romper esse selo de lignina e garantir a destruição da cristalinidade da celulose ^{89,90}.

Atualmente muitos pesquisadores estão concentrando esforços no desenvolvimento de pré-tratamentos eficientes, capazes de disponibilizar a maior quantidade de açúcar possível (pentoses e hexoses), no desenvolvimento de microrganismos geneticamente modificados que possam fermentar tanto pentoses quanto hexoses, na produção de plantas geneticamente modificadas com maior quantidade carboidratos ou plantas modificadas estruturalmente para facilitar a etapa de pré-tratamento em condições amenas e na utilização de processos integrados para reduzir o número de etapas do processo e consequentemente reduzir a demanda energética. (da SILVA, 2009⁹⁰ apud DIEN; COTTA; JEFFRIES, 2003⁹¹; JEFFRIES, 2006⁹²; HAHN-HÄGERDAL et al., 2006⁹³).

2.7 Pré-tratamento de Materiais Lignocelulósicos

A eficiente redução da estrutura resistente dos materiais lignocelulósicos e consequente liberação dos polissacarídeos estão dentre uma das mais urgentes e importantes prioridades nas áreas de pesquisa e desenvolvimento (P&D) em etanol celulósico e obtenção de substâncias químicas nas biorrefinarias⁹⁴⁻⁹⁷.

A hidrólise enzimática dos materiais lignocelulósicos é limitada por vários fatores tais como: área superficial acessível (porosidade do material), grau de polimerização e cristalinidade da celulose e umidade do lignocelulósico, por exemplo⁹⁸⁻¹⁰⁰. Dessa maneira, podemos dizer que as duas principais causas de resistência dos materiais lignocelulósicos à hidrólise enzimática são: a baixa acessibilidade às fibras celulósicas (micro)-cristalinas, as quais restringem a atividade das celulasas; e a presença de hemicelulose e principalmente de lignina na superfície da celulose, impedindo a ação das celulasas ao substrato^{96-98,101}. Portanto, a remoção desses compostos aumenta a porosidade do material, tornando mais eficiente o processo de hidrólise⁹⁸.

A quantidade variável de hemicelulose e lignina em um material lignocelulósico depende de fatores como o tipo da planta a partir da qual a biomassa é obtida, idade da colheita, dentre outros; demonstrando que não há nenhum método de pré-tratamento específico para diversos tipos de materiais lignocelulósicos⁸¹.

Nesse sentido, faz-se necessário a utilização de um processamento prévio e adequado destes materiais para que a parede celular (barreira física) seja rompida, e permitir a diminuição da cristalinidade da celulose e a associação com a lignina a fim de que as enzimas hidrolíticas possam acessar a macroestrutura da biomassa¹⁹, visando o aumento do rendimento de conversão da celulose em glicose^{40,75,102} (Figura 13).

Além disso, a etapa de pré-tratamento pode influenciar fortemente o custo do processo por determinar a toxicidade da fermentação, a taxa de hidrólise enzimática, carga de enzima utilizada no processo, o poder de mistura dos reagentes, a concentração e a purificação dos produtos obtidos, a demanda no tratamento de resíduos, entre outras variáveis. “Portanto, esta etapa de processamento da biomassa vegetal necessita ser de baixo custo, evitando-se o elevado consumo de reagentes químicos, a alta demanda energética e a intensa degradação dos materiais lignocelulósicos”^{75,90}.

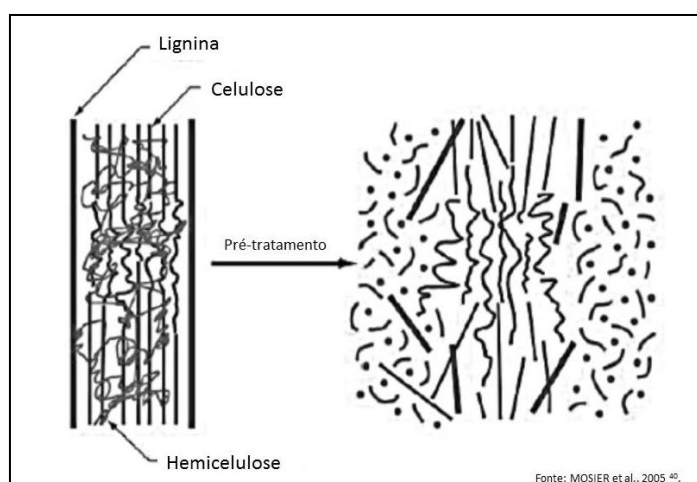


Figura 13: Esquema do efeito do pré-tratamento em materiais lignocelulósicos.

O pré-tratamento para a quebra da barreira física, portanto, pode ser físico (moagem da biomassa, aplicação de radiação), químico (pré-tratamento hidrotérmico, oxidação úmida (Wet Oxidation), pré-tratamento em meio ácido ou básico), biológico (microrganismos decompositores da madeira) ou a combinação desses^{19,40,74,103}; dependendo do grau de separação requerido e fim proposto^{104,105}.

A escolha das condições de pré-tratamento dependerá de uma avaliação global do processo, rendimento e manutenção da integridade dos insumos de interesse¹⁰⁶. Isso porque as condições de pré-tratamento podem ter dois efeitos opostos: se o pré-tratamento for muito drástico, ocorre a degradação da fração celulósica, além da formação de inibidores da fermentação; se for muito brando, a acessibilidade da enzima continuará sendo baixa^{79,86}. No primeiro caso, pode-se realizar um tratamento do hidrolisado para remover alguns inibidores como o furfural e hidroximetilfurfural (HMF), além do o ácido sulfúrico e ácido acético, aplicando-se, por exemplo, os métodos eletrodialise e carvão ativo. A formação de furfural e HMF é decorrente da reação de desidratação da xilose e da glicose, respectivamente. Esses compostos quando presentes no licor a ser fermentado podem danificar a parede celular dos microorganismos bem como prolongar a fase “lag” ou latência/ de adaptação de crescimento^{78,83}.

Dentre estes métodos citados, o mais comumente utilizado é o termoquímico, que desorganiza o complexo lignocelulósico; permitindo a dissolvência da hemicelulose e sua despolimerização quase que por completa. A lignina, por sua vez, pode ser separada utilizando-se solventes orgânicos. Assim, a celulose (material restante do tratamento) tem sua degradação aumentada, permitindo a ação das celulasas para hidrolise enzimática. Por conseguinte, a quebra da celulose gerará glicose, utilizada no processo fermentativo^{17,38,39,44,50,81,82,84}.

2.8 Hidrolise Enzimática da Celulose

Após a etapa de pré-tratamento e a remoção da lignina remanescente deste processo via solubilização dentro do próprio pré-tratamento é necessária uma etapa de hidrólise da celulose (Figura 14) para obtenção de açúcares fermentáveis. No processo enzimático a hidrólise da celulose é catalisada por enzimas específicas (denominadas de enzimas celulolíticas ou celulasas); cujas quais podem ser divididas em pelo menos três tipos principais:

1. *Endoglucanase (endo -1,4- β - glucanase ou β -1,4-endoglucanase ou EC 3.2.1.4 ou EGs)*: cliva randomicamente as regiões de baixa cristalinidade na fibra celulósica, i.e., os sítios amorfos (sítios β -1,4- glicosídicos nas porções internas das celulosas para reduzir o grau de polimerização da cadeia celulósica em pequenas subunidades); liberando oligossacarídeos com terminações redutoras e não redutoras livres^{70,107};
2. *Exoglucanase ou celobiohidrolase (1,4-D-glucano celobiohidrolase ou EC 3.2.1.91 ou CBHs)*: liga-se nas extremidades das cadeias (terminais redutores – CBHs tipo I; e não-redutores – CBHs tipo II), despolimerizando-as gradativamente pela remoção das celobioses terminais; gerando assim, principalmente glicose e celobiose^{65,69}; e
3. *β - glucosidase (EC 3.2.1.2 ou BGs)*: hidrolisam celobiose e oligossacarídeos solúveis de baixo grau de polimerização (menor que 7) à glicose a partir dos terminais não-redutores.

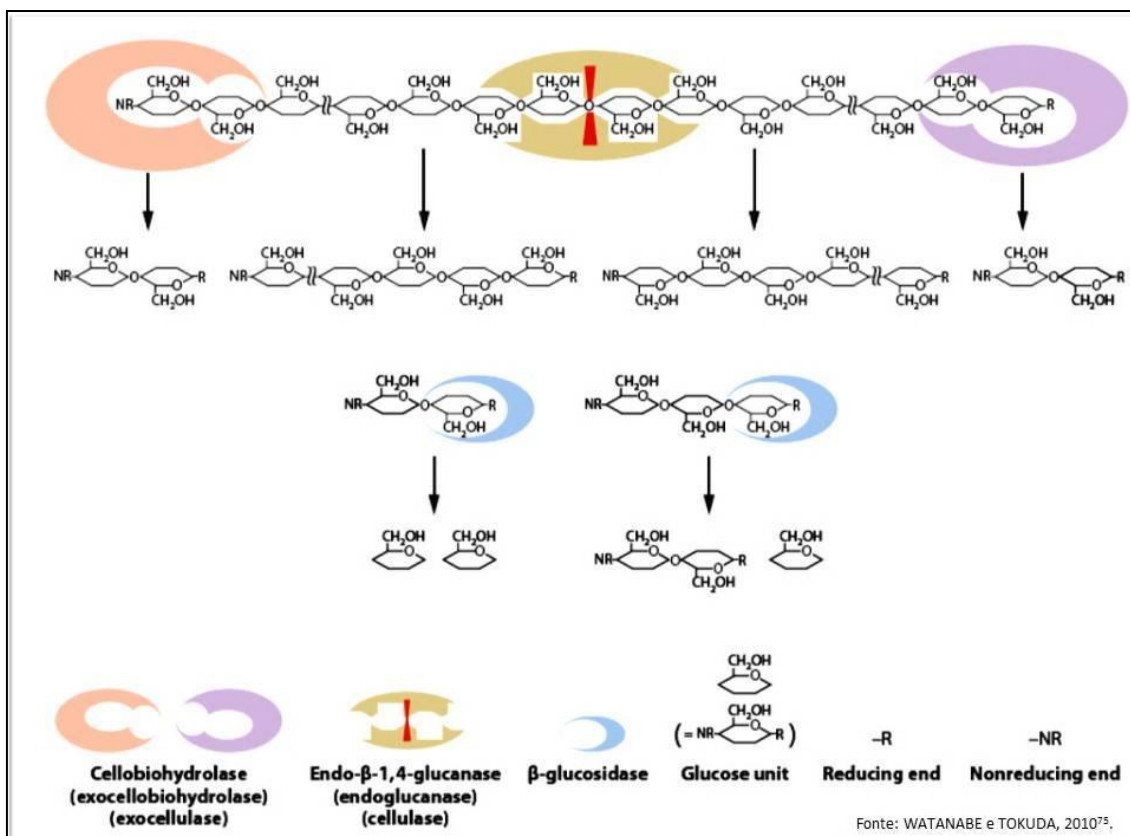


Figura 14: Visão esquemática da biodegradação da celulose.

Celobiohidrolases agem sobre os terminais não-redutores e redutores das fibras de celulose para, progressivamente, liberar celobiose. Endo- β -1,4-glicanase cliva aleatoriamente a cadeia da celulose. β -glicosidase hidrolisa celobiose ou oligossacarídeos em glicose a partir das terminais não-redutoras.

Diante da heterogeneidade da estrutura da cadeia celulósica, a qual apresenta regiões altamente ordenadas, estabilizadas por numerosas ligações de hidrogênio intra e intermoleculares, e áreas menos ordenadas ou amorfas, a sacarificação enzimática da biomassa necessita do sinergismo e complementaridade por parte destas celulasas, visando a hidrólise de todo carboidrato disponível na celulose^{65,81-83}. Além disso, o complexo enzimático precisa ser estável para altas atividades celulolíticas.

O sinergismo (Figura 15) pode ser classificado em ao menos três categorias^{11,67}:

1. **Sinergia endo-exo:** as endoglucanases atuam nas regiões amorfas, liberando terminais redutores e não redutores, nos quais atuarão as CBHs do tipo I e do tipo II, respectivamente.
2. **Sinergia exo-exo:** as CBHs I e CBHs II atuam simultaneamente nos terminais redutores e não redutores liberados pelas endoglucanases.
3. **Sinergia exo-BG e endo-BG:** as exoglucanases e endoglucanases liberam celobiose e oligossacarídeos, respectivamente, que são substratos da β -glucosidase.

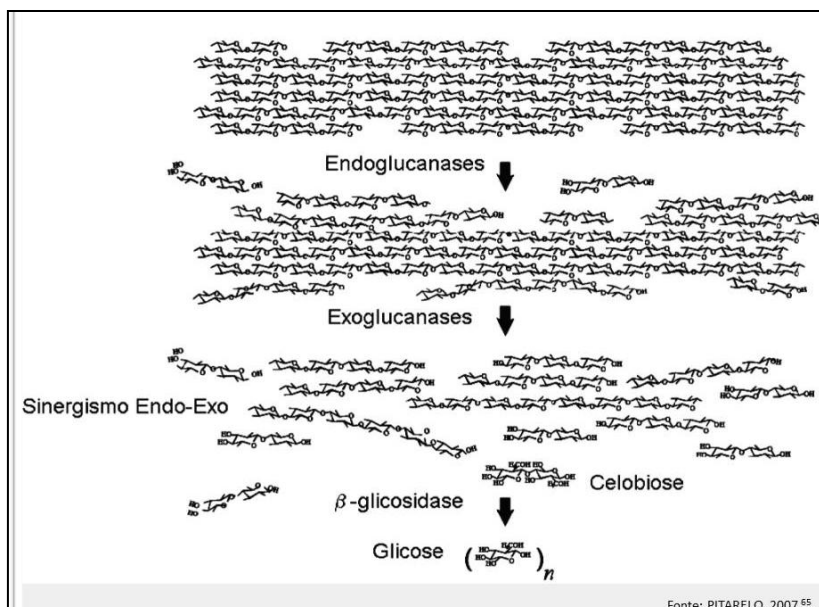


Figura 15: Sinergismo entre celulases.

Modo de ação das celulases resulta em um sinergismo entre as enzimas de modo a promover um sistema enzimático cooperativo na degradação da celulose. Pra isso, inicialmente as endoglucanase quebram as macromoléculas de celulose em partículas menores e, em seguida, há a ação das exoglucanases nos terminais redutores e não redutores e as β -glucosidases agindo sobre os fragmentos de celobioses. Essas três enzimas juntas resultam na degradação da celulose; de modo que uma é dependente da outra para a degradação completa.

As três classes enzimáticas, portanto, contribuem para a degradação da celulose; todavia, apenas a endo e exoglucanases são capazes de adsorver sobre o substrato e, portanto, são consideradas celulases verdadeiras^{65,108}.

Assim, devido a essa interação entre as enzimas, quando individualizadas, as enzimas do complexo celulósico não hidrolisam a celulose de modo eficiente; tal que, pode haver acúmulo de subprodutos, levando a inibição de uma ou outra enzima do complexo. Como exemplo, podemos citar o aumento da celobiose que leva a inibição das endoglucanases e exoglucanases e, portanto, a atuação das β -glicosidases para a hidrólise da celobiose à glicose se torna importante no processo para completar a degradação da celulose¹⁰⁹; sendo, portanto, passo limitante¹⁸. Em adição, sabe-se que a atuação das enzimas depende diretamente, não somente da sua eficácia (atividade enzimática), como também do substrato que é atacado, i.e., embora a composição enzimática seja encontrada para influenciar no grau de solubilização, tipo de açúcar liberado e o local de ataque enzimático dentro do polímero, a acessibilidade dos substratos é outro fator limitante principal para a hidrólise¹⁸. Pode-se citar aqui, também, como fatores limites as condições das reações (temperatura, pH, fonte de carbono, nitrogênio, fosfato, componentes fenólicos, por exemplo⁵⁹) e atividade da celulase.

Vários autores sugerem que a região amorfa da celulose, devido a sua maior área superficial, é mais acessível, i.e., mais suscetível à hidrólise enzimática do que a região ordenada ou cristalina, como já enfatizado ao final da seção 2.4. Todavia, complexos enzimáticos produzidos por vários tipos de organismos^a sugerem a capacidade de catalisar a hidrólise de celulose, tanto cristalina como amorfa, em açúcares solúveis de baixa massa molar, como glicose e celobiose¹⁰⁵.

a. Diz-se organismos, pois atualmente sabe-se que não somente microrganismos são capazes de realizar a quebra de celulose como também animais superiores, que detém celulasas endógenas capazes de realizar a quebra da celulose no trato digestivo. Esse tema, parte principal desta dissertação, será abordado em um tópico específico mais adiante.

Dessa maneira, os açúcares redutores produzidos a partir da hidrólise enzimática da palha e do bagaço, por sua vez, podem ser fermentados por leveduras para a produção de etanol. Contudo, o processo de bioconversão de fitobiomassas, atualmente, passa por dificuldades na implantação, tendo em vista o alto custo e a baixa atividade específica das enzimas necessárias à sacarificação da celulose^{1,28}.

2.9 *Celulases Animais*

Sabe-se, agora, que o interesse nessas enzimas aumentou nas últimas décadas devido a sua utilização na produção de combustíveis alternativos e participação no ciclo global de carbono¹¹⁰. As celulases são encontradas em 14 das 90 Famílias Glicosil Hidrolases conhecidas (GHFs), sendo são classificadas com base nas sequências similares de aminoácidos (GHF 5, 6, 7, 8, 9,10, 12, 26, 44, 45, 48, 51, 61 e 74^{111,112}; veja também CAZY - <http://afmb.cnrs-mrs.fr/~cazy/>). Todavia, a maioria das celulases não está agrupada com base em sua estrutura e sequência, indicando que sua capacidade de hidrólise tem evoluído de forma independente em várias ocasiões¹¹³. Cinco dessas famílias estão representadas no Metazoa.

Muitos microrganismos como exemplo bactérias, fungos e protozoários, podem sobreviver a partir de materiais celulósicos; mesmo assim, acredita-se que a maioria dos animais não pode utilizá-los. Evidências recentes e crescentes indicam que enzimas celulósicas são distribuídas em uma grande gama de animais invertebrados, como ascídias e nematoides, e alguns vertebrados utilizam a biomassa celulósica com ajuda de simbiontes intestinais¹¹⁴.

Essas evidências são corroboradas com a classificação dessas celulases entre as famílias, de modo que são estruturalmente e presumivelmente de famílias filogenéticas

distintas, onde temos a GHF5, GHF9 e GHF45 sendo encontradas, principalmente, em nematoides e besouros, cupins e baratas, besouros e molusco, respectivamente¹¹⁵⁻¹²². A GHF10 também é encontrada em Moluscos¹¹⁴.

Sistemas celulósicos de insetos, por exemplo, vem sendo estudados principalmente em cupins basais e baratas comedoras de madeira, uma vez que se encontram intimamente ligados filogeneticamente¹²³. Sabe-se que tais insetos possuem numerosos protozoários dentro de uma parte ampliada de seu intestino posterior, denominada “*paunch* ou *pança*”^{70,115}.

A teoria da dependência de simbioses foi proposta inicialmente por Cleveland (1923, 1924, 1925, 1934)¹²⁴⁻¹²⁷, em seu trabalho com os cupins, que demonstrou o desaparecimento da viabilidade de Reticulitermes flavipes após a retirada da fauna de protozoários do intestino posterior; concluindo, assim, sua responsabilidade pela digestão da celulose. Desde sua demonstração, a produção de celulases simbióticas tornou-se a explicação mais favorecida para a digestão da celulose nos animais superiores¹²⁸.

Na sequência de experiências de Cleveland, muitas pesquisas foram dedicadas à demonstração da produção de celulases simbióticas por bactérias e fungos, além de protozoários, que são abrigados em sistemas digestivos de invertebrados. Contudo, poucos relatórios foram apresentados demonstrando as relações simbióticas entre os invertebrados e os microrganismos⁷⁰.

Dois artigos, em 1978, no entanto, reportaram que cupins Macrotermes spp adquiriram exoglucanases através da ectosimbiose com o fungo Termitomyces spp. Seguindo este raciocínio, uma hipótese similar de aquisição de enzimas digestivas, incluindo celulases, a partir de fungos decompositores de madeira, foi reportado em vespas de madeira¹²⁹. Posteriormente, esse conceito de aquisição enzimática foi

expandido para espécies xilófagas e detritívoras, incluindo besouros, isópodos, e formigas¹³⁰.

Foi proposto que a aquisição enzimática foi um fato geral no curso da coevolução microrganismo-inseto¹³¹, através de transferência horizontal de gene de procariontes^{114,131-135}. Essa afirmativa é corroborada através da análise filogenética da sequência de aminoácidos de nematoides parasitas de plantas, que conduziram a hipótese de transferência horizontal de gene a partir de bactérias¹³⁶. Todavia tais especulações acerca desse processo evolutivo são, mesmo hoje, incertas.

Além dos artrópodes, vários estudos com celulases foram empreendidos em moluscos¹³⁷⁻¹⁴¹, incluindo caramujos, lesma marinha¹⁴², caracóis, gastrópodes comestíveis e alguns bivalves¹⁴³⁻¹⁴⁹; além de ouriços-do-mar e ascídias^{133,150}.

São reportados que a possível fonte de enzimas endógenas pode ser de hepatopâncreas^{135,142}, dentes gástricos¹⁵¹ e estilos cristalinos^{143,148,152}.

A alternativa para a presença de celulases nestes diversos animais é que eles teriam derivado a partir de genes de um eucarioto ancestral antigo e persistiriam apenas em algumas linhagens de metazoários (transferência vertical)^{119,136,153-155}.

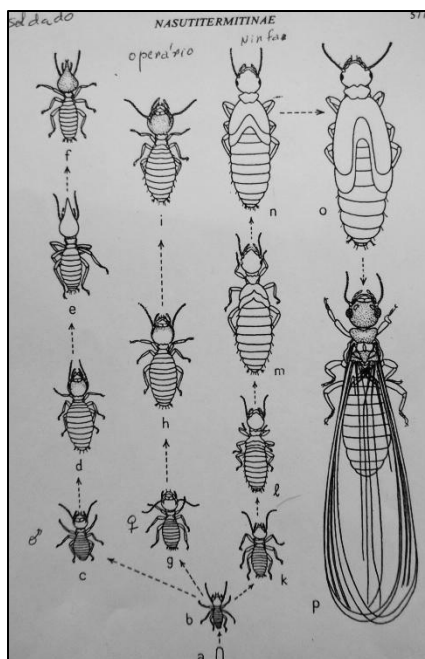
A explicação evolutiva para a presença de genes para celulases nestes diversos animais permanece obscura principalmente devido à precariedade de dados para realização de uma análise filogenética^{114,b}. Além disso, a maioria dos estudos tem focado nas características biológicas moleculares e na bioquímica dos seus produtos. Outro fator que remete a dificuldade para a formação evolutiva destes genes é a raridade com que aparecem entre os animais¹⁵⁵.

b. A análise filogenética para testar um origem antiga, com estes genes está comprometida pela falta de dados. Mesmo no caso de GHF5 e genes GHF6, a resolução filogenética é bastante pobre, presumivelmente porque os genes são curtos e saturada de substituição (LO, WATANABE, e SUGIMURA, 2003; MATTHYSSE *et al.*, 2004; NAKASHIMA *et al.*, 2004). No entanto, a família GHF 9 (que inclui beta- 1,4-endoglucanases) é excepcional, porque a sequência genética do núcleo é ao mesmo tempo relativamente longo (mais de 430 aminoácidos) e conservadas. Nesse sentido, torna-se primordialmente necessário um estudo mais detalhado da origem da aquisição destes genes no processo evolutivo; bem como um estudo mais amplo a cerca da biologia molecular e, porque não, estrutural destas celulases a fim de podermos descrever uma árvore filogenética.

Assim, antes de concluir por qual modelo tais genes foram adquiridos, faz-se necessário investigar com rigor as provas, de preferência usando uma abordagem gene-a-gene¹¹⁴.

2.10 Cupins e celulases endógenas

Pertencentes à Infraordem Isoptera (Ordem Blattaria), cuja qual detém cerca de 2370 espécies descritas no mundo, sendo aproximadamente 290 espécies (em 67 gêneros) registradas no Brasil, os cupins ou termita, insetos hemimetábolos, vivem em colônias altamente organizadas, onde o princípio básico é a sobrevivência desta e não do indivíduo. Se separarmos os indivíduos da colônia, eles morrem^{156,157}. Trata-se, pois, de insetos sociais onde os indivíduos, morfologicamente distintos, encontram-se agrupados em castas (Figura 16), sendo as principais definidas por Imagos ou Reprodutores (únicos responsáveis pela produção da prole); Soldados (realizam o papel de defesa da colônia, possuindo, dessa forma, adaptações morfológicas especiais, como fortes mandíbulas e cabeças esclerotizadas); e operários (são responsáveis pelas atividades de forrageamento, construção, reparo do ninho, alimentação das castas dependentes – jovens, soldados e reprodutores –, e cuidado com os ovos e cria¹⁵⁶⁻¹⁵⁸); existindo também ninfas (que dão origem aos reprodutores aladas), e larvas (que originam os operários e soldados, com asas totalmente ausentes), além dos ovos¹⁵⁷.



**Figura 16: Diferenciação morfológica das castas em cupins – Referência :
Nasutitermitinae.**

As imagens esquematizam, de maneira geral, a diferença morfológica entre rei/rainha, soldados, operários, ninfas e larvas; além de mostrar os ovos.

Apesar de definidas as castas, a diferenciação destas é um processo bastante complexo e não é bem conhecida entre os cupins. As castas são determinadas durante o desenvolvimento pós-embrionário e apesar dos ovos dos cupins terem capacidade de se diferenciar em qualquer casta (ovos totipotentes), cada indivíduo expressa apenas um único fenótipo que é afetado tanto por fatores extrínsecos, ou seja, pelas condições ambientais, quanto pelas interações sociais (intrínsecos)¹⁶⁰.

Aceita-se que a formação de castas nos cupins resulta da atividade de genes particulares, comuns a todos os indivíduos e que são expressos sob condições hormonais apropriadas, como citando em Tsuchiya, Watanabe e Maekawa (2008)¹⁶¹. Entretanto, recentes descobertas envolvendo a espécie Reticulitermes speratus, mostram que castas podem ser controladas tanto pelo ambiente, como por um complexo padrão de herança genética ligada ao sexo¹⁶². Assim, na referida espécie, depois do ovo, a

direção do desenvolvimento para ninfa ou operário é influenciada pelo genótipo da casta. Esta determinação genética de castas ligada ao sexo não é universal entre os cupins, já que este modelo não explica a origem de todas as castas, em outras espécies¹⁶³.

Estudos que visam hipóteses para evolução destes insetos também vêm sendo cogitadas em meio científico, como citado por Roux e Korb (2004)¹⁶⁴ e por Thorne, Breisch e Muscedere (2003)¹⁶⁵, além de Noirot e Pasteels (1987)¹⁶⁶; Noirot (1990)¹⁶⁷; Korb e Kantrantzis (2004)¹⁶⁸; Abe (1987)¹⁶⁹ dentre outros.

Em adição, os cupins, como insetos sociais, representam particularmente problemas interessantes para estudos de evolução de controle de doenças: frequentemente vivem em ambientes altamente povoados e em locais confinados (ninhos homeostáticos) que, em princípio, promoveria o crescimento de microrganismos, facilitando a transmissão de infecções para o ninho¹⁷⁰⁻¹⁷⁸. Apesar disso, os cupins, como clado, são extraordinariamente abundantes e diversos e causam grandes impactos ecológicos¹⁷⁹.

Nesse sentido, por desempenhar papéis ecológicos fundamentais nos ecossistemas naturais, alimentando-se basicamente de materiais de origem vegetal pobre (madeira, grama, serapilheira, algas, líquens e vegetais em decomposição, por exemplo), podendo ser estes herbívoros ou decompositores; são importantíssimos na reciclagem dos nutrientes acumulados nos tecidos vegetais. Atuam na aeração e drenagem do solo, bem como transporte de nutrientes entre os seus perfis e participam ativamente no processo de gênese de alguns tipos de solo^{180,181}.

Assim, os cupins são conhecidos por serem insetos comedores de madeira e por possuírem a alta habilidade de degradar materiais celulósicos. A madeira é inicialmente mastigada pelas mandíbulas e moela e então degradada por celulasas no trato digestório

¹⁸². Além disso, mantém complexas relações ecológicas (competição, simbiose, predação, parasitismo, comensalismo, etc) com diversas espécies de organismos, como citando por Warnecke (2007)¹⁸³ e por Sharf e Tartar (2008)¹⁸⁴. A celulose é decomposta através de dois sistemas celulósicos independentes: celulasas advindas do próprio cupim e celulasas de Protozoários/Bactérias simbiotes presentes no intestino grosso em cupins basais e na minoria dos cupins superiores^{69,132,181, 185-189}.

Apesar de muitas tentativas em isolar bactérias celulolíticas de cupins em meados dos anos 80, nenhuma evidência convincente da digestão de celulose por bactérias em cupins foi provada. Posteriormente, a produção de celulase bacteriana nos cupins livres de flagelados intestinais Nasititermes exitiosus e Nasititermes walkeri foi excluída pela demonstração que a celulase foi liberada a partir do intestino posterior e não por sonicação nem por tratamento lítico bacteriano (HOGAN et al, 1988, tradução do autor)¹⁸⁰.

Com exceção dos Termitidae (família derivada), os membros das demais famílias (basais) de cupins necessitam dos protozoários simbiotes presentes em seu intestino posterior para que a celulose seja degrada^{70,127,132,190-194}. Assim nos cupins superiores, pertencentes à família Termitidae, as partículas de madeiras ingeridas devem passar somente pelo sistema celulósico endógeno do intestino, não havendo necessidade de simbiose protozoário/cupim.

Em adição a bioquímica das enzimas celulósicas citada, as funções físicas e fisiológicas do canal alimentar também afetam, significativamente, o processo celulósico em insetos. Isso porque, segundo Watanabe e Tokuda (2010)⁷³, no curso da adaptação para a digestão simbiótica de celulose no intestino, um aumento no volume do lúmen ocorreu de modo frequente. Nesse contexto, a redução corporal e do tamanho da cavidade onde a digestão de celulose por celulasas endógenas ocorrem, levaram a um aumento na concentração das enzimas endógenas. Como resultado, o fornecimento de açúcares redutores foi maior e mais rápido^{73,190}.

Dessa maneira, levando em consideração que 75 % das espécies de cupins não abrigam flagelados celulósicos em seu canal alimentar; temos que tal sistema (endógeno) seja a única fonte de glicose dos próprios cupins; tornando-o e a seus componentes de interesse geral para futuras pesquisas.

A eficiência do aparelho digestório dos cupins é de 74 a 99 % para a digestão de celulose e 65-87 % para hemicelulose¹⁹¹. Essa alta eficiência, sobretudo tocante à digestão celulósica, tem fascinado muitos pesquisadores durante quase um século. A ideia de imitar o sistema digestório destes insetos parece ser bastante antiga; todavia, aplicações industriais empregando seus sistemas celulósicos ainda não são realizadas¹⁹².

No Brasil, duas enzimas presentes na espécie Coptotermes gestroi (espécie subterrânea originária da Ásia): a endo- β -glicosidase e a β -glucanase, foram clonadas e expressas, de modo a imitar o que o cupim faz, mostrando potencial, num futuro próximo, de degradar rejeitos como a palha da cana de açúcar e como o bagaço para serem transformados em energia. Este estudo, realizado entre 2006 e 2010 nos Laboratórios de Hemoglobina e Genoma do Hemocentro e de Genômica e Expressão do IB (UNICAMP – Campinas, SP), foi coordenado por Gonçalo Amarante Guimarães Pereira¹⁹⁵.

Dentro deste contexto, identificar celulasas estáveis e viáveis para o processo de fermentação presente hoje no Brasil é uma importância impar para desenvolver tecnologia industrial adequada à produção de etanol de terceira geração. Isso porque se faz necessário a busca por enzimas que trabalhem nas condições de pH (cerca de 4 a 5) e temperatura (entre 30 a 35 °C) presentes nas dornas de fermentação; até então não descobertas^{5,69}.

Ao propor o uso de celulasas de cupins, espera-se encontrar tal(is) enzima(s); dado a semelhança entre a biologia do cupim no tocante a digestão da celulose e o

processo industrial na conversão de biomassa, como citado acima. Além disso, pesquisas demonstram que as celulases de cupins superiores tendem a ser mais estáveis⁵ e mais eficientes, além de possuir maior adaptabilidade na quebra de celulose – única no reino animal. A alta eficiência pode ser explicada porque os cupins superiores desempenham um papel mais significativo do que os cupins basais na manutenção do balanço de carbono na natureza^{67,193}.

Um organismo interessante para a busca dessas enzimas é o cupim da espécie Velocitermes heteropterus Silvestri (Termitidae: Nasutitermitinae) (Figura 17). De ocorrência comum no Cerrado brasileiro – em especial no Brasil central e Sudeste, e também em algumas savanas amazônicas^{196,197} –, constrói ninhos em montículo de forma piramidal¹⁵⁸ e subterrânea; mas às vezes parecem construir pequenas cartonadas de consistência mole¹⁹⁷. Estes indivíduos possuem pernas longas (sobretudo os soldados) e se movimentam com rapidez e a casta neutra desta espécie apresenta polimorfismo, com a ocorrência de soldados trimórficos (pequenos, médios e grandes) e operários dimórficos (pequenos e grandes). Geralmente, o polimorfismo na casta neutra está associado a uma divisão de trabalho entre os diferentes indivíduos, principalmente nas atividades de forrageamento^{158,201}.

A maioria das espécies vive em áreas de vegetação aberta, e se alimentam das folhas da serapilheira, forrageando a céu aberto, no período noturno¹⁹⁶⁻²⁰⁰. Como a alimentação desses cupins assemelha-se em composição com a palha encontrada na cana-de-açúcar e pelo fato de serem insetos que forrageiam no período noturno existe uma boa probabilidade que as celulases encontradas neste gênero tenham uma alta atividade a temperaturas mais baixas que as atualmente isoladas; uma vez que estes insetos não possuem controle de temperatura corporal, obedecendo a temperatura externa.

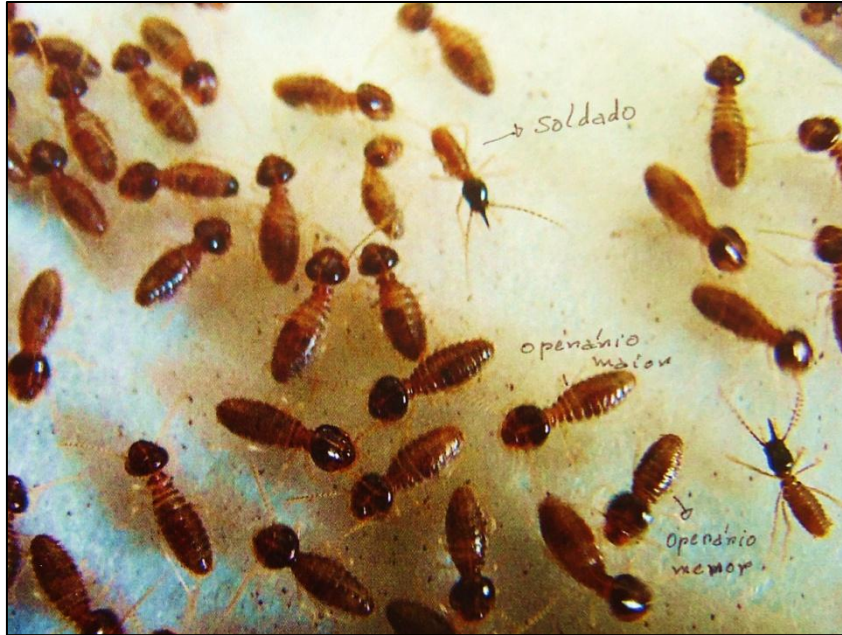


Figura 17: Velocitermes heteropterus - Termitidae: Nasutitermitinae.

A espécie é comum na região Central e Sudeste brasileira, particularmente em domínios de Cerrado. Na imagem veem-se soldados (os cupins menores com mandíbulas) e operários (cabeça mais protuberante).

3 Objetivo

O presente estudo teve como objetivos:

3.1 Objetivo Geral

Isolamento, Identificação e Expressão da celulase β -1,4-endoglucanase do cupim Velocitermes heteropterus em cepa de S. cerevisiae laboratorial.

3.2. Objetivos Específicos

1. Análise *in silico* visando à busca pela β -1,4-endoglucanase de V. heteropterus por meio de transcriptoma.
2. Isolamento do cDNA que codifica a celulase β -1,4-endoglucanase de cupim V. heteropterus por meio da extração de RNA total e PCR utilizando *primers* específicos desenhados com base no transcriptoma deste cupim.
3. Identificação desta celulase utilizando ferramentas de bioinformática para validação de sequência de nucleotídeos, aminoácidos, predição de estrutura terciária e Filogenética.
4. Construção de Plasmídeo de Expressão contendo a sequência para a celulase.
5. Expressão recombinante da enzima em S. cerevisiae FGSC 9721.
6. Ensaio da atividade proteica por meio de testes CMC, SDS-PAGE.

4 Material e Métodos

Para nossos experimentos, utilizamos os seguintes materiais e as seguintes técnicas a serem descritas nas seções seguintes:

4.1 Material biológico

Utilizamos os seguintes materiais biológicos a serem descritos:

4.1.1 Cupins

Os cupins da espécie V. heteropterus utilizados neste estudo, foram gentilmente cedidos pela Prof^a Dr^a Ana Maria Costa Leonardo da Universidade Estadual Paulista

“Júlio Mesquita” (UNESP) Campus de Rio Claro. Coletamos dois diferentes tipos de castas (soldados e operários).

4.1.2 Linhagens para clonagem e expressão

A linhagem de *E. coli* utilizadas nesse estudo foi a **DH5 α** (F⁻ Φ 80lacZ Δ M15 Δ (lacZYA-argF) U169 recA1 endA1 hsdR17 (rK⁻, mK⁺) phoA supE44 λ ⁻ thi-1 gyrA96 relA1)^c.

A linhagem de levedura utilizada foi a **FGSC 9721** (MAT α his3 Δ 200 ura3-52 leu2 Δ 1 lys2 Δ 202 trp1 Δ 63)^{d,203}.

4.2 Meios de cultura

- 1. LB (Luria-Bertani):** 10 % Triptona, 5 % Extrato de levedura e 5 % Cloreto de sódio, pH 7,0; complementados com antibióticos para seleção dos clones transformantes (ampicilina – 100 μ g/mL) ou com IPTG (200 U/ μ L) para indução da expressão da proteína. Para meio sólido é adicionado 10 % de ágar bacteriológico.

c. “lacZ Δ M15” permite uma seleção azul/branco para células recombinantes; “recA1” reduz recombinação homóloga; “endA1” reduz a digestão endonuclease de plasmídeo para maior rendimento; hsdE17 gera mutação no sistema de metilação e restrição que permitem que a *E. Coli* reconheça o DNA como estrangeiro; “supE44” tRNA supressor de glutamina; “thi-1” requer tiamina para crescimento em meio mínimo; “gyrA96” DNA girase mutante produz resistência ao ácido nalidíxico; “relA1” regula o acoplamento entre transcrição e tradução;

d. “his3”, “leu2”, “lys2” e “trp1” são marcas autotróficas. É derivada da linhagem S288C, cuja qual não contém mutações nessas regiões. O “MAT” reprime ou ativa transcrição de gene específico²⁰³.

2. **YPD (Levedura Peptonada Dextrose):** com 10 % Extrato de Levedura, 20% Glicose, 20 % Peptona, pH 6,5. Para meio solido, adicionou-se 20 % ágar bacteriológico.
3. **scTrip⁻** : 7 % YNB – Nitrogênio Levedura Base; 2% Glicose; 550 µM Lisina; 760 µM Leucina; 322 µM Histidina; 5,3 mM Uracila; 2,46 mM Uridina. Para meio sólido, acrescentou 20 % ágar bacteriológico.
4. **scLeu⁻** : 7 % YNB – Nitrogênio Levedura Base ; 2% Glicose; 550 µM Lisina; 2,46 mM Uridina; 322 µM Histidina; 486 µM Triptofano; 2,46 mM Uracila. Para meio sólido, acrescentou 20 % ágar bacteriológico.
5. **CMC ágar:** 1% CMC (Carboximetilcelulose), 2 % ágar.

4.3 Soluções

Ampicilina (10 mg/mL – estoque): 3 mM ampicilina;

Bis-acrilamida 30 %: 422 mM acrilamida; 5,2 mM N’N-metilenobisacrilamida

Coomassie coloidal: 600 mM Ácido Acético; 0,5 M Metanol; 242 µM Comassie blue R-250

Loading dye (6x): 700 mM glicerol; 373 µM Azul de bromofenol; 464 µM xilencianol

MOPS (10X – Solução estoque): 0,2 M MOPS; 20 mM Acetato de sódio, 10 mM EDTA

Solução I para minipreparação plasmidial : 10 mM EDTA pH 8,0; 25 mM Tris-HCl pH 8,0

Solução II para minipreparação plasmidial: 0,2 M NaOH; 3,47 mM SDS

Solução III para minipreparação plasmidial: 3M Acetato de Potássio; 115 mM Ácido acético glacial

Solução de Descorante de Coomassie: 100 mM de Ácido Acético; 0,3 M Etanol absoluto

TAE 1X (Tris-Acetato-EDTA 1X): 40 mM Tris acetato; 1 mM EDTA pH 8,0

TE 10X (solução estoque): 10 mM EDTA; 100 mM Tris

Tampão de Extração de DNA: 25 mM EDTA pH 8,0; 200 mM Tris-HCL pH 8,0; 5 mM SDS; 250 mM NaCl pH 8,0

Tampão de Extração Proteica: 50 mM Tris pH 7,5; 150 mM NaCl; 100 mM Glicerol; 0,1 % SDS; 15 mM EGTA pH 7,5; 10 mM triton x-100

Tampão de Laemilli: 89,5 mM Azul de bromofenol; 490 mM glicerol; 4,2 mM SDS; 600 mM Tris-HCl pH 6,8; 6 mM Ditriotreitol (DTT)

Tampão Tris-Glicina 1X: 25 mM tris-base; 192 mM Glicina; 1 mM SDS

4.4 Plasmídeos

pGemT-Easy (Promega®): Plasmídeo comercial utilizado para clonagem de produtos de PCR. Promotores T7 e SP6 e M13 flaqueiam região de Clonagem Múltipla. Possui região codificante lacZ, permitindo seleção por coloração diferencial; e resistência a ampicilina pela presença da marca de seleção *Amp^r*.

pYADE4: Plasmídeo contendo o promotor álcool desidrogenase 2 (ADH2), marca de seleção *Amp^r* (resistência a ampicilina) para bactérias e auxotrofia Trp1- (Triptofano).

4.5 Primers

Para a amplificação do fragmento da Celulase desejado; verificação de ligação em plasmídeo de clonagem e de expressão foram utilizados os seguintes *primers*:

Tabela 4: Sequências de primers utilizados e sua finalidade

Primers	Sequência	Função
β-actF	5' GTTGTCAAATGTAATCCATCC 3'	Amplifica região de conservada de β -actina (gene endógeno)
β-act R	5' CTGACGTACTGTGCGCTGT 3'	Amplifica região de conservada de β -actina (gene endógeno)
VEgF	5' ATC TCT <u>GGA TCC</u> CAG GAG GAT ACT ATG ACG 3'	Amplificar β -1,4-endoglucanase
VEgR	5' ACT TCT <u>CTC GAG</u> TGT TCT GTG AGT TTA ATA GCC 3'	Amplificar β -1,4-endoglucanase
M13F	5' GTA AAA CGA CGG CCA GT 3'	Flanqueia regiões adjacentes MCS do pGemT-Easy
M13R	5' GCG GAT AAC AAT TTC ACA CAG G 3'	Flanqueia regiões adjacentes MCS do pGemT-Easy
ADH2F	5' AAA GGG GCA AAA CGT AGG GGC 3'	Flanqueia regiões externa ao MCS do pYADE4
pYADE4R	5' ACT CCT TCC TTT TCG GTT AG 3'	Flanqueia regiões externa ao MCS do pYADE4

4.6 Extração de RNA total

Para a extração do RNA total, os cupins foram, inicialmente, pulverizados na presença de nitrogênio líquido e em seguida, adicionado 2 mL *Trizol*[®] (Invitrogen) para cada 50 cupins pulverizado. A extração ocorre segundo normas do fabricante.

A amostra homogeneizada foi posta em tubos de 1,5 mL e centrifugada por 12 minutos a 13.300 rpm à 4 ° C. Após, a fase aquosa obtida foram transferida para novos tubos e a estes acrescentam-se 200 μ L de clorofórmio; passando por posterior agitação. Os tubos foram centrifugados por 17 minutos a 13.300 rpm, 4 ° C.

O sobrenadante, contendo o RNA total, foi recuperado e transferido para novos tubos; acrescentando 200 μ L de fenol-clorofórmio-álcool isoamílico (25:24:1 v/v) ou

fenol-clorofórmio (1:1), sendo agitados e centrifugando por mais 17 minutos a 13.300 rpm, 4°C. Ao sobrenadante foram acrescentado 500 µL de isopropanol para a precipitação do RNA. O tubo foi agitado e centrifugado por 12 minutos a 13.300 rpm, 4°C. O sobrenadante foi descartado e, ao precipitado foram acrescentados 1 mL de etanol 75%, sendo agitado e centrifugado por 6 minutos a 4°C em rotação de 8.000 rpm. O precipitado foi ressuspendido em 20 µL água com dietilpirocarbonato (água com DEPC) e estocadas em freezer a - 80 ° C.

O critério de integridade do RNA foi à visualização intacta das bandas de RNA ribossomal em gel desnaturante 1,2 %.

4.7 Síntese de DNA complementar (cDNA)

Para a síntese do cDNA utilizamos o *kit Super Script II Reverse transcriptase* (Sigma®), seguindo instruções do fabricante. Resumidamente, a reação ocorre utilizando 1ng a 5ng de RNA total e pela adição de 1µL de oligo (dt)₁₂₋₁₈ *Primer*, 1 µL de dNTP Mix 10 mM e água para completar 12 µL. Esta reação é, então aquecida a 65 °C por cinco minutos e em seguida resfriada em gelo rapidamente.

Em seguida, foram adicionados 4 µL de tampão 5X First-Strand, 2 µL de dtt 0,1 M e 1 µL de RNase Out (40U/ µL). Incubou-se por dois minutos a 42 °C e adicionou 1 µL de Super Script II RT (200U). A reação foi incubada por 50 minutos à 42 °C e inativada à 70 °C por cinco minutos.

Para a preparação da segunda fita, adicionou 1 µL de RNase H (2U) e incubou à 37 °C por 20 minutos. As amostras foram armazenadas em freezer – 80 °C ou -20 °C quando utilizadas em um curto espaço de tempo.

4.8 Verificação da síntese de DNA complementar

A verificação da síntese de cDNA foi feita através da reação em cadeia da polimerase (PCR) para o gene da beta-actina (β -actina). Para esses experimentos utilizamos a sequência descrita para o gene beta actina de Coptotermes gestroi. As reações de PCR foram realizadas com: 1,5 μ L de tampão para PCR 10X; 0,5 μ L de dNTPs 10mM; 0,75 μ L de $MgCl_2$ 50 mM; 1 μ L para cada um dos *primers* β -actF e β -actF R a 10 pmol; 0,2 μ L de Taq DNA polimerase 5U/ μ L - Fermentas[®]; 0,5 μ L de cDNA; 9,85 μ L de água estéril. O programa é iniciado por 2 minutos à 94°C, seguido de 40 ciclos (94 °C por 15 segundos, 55 °C por 20 segundos e 72 °C por 2 minuto). Os produtos da PCR são submetidos à eletroforese em gel de agarose 1,5% para verificar a amplificação de um fragmento de aproximadamente 100 pares de base (pb) (tamanho esperado do amplificado).

4.9 Construção dos primers para amplificação do fragmento que codifica a β -1,4-endoglucanase

Para a amplificação do fragmento que codifica a celulase (β -1,4-endoglucanase) foram desenhados *primers* (VEgF e VEgR) com base na sequência para β -1,4-endoglucanase encontrada no transcriptoma^e (Apêndice) de V. heteropterus e comparadas contra banco de dados para confirmação (veja seção 4.5).

e. O transcriptoma visa o reconhecimento de diversos genes importantes para o entendimento da biologia deste cupim, todavia, para fins iniciais, foram reconhecidos apenas celulasas. Análise deste transcriptoma encontra-se em andamento.

Aos *primers* foram acrescentadas regiões para sítios enzimáticos (sublinhados) onde as enzimas BamHI e XhoI poderiam ser utilizadas nas regiões *forward* e *reverse*, respectivamente. Essa estratégia foi feita visando a posterior inserção do fragmento em plasmídeo de expressão pYADE4.

O fragmento de amplificação esperado para esta combinação de *primers* deveria ser de aproximadamente 1200 pbs. Neste caso, a construção de um transcriptoma foi necessário haja visto que este cupim não possui qualquer descrição molecular em literatura.

4.10 Amplificação do cDNA que codifica a β -1,4-endoglucanase de V. heteropterus

O fragmento que codifica a β -1,4-endoglucanase de V. heteropterus foi amplificado utilizando a *taq Phusion DNA polymerase hot start*, uma enzima que se caracteriza por ter alta fidelidade, tal que sua taxa de erro é 50 vezes menor que a da *Taq DNA polymerase*. Além de possuir atividade polimerase 5'→3', também possui atividade exonuclease 3'→5'; garantindo que o *amplicon* não tenha bases mutadas ou ausentes. Isso garante que a proteína a ser codificada a partir do *amplicon* seja o mais próximo possível da proteína expressa pelo organismo de origem.

Assim, para a reação é utilizado 2 μ L de cDNA (transcrito a partir de 1 μ g de RNA); 4,0 μ L de 5X *Phusion*[®] *HF buffer* (Finnzymes *Phusion*[®]); 0,4 μ L de dNTPs mix 10 mM; 1 μ L do *primer* VEGF; 1 μ L do *primer* VEGR; 0,2 μ L de *Phusion DNA polymerase hot start* (2u/ μ L) (Thermo Scientific – Finnzymes PCR[®]). Os fragmentos são amplificados em Termociclador automático C1000 Touch[™] Thermal Cycler - BioRad[®]) e seguem o seguinte ciclo: 95°C por 4 minutos; 40 x (95 °C por 1 minuto; 55

°C por 1 minuto; 72 °C por 1,5 minuto); 72 °C por 10 minutos. A verificação da amplificação é feita por eletroforese em gel de agarose 0,7 % com de corrida tampão TAE 1X; 100 v por 30 minutos.

4.11 Clonagem do fragmento que codifica β -1,4-endoglucanase de V. heteropterus em vetor pGemT-Easy

O fragmento amplificado foi clonado em vetor *pGemT-Easy* (Promega®), seguindo as instruções do fabricante. Para tanto, inicialmente, o fragmento amplificado foi tratado com adenina para clonagem em pGemT-Easy. Após a ligação, a construção foi inserida em células de *E. coli* DH5 α eletrocompetente utilizando a técnica de eletroporação (1,8 KVA). Os clones positivos foram selecionados em meio LB sólido contendo ampicilina (100 μ g/ μ L), x-gal (20 μ L/mL) e IPTG (200 mM). Isso porque o *pGemT-Easy* contem região *lacZ*, que proporciona a seleção pela coloração das colônias (branco/azul). Após o crescimento foram selecionadas 20 colônias para PCR de colônia.

A PCR foi feita da seguinte maneira: 2 μ L de tampão para PCR 10X, 0,4 μ L de dNTP 10 mM, 0,6 μ L de MgCl₂ 50 mM; 0,4 μ L de cada *primer* M13F e R; 0,1 μ L taq DNA polimerase (5U/ μ L); 2 μ L do inóculo aquecido. O ciclo foi: 95 °C por três minutos, 25x(94 °C por 30 segundos, 60 °C por 30 segundos e 72 °C por um minuto). A amplificação foi visualizada em gel de agarose 1 % usando tampão de corrida TAE 1X. Das colônias positivas, foi extraído DNA plasmidial.

4.12 Minipreparação de plasmídeo pGemT-Easy+EG e pYADE4+EG

As colônias positivas foram inoculadas em meio LB líquido e crescidas durante 16 horas sob agitação constante a 37 °C e em seguida o DNA plasmidial foi extraído utilizando o protocolo como realizado por Sambrook (2000)²⁰⁴: a cultura foi centrifugada à 13.000 rpm por 2 minutos e ao precipitado foram adicionados 100 µL da Solução I e agitado. Acrescentou-se, posteriormente 200 µL da Solução II preparada no momento do uso. A mistura agitada novamente, permanecendo em gelo por 10 minutos e, posteriormente, acrescentou-se 150 µL da Solução III. Uma nova centrifugação à 13.000 rpm ocorre por 5 minutos.

Ao sobrenadante foi acrescido fenol-clorofórmio (1:1 v/v); centrifugando por 5 minutos à 12.000 rpm. O precipitado foi descartado e ao sobrenadante adicionou 1 mL de etanol absoluto gelado para precipitação do DNA; centrifugando à 12.000 rpm por 10 minutos, 4 °C. O precipitado foi lavado com etanol 70 % e centrifugado novamente. O precipitado foi eluído em 30 µL de água MilliQ.

4.13 Sequenciamento dos clone recombinantes

O sequenciamento foi realizado pela empresa Macrogen Inc. (<http://www.macrogen.com/eng/>)²⁰⁵ pelo método de Sanger, através da utilização do serviço *Standard Sequencing*; e pela empresa QGene (São Carlos, SP) através do equipamento MEGABaceTm 1000. As construções pEG e yEG foram enviadas para ambas as empresas bem como os *primers* ADH2F, pYADE4R, VEgF, VEgR; M13F e M13R.

4.14 Bioinformática e análise dos dados

As sequências geradas foi comparada utilizando a ferramenta BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) contra banco de dados do GenBank ²⁰⁶ para a verificação e alinhadas com a sequência encontrada no transcriptoma de *Velocitermes* sp. Testes filogenéticos com base nesta sequência também foram realizados fazendo uso do programa MEGA 5.05²⁰⁷. Com base nas sequências (do transcriptoma e plasmídeos sequenciados), foram realizadas predição da estrutura terciária desta proteína utilizando programa Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>)²⁰⁸.

*4.15 Expressão heteróloga da β -1,4-endoglucanase em levedura *S. cerevisiae* FGSC 9721*

Para a construção do vetor de expressão, os plasmídeo pGemT-Easy+EG, bem como o vetor pYADE4 foram digeridos com as enzimas XhoI e BamHI *fast digest* (Fermentas[®]) por 30 minutos. Os fragmentos correspondentes ao inserto (denominado EG) e ao plasmídeo pYADE4 digeridos foram purificados do gel de agarose 0,7 % utilizando kit *GeneJet Gel Extration Kit* (Fermentas[®]) , seguindo as normas do fabricante.

Posteriormente, a ligação entre o pYADE4 e o inserto EG foi realizada utilizando T4 DNA ligase (200 U) e 10 X T4 DNA ligase *buffer* (Fermentas[®]). Essa reação foi deixada à 4 °C por cerca de 16 horas. A transformação de ligação foi inserida em bactérias DH5 α através de eletroporação; e a solução, plaqueada em meio LB contendo ampicilina (100 μ L/mL). As placas foram incubas à 37 °C por cerca de 24 horas.

Das colônias, selecionou-se cerca de 15 colônias para a realização de PCR de colônia, seguindo protocolo: 2 µL de tampão para PCR 10X, 0,4 µL de dNTP 10 mM, 0,6 µL de MgCl₂ 50 mM; 0,4 µL do *primer* ADH2F e 0,4 µL do *primer* pYADE4R; 0,1 µL taq DNA polimerase (5U/ µL); 2 µL do inoculo aquecido. O ciclo foi: 95 °C por três minutos, 25x(94 °C por 30 segundos, 55 °C por 30 segundos e 72 °C por 1,5 minutos). A amplificação foi visualizada em gel de agarose 1 % usando tampão de corrida TAE 1X. Das colônias positivas, foi extraído DNA plasmidial utilizando uma Midipreparação (ver item 4.15.1).

A estratégia utilizada para a clonagem consistiu na tentativa de se clonar o fragmento no vetor utilizando extremidades *stick-end*, uma vez que o pGEMT-Easy não possui regiões para tais enzimas; em detrimento do pYADE4. Portanto, tais regiões foram previamente desenhadas nos respectivos *primers*.

4.15.1 Midipreparação de plasmídeos pGemT-Easy+EG e pYADE4+EG

As colônias positivas foram inoculadas em meio LB líquido e crescidas durante 16 horas sob agitação constante a 37 °C e em seguida o DNA plasmidial foi extraído utilizando o seguinte protocolo realizado pelo Sambrook (2000)²⁰⁴: a cultura foi centrifugada à 4.000 rpm por 10 minutos e ao precipitado foram adicionados 200 µL da Solução I e agitado. Acrescentou-se, posteriormente 400 µL da Solução II preparada no momento do uso. A mistura, então, agitada novamente, permaneceu em temperatura ambiente por 5 minutos e, posteriormente, acrescentou-se 300 µL da Solução III. Uma nova centrifugação a 13.000 rpm ocorreu por 10 minutos, 4 °C. Ao sobrenadante, foram adicionados fenol:clorofórmio (1:1 v/v) e uma agitação vigorosa foi realizada.

Essa solução, então, foi centrifugada novamente. O precipitado foi descartado e ao sobrenadante, adicionado 500 μ L de isopropanol, sendo também centrifugado.

Após isso, adicionou etanol 70% passando por nova centrifugação. Essa etapa foi realizada até que todo o fenol fosse retirado. Uma última etapa foi feita utilizando etanol 100%. Ao precipitado seco foram adicionados 100 μ L de água estéril.

4.15.2 Transformação em linhagem de levedura S. cerevisiae PGSC 9721

A linhagem PGSC 9721 de S. cerevisiae foi previamente inoculada em YPD e incubada a 30 °C por 24 horas. Deste pré-inóculo, cerca de 500 μ L foram inoculados em 50 mL de YPD e deixados sob agitação constante a 30 °C até a OD₆₀₀ de aproximadamente 0,6.

Essa cultura, então, foi centrifugada por 2 minutos a 1000 xg e ao precipitado foi adicionado 20 mL de TE 1X, pH 7,5. As células foram decantadas por centrifugação e ao precipitado, acrescentou 800 μ L de solução TE1X/Acetato de Lítio 1X. Incubou-se por 1 hora a 30°C sob agitação moderada.

Posteriormente, alíquotas de 200 μ L foram postas em tubos de 1,5 mL; de modo que, em dois tubos foram acrescentados 10 μ L da construção pYADE-EG; em um tubo somente o plasmídeo pYADE4 e no outro somente havia a linhagem sc9721. Em cada tubo foram adicionados, também, 20 μ L de Esperma de Salmão (10 mg/mL). Esses tubos foram deixados sob agitação moderada por 30 minutos a 30 °C.

Após decorrido tempo, foram adicionados 1,5 mL de solução Acetato de Lítio 1X/ PEG 3350 40% em cada tubo, sendo misturados por inversão e incubados a 42 °C por 30 minutos em banho úmido. Os tubos foram centrifugados por 1 minuto a 2500 xg

e ao precipitado adicionado 1 mL de TE 1X. Outra centrifugação foi realizada e o precipitado ressuspensionado em 100 µL de TE 1X.

O eluído foi plaqueado em meio scTrip⁻, uma vez que o vetor pYADE4 possui marca auxotrófica para triptofano; sendo portanto, um modo de seleção para aquisição de colônias positivas. As placas foram incubadas a 30 °C por 72 horas.

Como a linhagem utilizada possui marcas de auxotrofia para os aminoácidos Leucina e Triptofano (além de Uracila e Histidina), colônias aleatórias foram transferidas para placas contendo meio scLeu⁻, YPD e scTrip⁻, a fim de se fazer uma melhor triagem de possíveis colônias positivas. Dessa forma, colônias positivas deveriam crescer apenas se tivessem incorporado o plasmídeo nas placas Trip⁻ (marca constante no plasmídeo construído) enquanto não deveriam crescer na ausência de Leucina (placas scLeu⁻). A partir desta pré-seleção, foram selecionadas linhagens para os testes a seguir.

4.16 Produção de proteína recombinante

A extração da proteína das linhagens selecionadas se procedeu da seguinte forma: um pré-inóculo em scTrp⁻ foi deixado durante 18 horas a 30 °C sob agitação constante. Em seguida, este foi centrifugado e ao precipitado de células foram adicionados tampão de extração proteica juntamente com Complete Mini e pmSF (inibidores de proteases). A solução foi adicionada pérolas de vidro e postas em agitação por cerca de 15 minutos. Após, a solução aquosa foi transferida para novo tubo e centrifugada por 30 minutos. O extrato proteico (sobrenadante) foi armazenado em freezer -80 °C.

O extrato proteico da linhagem selecionada foi corrido em gel de poliacrilamida 12%, pelo método de SDS-PAGE, a 150 V por cerca de 1,5 horas; e corados com Coomassie *blue*.

4.17 Teste de atividade em meio CMC

Para os testes de CMC, o mesmo protocolo de extração proteica (veja seção 4.16) foi usado para as linhagens selecionadas, de modo que cerca de 20 μ L foram utilizados para o teste e as placas foram deixar por cerca de 24 a 26 horas a 30 °C. Após este período, o meio foi corado com vermelho de Congo 28,7 mM por 30 minutos e lavado com NaCl 1M. A aparição de um halo indica a presença de celulase. Isso porque o corante se agrega as moléculas do meio de cultura. Caso a proteína metabolize o CMC, a agregação fica dificultada e assim, há a formação do halo. O extrato proteico que possibilitou um halo maior, então foi escolhido para os experimentos seguintes.

4.18 Teste comparativo e nível proteico

O teste comparativo consistiu em traçar curvas de crescimento e viabilidade, bem como consumo de glicose entre a linhagem que apresentou melhor expressão em meio CMC e SDS-PAGE com a linhagem selvagem PE-2 e linhagem controle.

Nesse tocante, foram feitos pré-inóculos de ambas as cepas, deixado por 18 horas a 30 °C sobre agitação. Em seguida, este pré-inóculo foi colocado em meio YPD contendo 4 % de glicose, com uma OD₆₀₀ de aproximadamente 1,0. Alíquotas foram retiradas a cada 2 horas. Destas, foram feitos testes de viabilidade e crescimento, bem como consumo de glicose utilizando o kit Glicose (Glicose Teste Enzimático LarborLab[®]).

4.19 Extração de DNA plasmidial de levedura para verificação da positividade da transformação

A cepa selecionada foi inoculada em $scTrp^-$ e incubada a 30 °C por 16 horas em agitação constante. Posteriormente, foi centrifugada por 5 minutos a 4.000 rpm e o precipitado ressuspendido em 500 μ L de tampão de extração de DNA. Acrescentou-se perolas de vidro a fim de que a parede celular fosse quebrada durante a agitação por 10 minutos. A solução aquosa foi posta em tubo de 1,5 mL e acrescentou-se 400 μ L de solução de fenol:clorofórmio (1:1 v/v), sendo novamente agitado por 10 minutos. Posteriormente, retirou-se a fase aquosa acima da linha de proteínas, e a esta foi adicionada 300 μ L de isopropanol. Essa etapa foi deixada em freezer -80 °C por 30 minutos e posteriormente centrifugada. Ao precipitado, foram acrescentados 300 μ L de etanol 70 % e novamente centrifugado. A eluição foi feita em 80 μ L de água estéril. A armazenagem foi feita em freezer – 20 °C.

Essa técnica normalmente é utilizada para a extração de DNA genômico; no entanto, se existir também a presença de DNA plasmidial, este também será extraído conjuntamente. Para verificar se o plasmídeo ainda se encontrava presente na levedura selecionada, o eluato da extração foi utilizado para transformação de bactérias eletrocompetentes DH5 α fazendo uso da mesma metodologia citada na seção 4.11. A presença de transformantes; bem como sua verificação por meio da PCR do gene heteróloga, evidencia a presença da construção plasmidial. Para a verificação de que a sequência não sofreu mutações, o plasmídeo extraído dessas bactérias também foi sequenciado, utilizando a metodologia da seção 4.12 e 4.14.

5 Resultados e Discussões

A partir dos protocolos utilizados, obtivemos os seguintes resultados que serão discutidos a seguir:

5.1 Extração de RNA

Os cupins, sendo insetos eusociais se dividem em castas, onde os indivíduos se apresentam morfologicamente distintos¹⁸¹. No presente estudo, trabalhamos com a casta de operários; onde foi isolado o RNA total cuja integridade foi verificada por meio de gel Desnaturante. A visualização das bandas de RNA ribossomais 28S e 18S evidenciam uma extração de excelente qualidade (Figura 18). A concentração (ng/ μ L) deste RNA foi feita por leitura em espectrofotômetro *Nanovue* (GE Healthcare), atingindo 2469,2 ng/ μ L e as contaminações foram comparadas através da razão entre a leitura a 260 com a leitura a 280; cuja qual teve relação de 1,81 indicando uma boa qualidade deste RNA (Razão ideal -2).

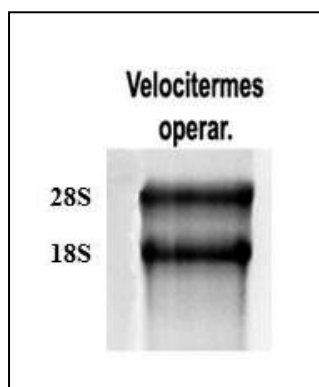


Figura 18: Integridade do RNA total extraído de Operários de *V. heteropterus* sp.

Nota-se a presença das bandas de RNA ribossomal (28S e 18S) íntegras evidenciando a boa qualidade da extração. Por se tratar de gel desnaturante e pela adição de brometo ser diretamente na amostra e não durante a produção do gel, o marcador de peso molecular não é visível.

5.2 Síntese de cDNA e Verificação da eficiência da transcrição

Para a comprovação da eficiência da transcrição, foi realizada uma reação de polimerase em cadeia (PCR) para a amplificação do gene endógeno da β -actina; Conforme evidenciado na figura 19, obtivemos a simplificação de um fragmento de aproximadamente 100 pb, como esperado; comprovando o sucesso da síntese do cDNA. A utilização deste gene endógeno como controle da eficiência se deve ao fato deste ser constitutivo e expresso em todas as células do organismo, sendo essencial; sendo necessariamente detectável após a transcrição reversa.

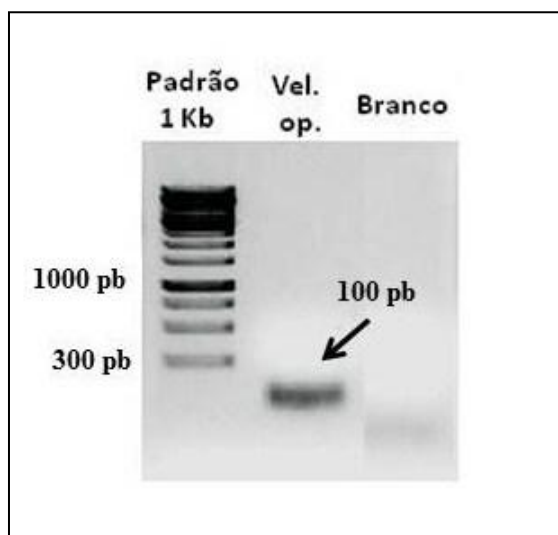


Figura 19: Verificação da síntese do DNA complementar com o uso de primer β -actina.

A amplificação de um fragmento de aproximadamente 100 pb na amostra de cDNA de operários de *V. heteropterus* mostra a eficiência da síntese de cDNA. Como controle negativo de amplificação foi utilizado um branco, no qual todos os reagentes, exceto cDNA, foram adicionados a reação. Conforme esperado, não houve a amplificação de nenhum fragmento. O padrão de peso molecular utilizado foi de 1Kb.

5.3 Amplificação do fragmento que amplifica a celulase β -1,4-endoglucanase

Como objetivo de amplificar o gene da β -1,4-endoglucanase, o cDNA obtido foi utilizado como molde em uma reação de PCR, conforme descrito na seção 4.10. A

amplificação gerou um fragmento de aproximadamente 1200 pb. (Figura 20); condizente com o esperado segundo sequência encontrada no transcriptoma dessa espécie (Apêndice).

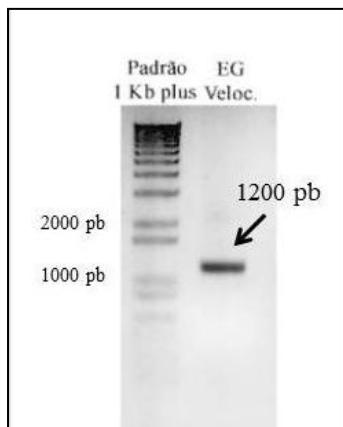


Figura 20: Amplificação do cDNA de β -1,4-endoglucanase de operários de V. heteropterus.

A amplificação utilizando *primers* específicos VEgF e VEgR resultou em um fragmento de aproximadamente 1200 pb; condizente com o esperado segundo o transcriptoma desta espécie.

5.4 Clonagem de β -1,4-endoglucanase de V. heteropterus no plasmídeo pGemT-Easy

O fragmento amplificado foi clonado em vetor pGemT-Easy segundo as normas do fabricante. A construção, aqui denominada de pGemT-Easy+EG, (Figura 21) foi inserida em E. Coli DH5 α eletrocompetentes seguindo protocolo padrão, e plaqueadas em meio LB com ampicilina. Colônias supostamente positivas passaram por PCR de colônia para confirmação e, posteriormente, foi realizada extração de DNA plasmidial.

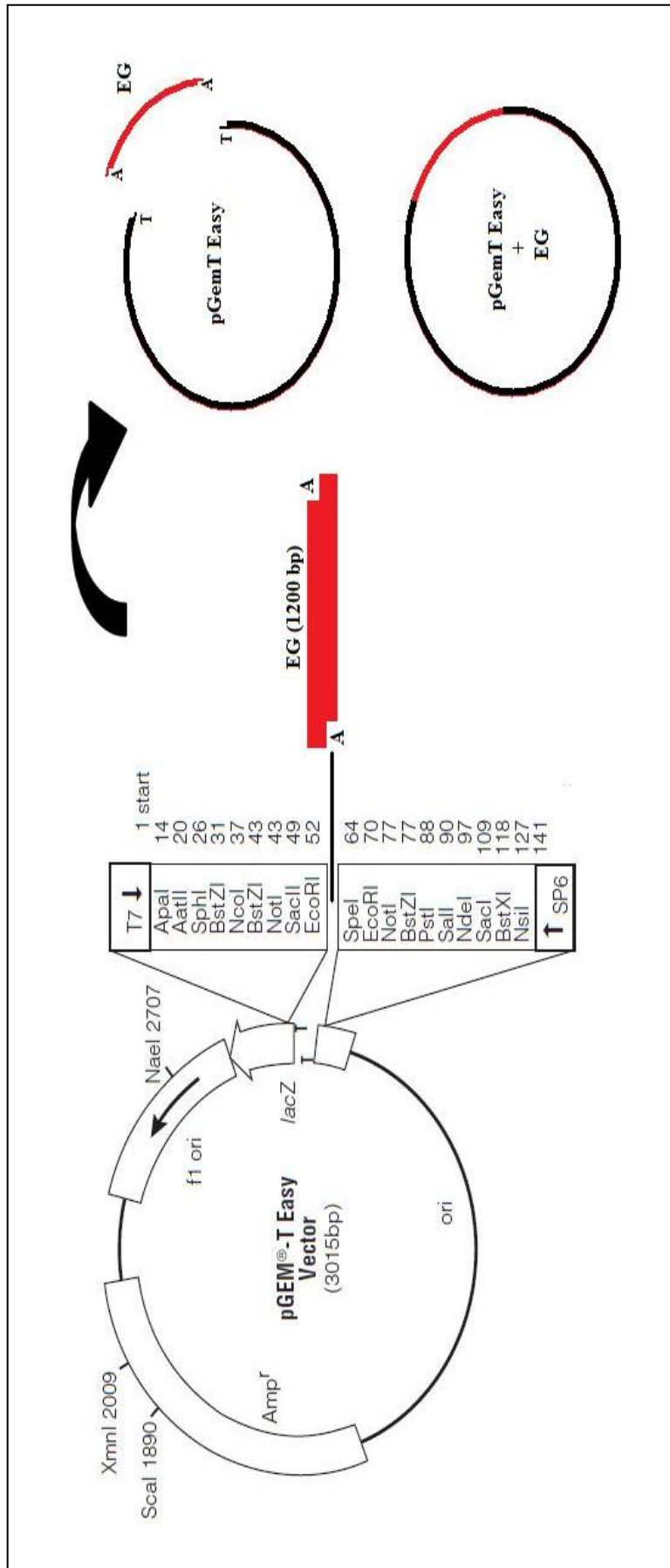


Figura 21: Esquemática da Construção do pGemT-Easy+EG.

A construção consistiu na ligação entre o pGemT-Easy (3015 bp) e o fragmento EG de 1200 bp.

5.5 PCR de colônia para verificação de clones positivos e Minipreparação de plasmídeos

A PCR das colônias positivas, descrita na seção 4.11 permitiu uma confirmação inicial de que realmente estas colônias possuíam o pGemT-Easy+EG. Dentre as quatro colônias, obtivemos uma colônia com amplificação de 1200 pb (Figura 22). A utilização dos *primers* M13 *forward* e *reverse* permitiu a confirmação da ligação do inserto (iEG) ao vetor de clonagem (pGemT-Easy), já que estes *primers* flanqueiam a região do MCS.

Desta colônia foi feito extrato plasmidial e seu posterior tratamento com RNase (Figura 23) visando a digestão enzimática. O tratamento garante que apenas o DNA plasmidial esteja íntegro; não havendo, portanto, RNA que pode vir atrapalhar a posterior digestão. O mesmo procedimento foi realizado para o vetor pYADE4 (Figura 24).

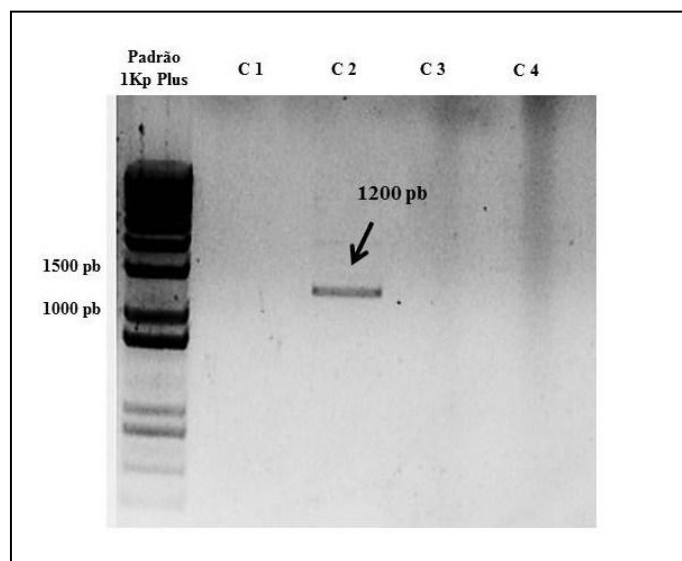


Figura 22: PCR da colônia positiva para inserção do plasmídeo pEG.

Em apenas uma colônia pudemos observar a amplificação de um fragmento de 1200 pb aproximados, condizente com o esperado para uma construção correta. Esta colônia foi utilizada para a propagação deste plasmídeo para as etapas posteriores de clonagem. Gel de agarose 1%.

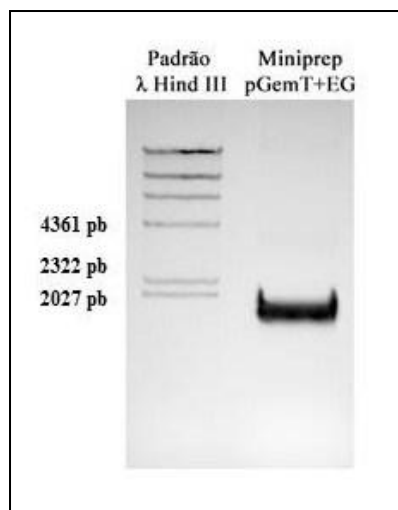


Figura 23: Análise da qualidade de extração plasmidial da colônia positiva.

Para a análise da qualidade da extração, 3 μ L do extraído foram verificados por eletroforese em gel de agarose 1%. O resultado evidencia tanto a qualidade de extração quanto eficiência do tratamento.

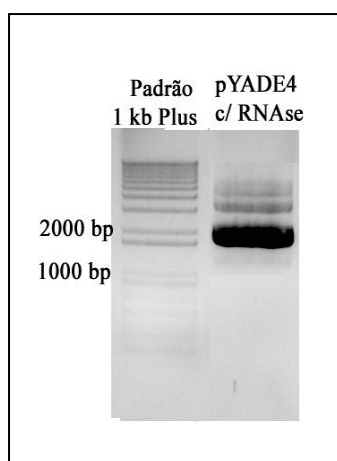


Figura 24: Extração plasmidial do pYADE4.

A extração plasmidial foi realizada de modo que 3 μ L da amostra tratada com RNase passaram por eletroforese em gel de agarose 1%.

5.6 Sequenciamento dos clones

O DNA plasmidial foi sequenciado tal que, destes, obtivemos trechos do fragmento total. Para sequenciamento foram utilizados os *primers* M13 e *primers*

específicos utilizados na clonagem (VEgF e VEgR). Esses trechos comparados; mostram uma sequência consenso de 1170 pb; condizente com o esperado a partir do transcriptoma desse cupim (Figura 25).

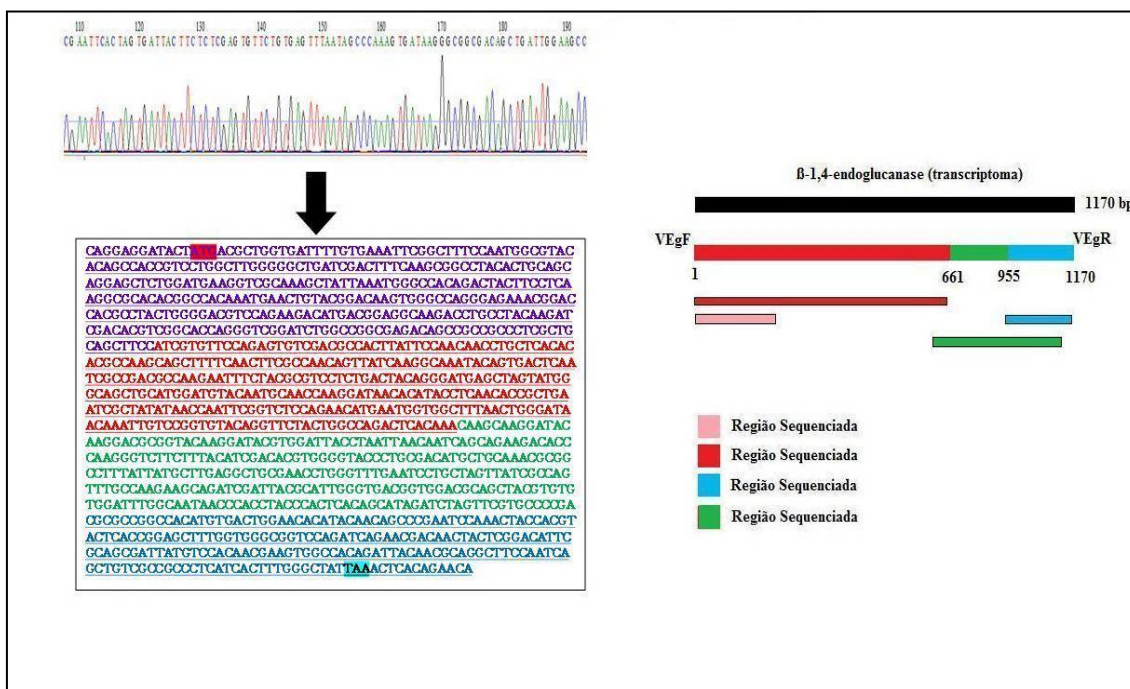


Figura 25: Sequência Consenso da β -1,4- endoglucanase de V. heteropterus encontrada a partir do alinhamento de trechos sequenciados a partir da construção pEG.

A partir dos trechos sequenciados e baseado na sequência do transcriptoma deste cupim, pudemos encontrar uma sequência consenso de aproximadamente 1170 pb. Ao alinhar os trechos encontrados a partir dos sequenciamentos, verificamos que esta alinhamento forma um *contig* de 1170 pb, condizente com o esperado; indicando que esta sequência foi integrada completamente ao plasmídeo pGemT Easy com sucesso. Os trechos mostrados em vermelho e azul indicam as regiões obtidas com base no sequenciamento.

A sequência consenso, obtida a partir do sequenciamento dos clones, foi comparada com o banco de dados utilizando a ferramenta BLAST-N do NCBI e mostra semelhanças com outros cupins (Figura 26); sendo que a comparação indica maior similaridade com o mRNA para β -1,4-endoglucanase do cupim Nasutitermes takasagoensis de GI:6253262 (Termitidae: Nasutitermes) com 85 % de alinhamento idêntico (Figura 27 e 28).

Select: All None Selected: 0							
Alignments Download GenBank Graphics Distance tree of results							
	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	Nasutitermes takasagoensis mRNA for endo-b-1,4-glucanase, complete cds	1290	1290	100%	0.0	85%	AB013272.2
<input type="checkbox"/>	Nasutitermes walkeri mRNA for NwEG, complete cds	1265	1265	99%	0.0	84%	AB013273.1
<input type="checkbox"/>	Odontotermes formosanus OfEG3 mRNA for putative endo-beta-1,4-glucanase OfEG3, partial	1020	1020	100%	0.0	80%	AB118802.1
<input type="checkbox"/>	Macrotermes barnevi endoglucanase mRNA, complete cds	1007	1007	98%	0.0	80%	JQ292846.1
<input type="checkbox"/>	Coptotermes acinaciformis cellulase mRNA, complete cds	1000	1000	99%	0.0	79%	AF336120.1
<input type="checkbox"/>	Odontotermes formosanus endo-beta-1,4-glucanase mRNA, complete cds	998	998	100%	0.0	79%	GU326330.1
<input type="checkbox"/>	Coptotermes gestroi endo-beta-1,4-glucanase mRNA, complete cds	966	966	98%	0.0	79%	KC891003.1
<input type="checkbox"/>	Coptotermes formosanus endo-beta-1,4-glucanase (EG3b) mRNA, complete cds	964	964	99%	0.0	79%	GU017483.1
<input type="checkbox"/>	Reticulitermes speratus RsEG mRNA for salivary cellulase, complete cds	902	902	96%	0.0	78%	AB008778.2
<input type="checkbox"/>	Reticulitermes flavipes endogenous cellulase mRNA, complete cds	884	884	96%	0.0	78%	AY572862.2
<input type="checkbox"/>	Reticulitermes speratus mRNA for Endoglucanase 2, complete cds	875	875	96%	0.0	77%	AB019095.1
<input type="checkbox"/>	Coptotermes formosanus CfEG3 mRNA for endo-b-1,4-glucanase, complete cds	857	857	97%	0.0	77%	AB058670.1
<input type="checkbox"/>	Coptotermes formosanus CfEG4 mRNA for endo-b-1,4-glucanase, complete cds	848	848	97%	0.0	77%	AB058671.1

Figura 26: Sequência consenso comparada contra bando de dados utilizando a ferramenta BlastN.

A sequência consenso obtida através do sequenciamento juntamente com os dados do transcriptoma mostrou uma maior similaridade com a espécie Nasutitermes takasagoensis; seguido de Nasutitermes walkeri com similaridade de 85 e 84 %, respectivamente. A sequência condiz com o mRNA de β -1,4-endoglucanase (EC: 3.4.1.2).

Nasutitermes takasagoensis mRNA for endo-b-1,4-glucanase, complete cds					
Sequence ID: dbj AB013272.2 Length: 1685 Number of Matches: 1					
Range 1: 312 to 1480 GenBank Graphics ▼ Next Match ▲ Previous Match					
Score	Expect	Identities	Gaps	Strand	
1290 bits(1430)	0.0	994/1173(85%)	5/1173(0%)	Plus/Plus	
Query	1	CAGGAGGATACTATGACACTGGTGAATTTTGTGAAAATCGGCTTTCCAATGGCGTACACAG			60
Sbjct	312	CAGGAGGATACTTTGACGCTGGTGAATTTTGTGAAAATCGGCTTTCCAATGGCGTACACAG			371
Query	61	CCACCGTCCTGGCTTGGGGCTGATCGACTTTCAAGCGGCCTACACAGCAGCAAGAGCTC			120
Sbjct	372	CCACCGTCCTGGCTTGGGGCTGATCGACTTTGAAGCGGGCTACTCTTCAGCAGGAGCTC			431
Query	121	TGGATGAACGTCCTCAAGCTATTAATGGGCCACAGACTACTTCTCAAGCGCACAGG			180
Sbjct	432	TGGATGATGGTCGCAAGCTGTTAAATGGGCCACAGACTATTTTCATCAAGCGCACAGCT			491
Query	181	TCACAGATGAACTGTAAGGACAAAGTGGGCCACGGAGAAACGGACCAGCCTACTGTGGGC			240
Sbjct	492	CCCAAAATGAATTTACGGACAAAGTGGGCCAGGGAGATGCGGACCAGCCTTCTGGGGAC			551
Query	241	GTCCAGAAGACATGACGGAGGCAAGACCTGCCTACTAGATCGACCCGTCGGCACCCTGT			300
Sbjct	552	GTCCAGAAGACATGACAAATGGCAAGACCTGCCTACAAGATCGACACGCTTAGACCAGGGT			611
Query	301	CTGATCTGACCGGCGAGACAGAGCGCCGCCCTCGCTGACAGCTTCCATCGTGTTCAGAGT			360
Sbjct	612	CGGATCTGGCCGGCGAAACAGC-CGCCGCCCTCGCTGACAGCTTCCATCGTGTTCAGGAAT			670
Query	361	GTAGACGCCACTTATTCCTACTACCTGCGCGCCAAACCCCAAGCATCTTGTCAACATCGAC			420
Sbjct	671	GTCGACGGCACTTATTCCTAACACCTGCTCACACAGCCAGGAGCTTTTCGACTTCGCC			730
Query	421	AATACTTACCCACCCAAAGACAATGACTCAGTGTCCGACGCCAGAAATTCACCGCTCA			480
Sbjct	731	AACAATTTATCGCGCCAAATACAGTGAATCAATCACCAGCCAGGAATTTCTACGCGTCC			790
Query	481	TCTGACTACCAGGATGAACTCGTATGGGCCGCTAAATAGCTCTACTGAACAACCAACGGT			540
Sbjct	791	GCGGACTACAGGGATGAATTAGTATGGGCAGCCGATGGCTCTACAGGGCAACCAACGAC			850
Query	541	ATCCCCTACCTCA-CACTCGCGTGAATCGGTATGTAACCAATTCGGGATCCAGAACAAAGC			599
Sbjct	851	ARCACCTACCTCAACAC-CGC-TGAATCGCTATATGATGAATTCGGTCTCCAGAACTGGG			908

Figura 27: Sequência consenso alinhada com a β -1,4-endoglucanase de Nasutitermes takasagoensis.

Após o alinhamento, notamos que a maior similaridade com a espécie Nasutitermes takasagoensis com 85 % de identidade com a β -1,4-endoglucanase deste cupim. Na imagem apenas uma parte deste alinhamento é mostrado para fins de registro. O *Query* indica a sequência de N. takasagoensis e o *Subject*, a sequência consenso.

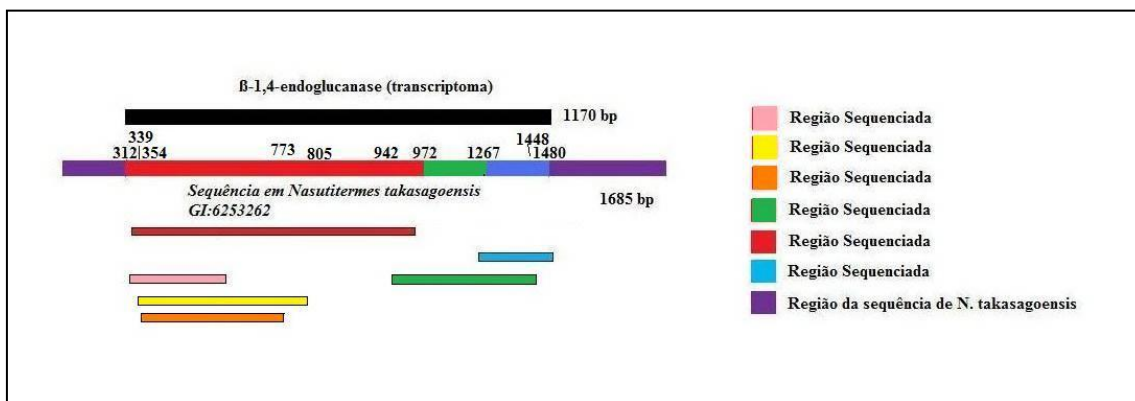


Figura 28: Esquema do alinhamento entre a sequência consenso de β -1,4-endoglucanase de V. heteropterus sp com sequência de Nasutitermes takasagoensis.

O esquema acima mostra o alinhamento entre as sequências de β -1,4-endoglucanase para os cupins V. heteropterus e Nasutitermes takasagoensis. No caso, o tamanho da sequência de N. takasagoensis é de 1650 pb contra os 1170 pb encontrados para a sequência de V. heteropterus. Na figura, as cores rosa, vermelha e azul indicam os *uniques* resultantes independentes do sequenciamento. A região em roxo indica trechos encontrados apenas em N. takasagoensis e indicam regiões divergentes do próprio cupim.

A figura 29 mostra esta mesma sequência quando comparada com a sequência encontrada no transcriptoma; de modo que seu alinhamento demonstra 92% de identidade. Isso pode ser devido ao baixo número de sequências obtidas no sequenciamento e que provavelmente as regiões de disparidade sejam regiões de baixa qualidade e não a possíveis mutações, levando em consideração a utilização da *Phusion* (veja seção 4.10). Um novo sequenciamento desta região está sendo feito com o objetivo de se identificar a relação desta sequência.

Score	Expect	Identities	Gaps	Strand
1674 bits(1856)	0.0	1079/1173(92%)	5/1173(0%)	Plus/Plus
Query 1	CAGGAGGATACTATGACGCTGGTGTATTTGTGAAATTCGGCTTCCAATGGCGTACACAG	60		
Sbjct 1	CAGGAGGATACTATGACACTGGTGTATTTGTGAAATTCGGCTTCCAATGGCGTACACAG	60		
Query 61	CCACCGTCCTGGCTTGGGGGCTGATCGACTTTC AAGCGGCCCTACACTGCAGCAGGAGCTC	120		
Sbjct 61	CCACCGTCCTGGCTTGGGGGCTGATCGACTTTC AAGCGGCCCTACACAGCAGCAAGAGCTC	120		
Query 121	TGGATGAAGGTCGCAAAGCTATTAATGGGCCACAGACTACTTCCTCAAGGCGCACACGG	180		
Sbjct 121	TGGATGAACGTCCCAAAGCTATTAATGGGCCACAGACTACTTCCTCAAGGCGCACACGG	180		
Query 181	CCACAAATGAACTGTACGGACAAGTGGGCCAGGGAGAAACGGACCACGCCTACTGGGGAC	240		
Sbjct 181	TCACAGATGAACTGTAAGGACAAGTGGGCCACGGAGAAACGGACCACGCCTACTGTGGGC	240		
Query 241	GTCCAGAAGACATGACGGAGGCAAGACCTGCCTACAAGATCGACACGTCGGCACCAGGGT	300		
Sbjct 241	GTCCAGAAGACATGACGGAGGCAAGACCTGCCTACTAGATCGACCCGTCGGCACCACACTGT	300		
Query 301	CGGATCTGGCCGGCGAGACAGC-CGCCGCCCTCGCTGCAGCTTCCATCGTGTTCAGAGT	359		
Sbjct 301	CTGATCTGACCGGCGAGACAGAGCGCCGCCCTCGCTGCAGCTTCCATCGTGTTCAGAGT	360		
Query 360	GTCGACGCCACTTATTC AACAACCTGCTCACACACGCCAAGCAGCTTTTCAACTTCGCC	419		
Sbjct 361	GTAGACGCCACTTATTC AACTACCTGCGCGCCACCCCAAGCATTTGTCAACATCGAC	420		
Query 420	AACAGTTATCAAGGCAAATACAGTACTCAATCGCCGACGCCAAGAATTTCTACGCGTCC	479		
Sbjct 421	AATACTTACCCACCCAAAGACAATGACTCAGTGTCCGACGCCAGAATTTCAACGCGTCA	480		
Query 480	TCTGACTACAGGGATGAGCTAGTATGGGCAGCTGCATGGATGTACAATGCAACCAAGGAT	539		
Sbjct 481	TCTGACTACCAGGATGAACTCGTATGGGCCGCTA AATAGCTCTACTGAACCAACCAAGGAT	540		

Figura 29: Sequência consenso obtida pelo sequenciamento do plasmídeo pEG comparada contra Sequência do transcriptoma de V. heteropterus.

Uma identidade de 92 % foi encontrada quando comparamos a sequência consenso obtida com base no sequenciamento com a encontrada na análise do transcriptoma do cupim V. heteropterus. A sequência condiz com o mRNA de β -1,4-endoglucanase (EC: 3.4.1.2); como já previamente analisado. Na imagem apenas parte deste alinhamento é mostrado para fins de registro. O *Query* indica a sequência retirada do transcriptoma e o *Subject*, a sequência consenso. As regiões de disparidade provavelmente representam regiões de baixa qualidade do sequenciamento.

Na comparação entre as sequências do V. heteropterus e N. takasagoensis notamos que o alinhamento se inicia no par de base de número 312 na sequência do N. takasagoensis e de número 1 na sequência do V. heteropterus; terminando no par de base 1480 e 1172, respectivamente. Além disso, verifica-se um escore de 1260 (escore máximo 1430), comprovando que o número de pontos similares é superior ao número de *gaps*, *gaps extension* e diferenças. Nesse sentido, podemos afirmar que, em relação ao fragmento do mRNA da enzima β -1,4-endoglucanase, tanto o Nasutitermes takasagoensis como do V. heteropterus são possivelmente similares. Assim, acredita-se

que tal fragmento codifica, provavelmente, aminoácidos semelhantes na maioria da proteína, haja visto as poucas divergências no alinhamento.

Essa ligeira divergência no número de pares de base e, conseqüentemente, de aminoácidos, pode, no entanto, afetar as características enzimáticas, como temperatura e pH ideais. Todavia, testes devem ser realizados para tal comprovação. Mesmo assim, dado que a análise comparativa mostra que a sequência consenso abrange quase toda a extensão da sequência do mRNA de N. takasagoensis, pode-se dizer que, possivelmente, ambas possuem o mesmo sítio catalítico; desempenhando o mesmo papel.

Em literatura, a endoglucanase de N. takasagoensis é descrita como sendo uma enzima estável; com pH ótimo ser em torno de 6,0 e temperatura ótima ser em torno de 65 °C⁵ e de massa molecular aproximada de 37 KDa⁷³. Talvez, por a sequência de β-1,4-endoglucanase de V. heteropterus ser possivelmente menor e com algumas modificações na estrutura de aminoácidos, podemos esperar que as suas condições ambientais sejam diferenciadas em comparação a β-1,4-endoglucanase encontrada em cupins N. takasagoensis.

Em adição, quando analisado o Valor de E notamos que este é de 0 em relação ao Nasutitermes takasagoensis e, portanto, tal alinhamento não possui chances de ter ocorrido ao acaso. A semelhança de 85 % entre os dois fragmentos indica que são conservados evolutivamente em relação a ambos os cupins. Isso também pode ser comprovado quando analisamos a sequência de β-1,4-endoglucanase de V. heteropterus com de 7 espécies de cupins já descritos em literatura e que possuem identidade acima de 75 %, como mostrado na figura 26 (Nasutitermes walkeri, Coptotermes formosanus, Coptotermes acinaciformis, Sinocapritermes mushae, Reticulitermes speratus, Odontotermes formosanus, e Reticulitermes flavipes) Na figura 30 verificamos, através

A análise *in silico* indica, portanto, que entre as espécies ocorre grande similaridade entre os nucleotídeos; possuindo diversos sítios conservados entre si. Assim, a β -1,4-endoglucanase é uma enzima presente nestes cupins, de modo a configurar proteínas semelhantes; possivelmente com os mesmos domínios e estruturas secundária e terciária e domínio de ligação (veja seção 5.13).

5.7 Filogenética da β -1,4-endoglucanase

Uma análise filogenética, através do método Neighbor-Joining (Max. Seq. Difference = 0,75) (Figura 32), cujo qual se baseia na distância filogenética (Número relativo de similaridades e diferenças entre os grupos de sequências), comprova a similaridade citada na seção 5.6. No caso desta análise nota-se que o gênero V. heteropterus possui, de fato, mais similaridade com o grupo Nasutitermes sp em detrimento dos demais clados comparados; indicando que a identidade com a sequência deste no Blast-N está correta.

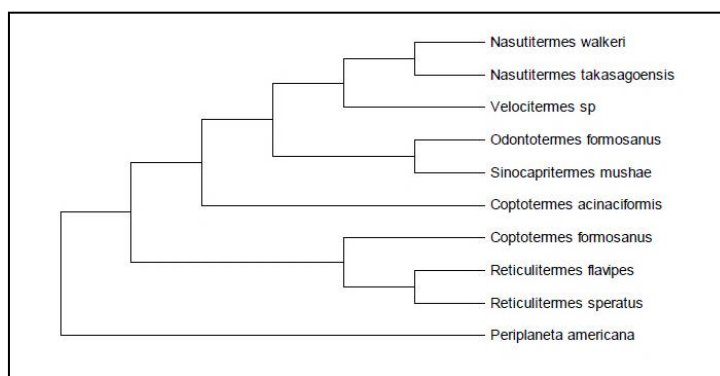


Figura 32: Análise filogenética utilizando o Método Neighbor-Joining.

A utilização do método permite, com base no número relativo de similaridades e diferenças, determinar a distância filogenética entre os clados alinhados para a celulase β -1,4-endoglucanase. Neste verificamos a proximidade do gênero V. heteropterus com a espécie Nasutitermes takasagoensis. A análise foi realizada utilizando-se o programa MEGA 5.05²⁰⁸.

Nota-se também que, apesar de se assemelhar, evolutivamente, mais com grupos superiores de cupins; a espécie V. heteropterus também possui certa relação com um grupo basal representados pelas espécies Reticulitermes speratus e Reticulitermes flavipes; que também possuem suas celulasas como alvos de estudos de interesse^{69,73,114}. Isso é devido, provavelmente, as regiões altamente conservadas presentes em todos os cupins analisados. Além disso, o Método Neighbor-Joining utiliza distâncias genéticas como uma métrica de comparação, produzindo árvores não enraizadas e não assumindo taxa constante de evolução em linhagens (Relógio Molecular)^{208,e} e, portanto, não proporciona indicativo de quando a divergência ocorreu a partir do alinhamento das sequências.

Uma Análise através do Método de Máxima Verossimilhança (Figura 33), contudo, demonstra que, apesar dessa relação existir, a análise probabilística de cada sítio indica que de fato as espécies R. flavipes e R. speratus são mais basais; sendo mais próximas do Coptotermes formosanus do que do próprio V. heteropterus e mais próxima da Periplaneta americana (Blattidae: Periplaneta); portanto, comprova o processo evolutivo em relação a celulase; como já citado por Davison e Blaxter (2005)¹¹⁵.

Esse tipo de análise é eficiente uma vez que utiliza todos os sítios, permitindo uma grande acurácia dos tamanhos dos ramos, perfazendo um modelo evolutivo específico que garante o relógio molecular e, portanto, permite dizer o tempo de divergência; ou seja, qual clado surgiu antes (basal) e qual é o mais recente (superior)²⁰⁹.

e. O Relógio Molecular consiste basicamente em uma técnica que relaciona a evolução molecular com o tempo de divergência entre 2 ou mais espécies com o número de diferenças moleculares medidas através da proteína ou DNA. Assim, o parentesco está diretamente relacionado ao tempo e, portanto, quanto mais próximo geneticamente, menor o tempo de separação entre as espécies. Essa teoria afirma que as moléculas evoluem de forma constante e permite contar as desconexões em sequências moleculares entre indivíduos existentes²⁰⁹.

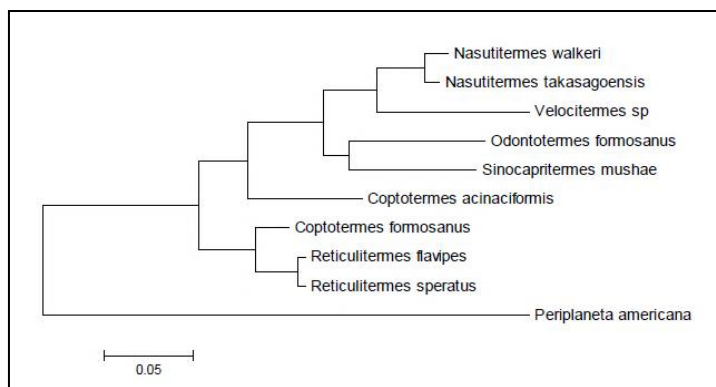


Figura 33: Análise filogenética utilizando o Método de Máxima Verossimilhança.

A utilização do método, baseado na taxa de transição e transversões encontradas nos sítios (bases), é feito de modo probabilístico; tal que permite a verificação do processo evolutivo. Notamos acima, portanto, que V. heteropterus é um grupo superior. A relação filogenética com o grupo basal representado pelas espécies Reticulitermes flavipes e Reticulitermes speratus comprova-se através das distâncias notadas na figura; mostrando ser estes últimos mais próximo da barata (Periplaneta americana) e da espécie de cupim Coptotermes formosanus do que dos demais cupins derivados. A análise foi realizada utilizando-se o programa MEGA 5.05²⁰⁸.

5.8 Construção do plasmídeo recombinante β -1,4-endoglucanase e Expressão heteróloga da em levedura S. cerevisiae FGSC 9721

A expressão heteróloga constituiu-se, inicialmente, da digestão tanto do pGemT-Easy+EG quanto do pYADE4 (Vetor de expressão) fazendo uso das enzimas *Bam*HI e *Xho*I, que clivam regiões presentes nos *primers* utilizados para amplificação do fragmento e MCS do vetor de expressão (nas posições 64 e 13; respectivamente), como esquematizado na Figura 34. Essa clivagem permite que o iEG (inserto EG) seja retirando do plasmídeo de clonagem (Figura 35); bem como a linearização do pYADE4 de modo que seja possível a posterior ligação entre o iEG e o pYADE4 (Figura 36). O iEG apresentou 1200 pb aproximadamente, condizendo com o esperado (tamanho da β -1,4-endoglucanase). Já o pYADE4 linear apresentou um tamanho de aproximadamente 6000 pb, confirmando sua linearização. A ligação é possível devido à clivagem do tipo *stick-end*; como mostrado na figura 35.

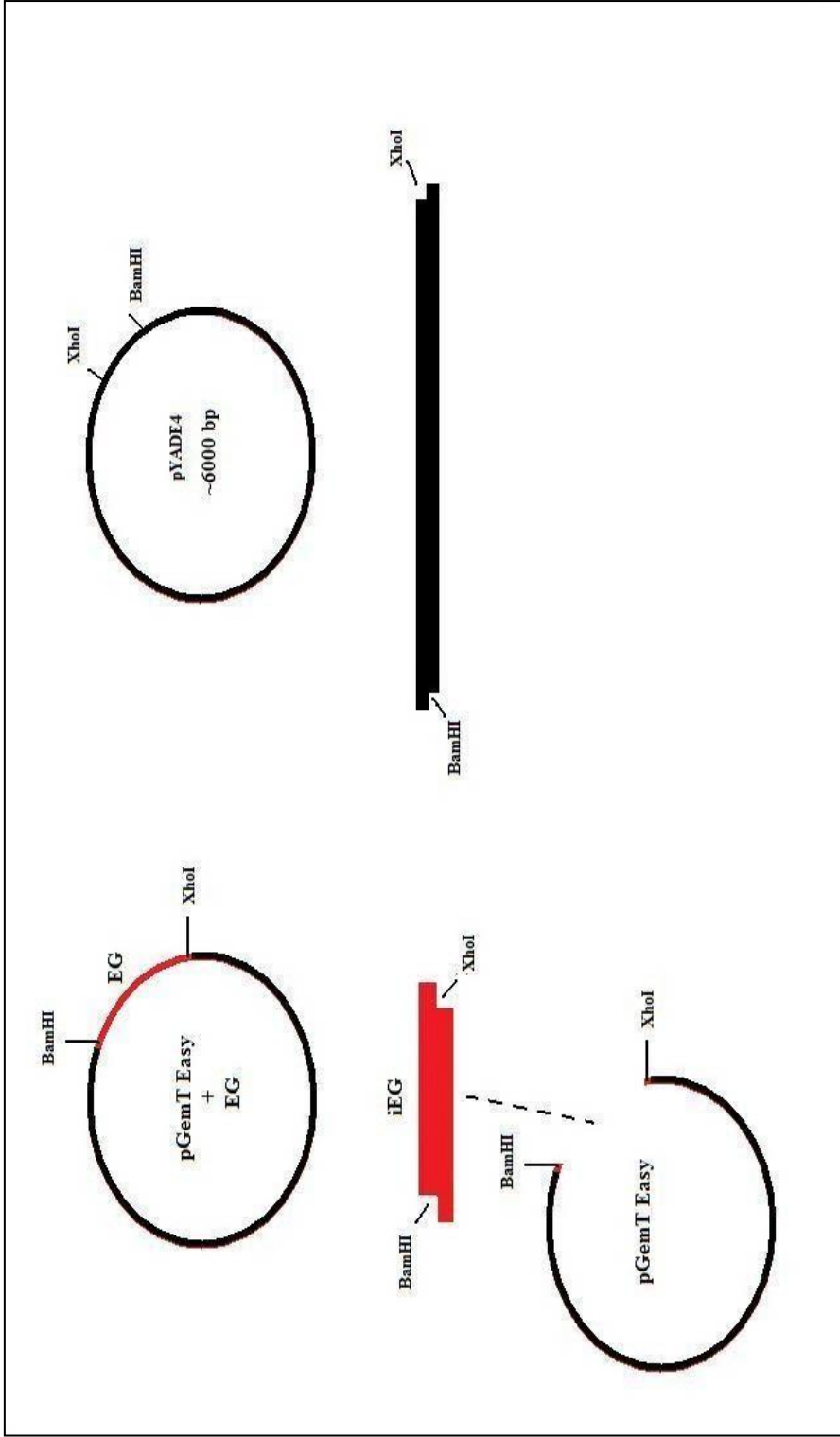


Figura 34: Esquematização da clivagem do pGemT Easy+EG e vetor de expressão pYADE4.

Esquema acima mostra as regiões de clivagem ao utilizar as enzimas de digestão BamHI e XhoI, respectivamente. Essa clivagem permite a linearização do pYADE4 e a retirada o iEG do plasmídeo de clonagem.

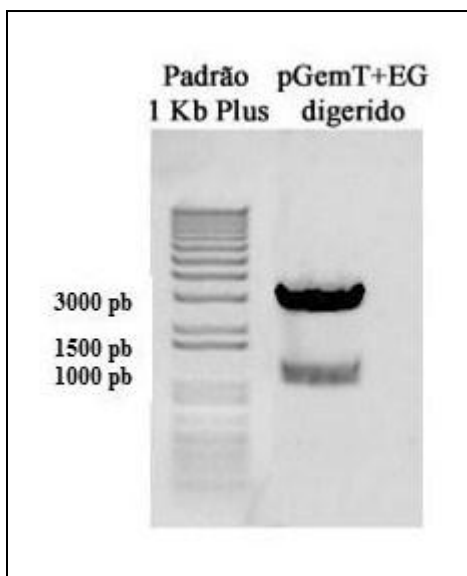


Figura 35: Digestão do pGemT-Easy+EG usando BamHI e XhoI.

A partir do gel de agarose 0,7 % é possível verificar a digestão completa do pEG de modo a mostrar que o iEG foi totalmente digerido e que este apresenta aproximados 1200 pb; condizente com a liberação do fragmento contendo o gene da β -1,-4endoglucanase conforme o esperado.

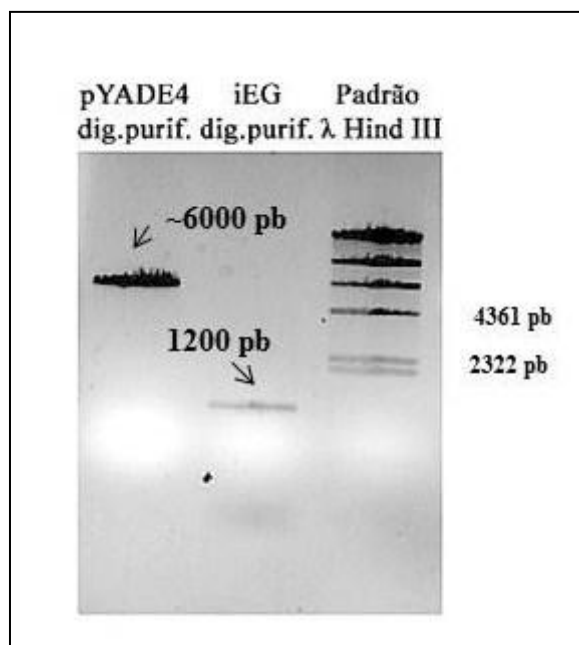


Figura 36: Purificação do pYADE4 e do inserto (iEG) após a digestão.

O fragmento (iEG) digerido e o plasmídeo pYADE4 digeridos foram purificados. A figura mostra a qualidade da purificação e tamanho respectivos de cada um. No caso do iEG um tamanho de 1200 pb e do pYADE4, um tamanho de um pouco mais de 6000 pb; condizente com o esperado.

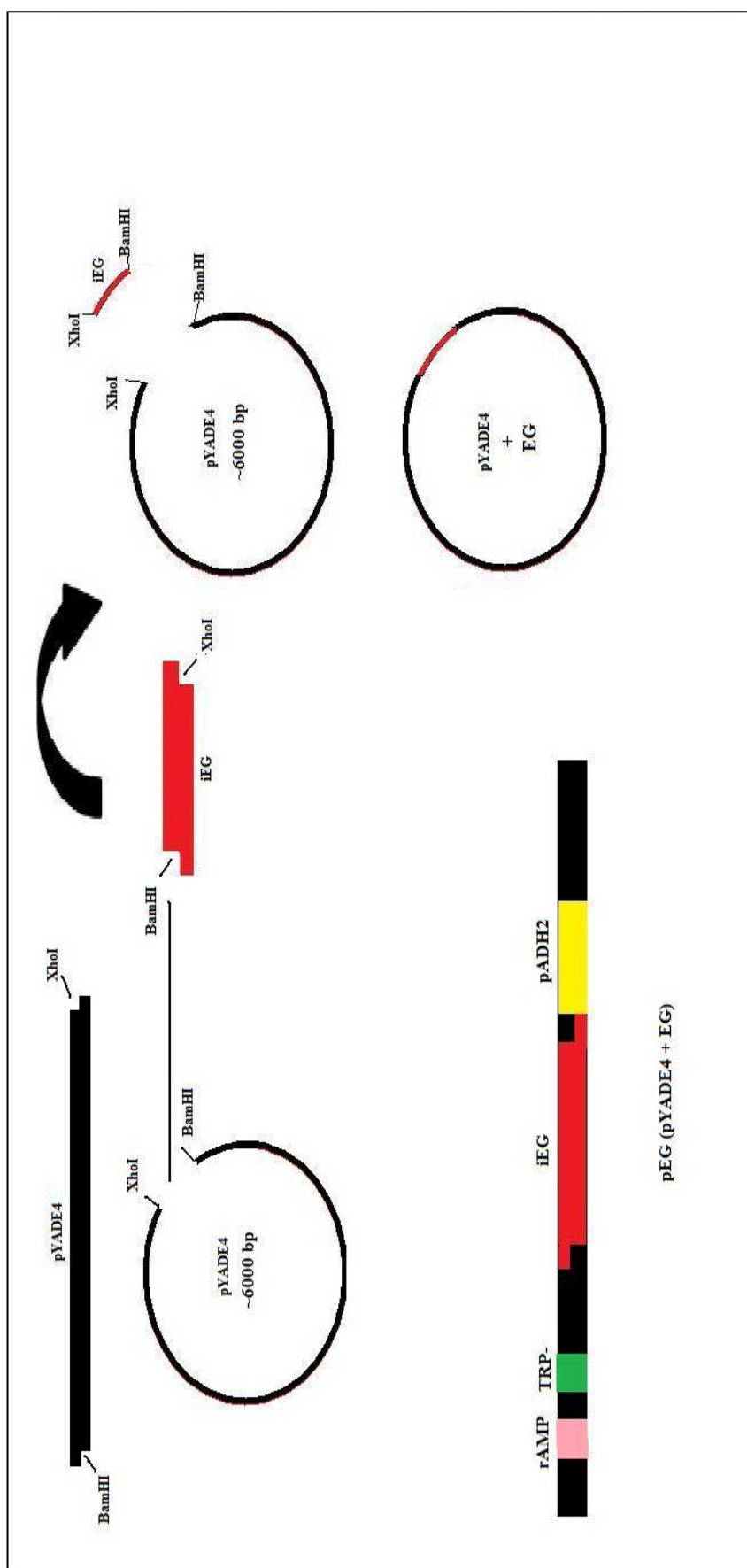


Figura 37: Esquematisação da ligação entre pYADE4 e o iEG.

O esquema mostra como se dá a ligação entre o inserto e o plasmídeo de expressão digeridos e purificados. Neste caso as regiões clivadas com as enzimas permitem um alinhamento perfeito entre as regiões, uma vez que se trata de uma clivagem do tipo *stick-end*. Assim, a região clivada com BamHI do pYADE4 se liga com a região do iEG clivada com esta mesma enzima. O mesmo ocorre em relação a clivagem pela XhoI.

A ligação entre o pYADE4 e o iEG, denominada aqui de pYADE4+EG, foi inserida em bactérias *E. coli* eletrocompetentes. Das colônias, foram selecionadas cerca de 10 para a realização da PCR de colônia seguindo protocolo padrão. A partir dessa PCR, obtivemos apenas uma colônia positiva (Figura 38). Desta foi extraído DNA plasmidial (Figura 39) e realizado novamente PCR utilizando os *primers* específicos (VEgF e VEgR) e *primers* ADH2F e pYADE4R (descritos na seção 4.5). Estes últimos flaqueam regiões externas ao MCS do pYADE4; i.e., presentes em regiões internas do plasmídeo. No caso, o ADH2F flaqueia região junto ao promotor ADH2 e o pYADE4R engloba uma parte do plasmídeo. Assim, a amplificação usando ADH2F e pYADE4R foi positiva, proporcionando um fragmento de cerca de 2000 pb. Caso esta amplificação fosse negativa, ocorreria uma amplificação de aproximadamente 800 pb. Para os *primers* específicos, o fragmento apresentou 1200 pb (Figura 40).

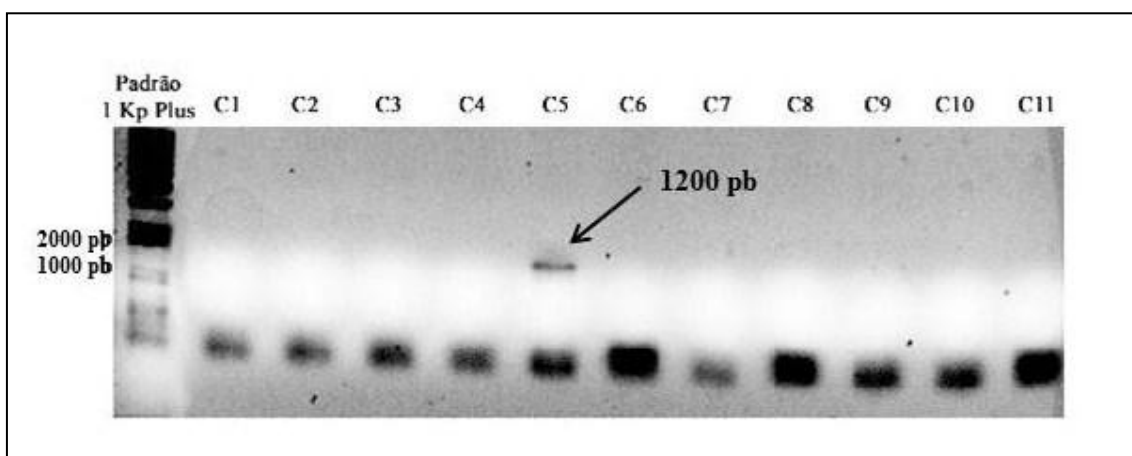


Figura 38: PCR de Colônia para as colônias selecionadas de bactérias para a verificação da construção do plasmídeo yEG (pYADE4+EG).

A PCR de colônia foi realizada com diversas colônias. Como evidenciado, houve amplificação do fragmento esperado de aproximadamente 1200 pb em apenas uma colônia. Para tanto foram usando *primers* específicos VEgF e VEgR. Gel agarose 1%.

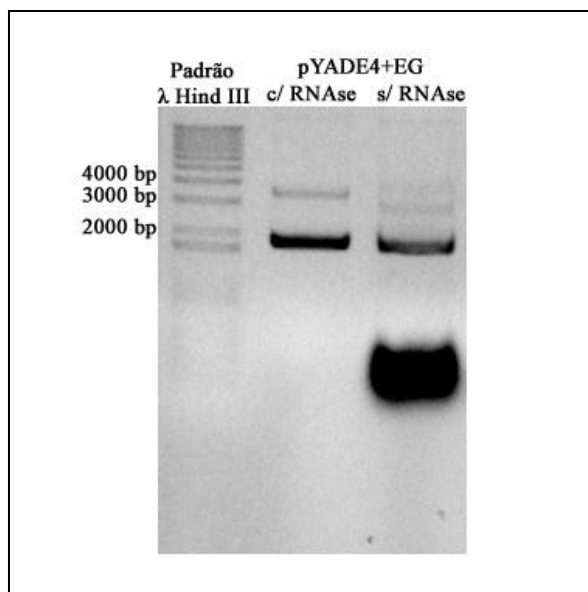


Figura 39: Midipreparação plasmidial do yEG.

Após a preparação do plasmídeo, obtivemos uma alta concentração e qualidade de DNA plasmidial. Acima é mostrado a *Midiprep* (aplicação de 4 μ L) tratada com RNase. Gel de agarose 0,7 %.

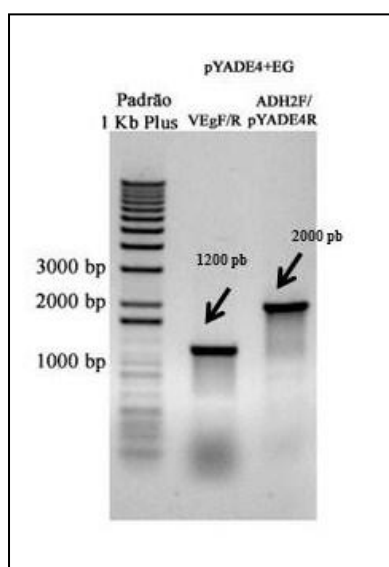


Figura 40: PCR para a verificação da construção do plasmídeo yEG.

A PCR utilizando os *primers* específicos VEGF e VegR mostra um fragmento de 1200 pb. Já a PCR usando *primers* ADH2F e pYADE4R mostra um fragmento de 2000 bp; como esperando. A não inserção do fragmento respectivo ao gene resultaria na amplificação de um fragmento de 800 pb; quando utilizado *primers* que hibridizam no vetor. Gel de agarose 0,7 %.

Linhagens de *S. cerevisiae* FGSC 9721 foram transformadas com a midipreparação do pYADE4+EG de acordo com a seção 4.15.2.

Tal linhagem foi escolhida por possuir marca autotrófica *trp1*, garantindo que as somente colônias que possuíssem o pYADE4+EG cresceriam em meio mínimo *trp*⁻, i.e, meio mínimo sem a presença do triptofano. Isso porque o plasmídeo de expressão pYADE4 possui marcação *trp*⁻ e, portanto, é capaz de sintetizar este aminoácido; garantindo a sobrevivência da levedura.

Dessa forma, levando em consideração tal fato, todas as colônias que cresceram em meio *trp*⁻; mas que não cresceram em meio *leu*⁻ (por não haver nenhuma marcação para tal aminoácido), teoricamente, seriam positivas.

5.9 Expressão da proteína recombinante β -1,4-endoglucanase

Através da técnica de SDS-PAGE verificou-se que a β -1,4-endoglucanase possui uma massa molecular de 43/44 kDa (Figura 41). Esse dado é corroborado com dados em literatura, cujos quais predizem um tamanho entre 37 a 50 kDa para esta celulase em alguns cupins já descritos^{107,116,117,132,136,143,144,153,155,181,182,184,191,192}.

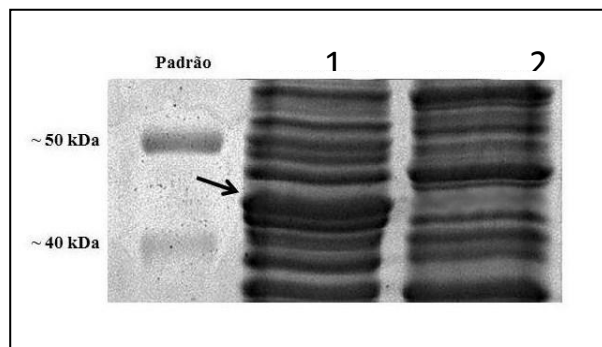


Figura 41: Análise de expressão de β -1,4-endoglucanase em *S. cerevisiae*.

Através do gel verificamos a presença de uma banda próxima a 40 kDa diferencialmente expressa entre (1) a extração de leveduras transformadas com o plasmídeo pYADE4+EG e (2) controle negativo – levedura não transformada; segundo marcador *Spectra Multicolor Broad Range Protein Ladder* (Thermo ScientificTM). A presença desta proteína na linhagem transformante indica a expressão desta proteína na linhagem de *S. cerevisiae* laboratorial conforme o esperado. Este dado corrobora com β -1,4-endoglucanases descritas de outras espécies de cupins. A extração proteica foi realizada após 18h de incubação das linhagens em meio *sc trp*⁻ à 30 °C sob agitação constante.

Em adição, com o objetivo de caracterizar e identificar a enzima β -1,4-endoglucanase (EG) recombinante produzida na *S. cerevisiae*, foi realizado uma eletroforese do extrato proteico intra e extracelular; sendo utilizado como controle a própria linhagem não transformada. Neste teste comprovou-se que a proteína encontrava-se internamente nas células de levedura como esperado; não sendo, portanto, exportada para o meio (Figura 42).

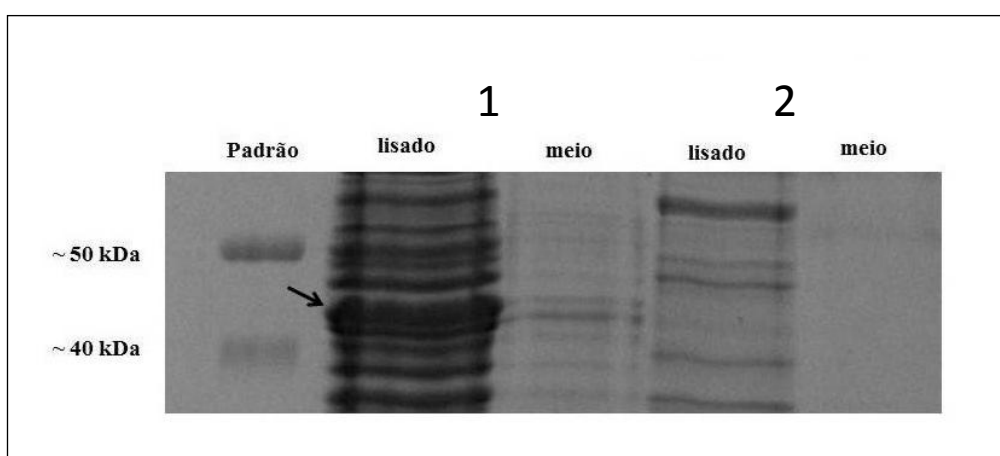


Figura 42: SDS-PAGE da linhagem recombinante e controle do lisado e do meio.

Este SDS-PAGE permitiu comprovar que a proteína está de fato presente no ambiente intracelular e, portanto, não é exportada para o meio. (1) transformantes; (2) não transformantes.

5.10 Teste de atividade em meio CMC

Cerca de oito colônias transformantes foram pré-selecionadas e passaram por teste em CMC a fim de confirmar a presença da proteína. A expressão desta proteína neste meio resultaria em um halo de degradação conforme descrito na seção 4.17. Para tanto extratos proteicos foram extraídos seguindo protocolo da seção 4.16. A partir destes extratos foram postos no meio CMC pelo período de aproximadamente 26 horas.

A presença de um halo garantiu a eficiência do teste, mostrando que a proteína está de fato presente (Figura 43). No controle negativo, ou seja, o extrato proteico da levedura não-transformada não apresentou halo (dado não mostrado).

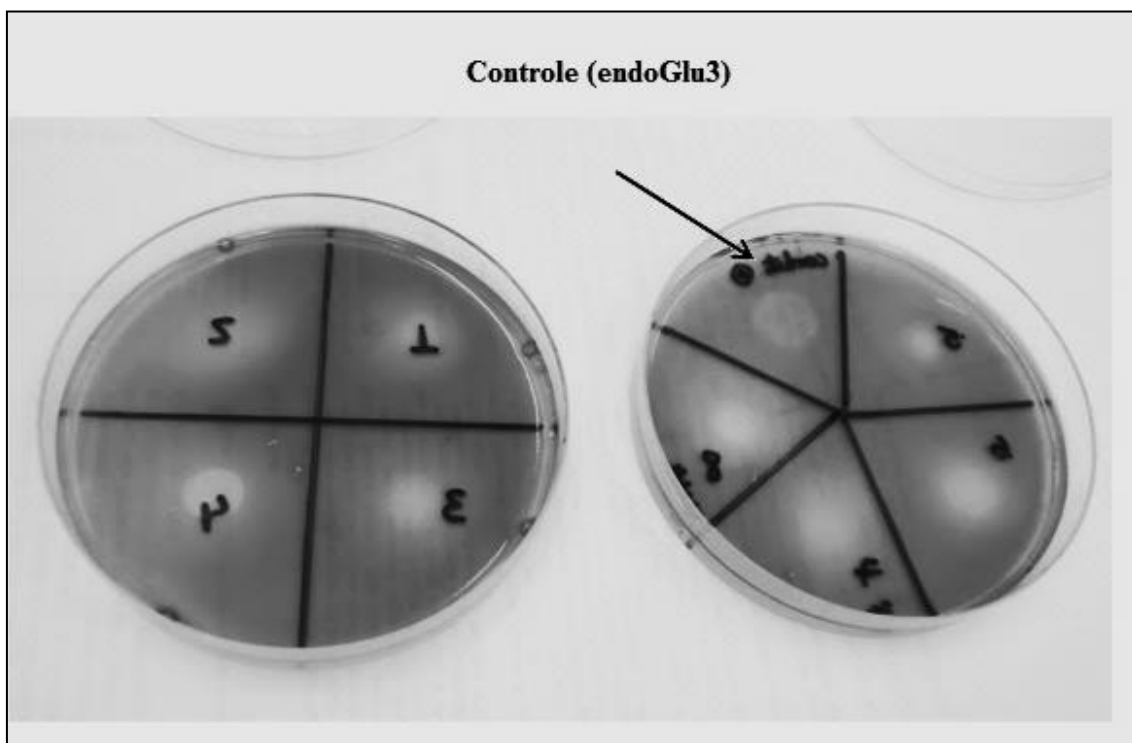


Figura 43: Teste CMC com linhagens transformantes selecionadas.

O teste em CMC permitiu a verificação da expressão proteica da β -1,4-endoglucanase nas linhagens selecionadas. A presença do halo (branco) indica a presença da celulase. Um controle positivo (endoGlu3) foi utilizado. Este controle refere-se à endoglucanase III de Trichoderma harzianum e foi cedido pelo grupo do Laboratório de Genética Molecular (UFSCar-São Carlos, SP) supervisionado pelo Prof. Dr. Flávio Henrique Silva^{210,f}. O controle negativo (não mostrado) não apresentou a presença do halo.

f. Para mais informações sobre esta celulase utilizada como controle, consultar GENEROSO, W.C. *et al.* **Recombinant expression and characterization of an endoglucanase III (cel12a) from Trichoderma harzianum (Hypocreaceae) in the yeast Pichia pastoris**. Brasil: Genetics and Molecular Research, 2012. Vol. 11, n. 2, p. 1544-1557²¹⁰.

5.11 Testes de crescimento e viabilidade celular e nível proteico

Uma análise comparada a linhagem selecionada com uma linhagem selvagem utilizada na fermentação em Usinas (PE-2) foi realizada para averiguação de possíveis diferenças entre si durante um teste fermentativo laboratorial. No caso, este teste consistiu na verificação do crescimento, viabilidade e consumo de glicose entre as linhagens. Esses fatores permitiram uma comparação de eficiência entre uma linhagem já comercialmente utilizada com a linhagem recombinante em teste.

A figura 44 mostram essas perspectivas em análise. Nestes verificamos que o consumo de glicose, com um inicial de 4%, foi relativamente semelhante, de modo que ambas as linhagens consumiram em aproximadamente oito horas toda a glicose presente no meio. Já em relação ao crescimento destas linhagens, nota-se que a PE-2 teve um crescimento menor quando comparada com a linhagem recombinante; i.e, o número de células totais foi maior nesta do que naquela. Este dado é comprovado pela OD_{600} medida em ambas as linhagens a cada 2 horas. Através desse fator, notamos que a PE-2 e a FGSC 9721 tiveram uma OD_{600} de aproximadamente 1 após 10 horas, permanecendo constante até as 24 horas; enquanto que no mesmo período citado, a linhagem transformada atingiu patamares próximos a 2 nesta categoria.

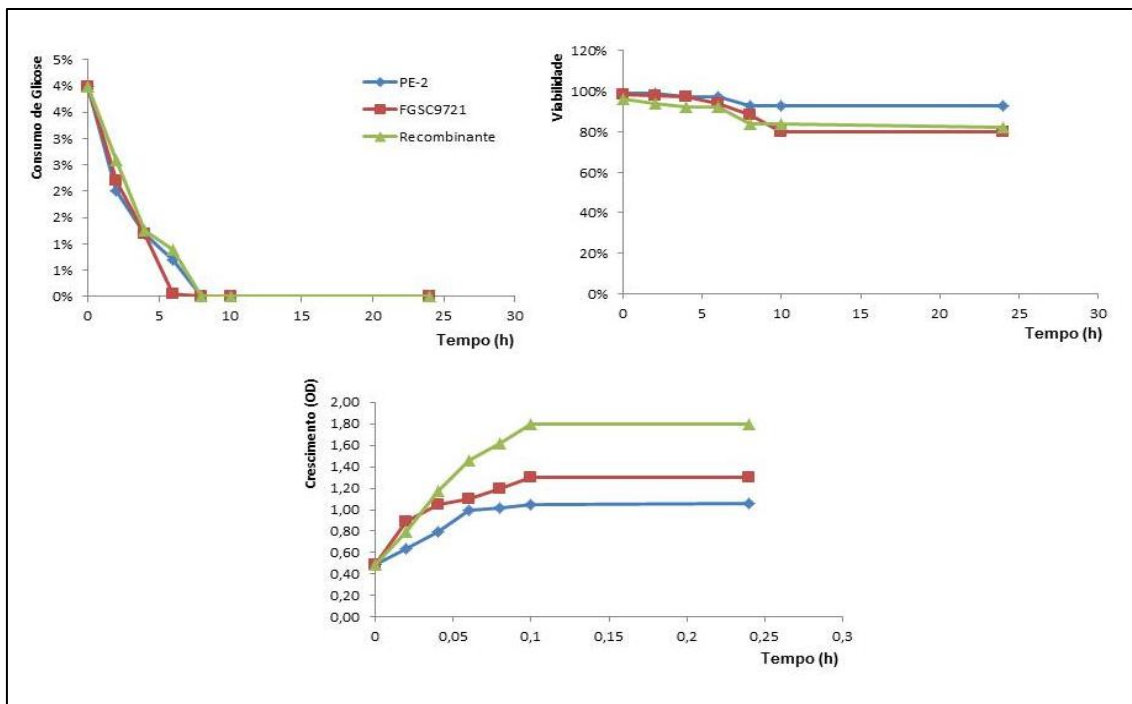


Figura 44: Comparação da análise Fermentativa da Linhagem Selvagem PE-2 com a linhagem recombinante.

O gráfico de viabilidade mostra que a linhagem PE-2 (azul) apresenta uma viabilidade ótima, com um inicial de aproximadamente 98% as 0 horas e, após 24 horas, essa taxa decai para 93%; de modo que o crescimento celular permanece crescente até as 10 horas e constante até as 24 horas após a inoculação. O consumo de glicose se faz em 8 horas, sendo que nas primeiras 2 horas este já tem redução de 50%; dado um inicial de 4% de glicose. Já no tocante a linhagem transformada (verde), há uma boa viabilidade, com um inicial de aproximadamente 97% as 0 horas e, após 24 horas, essa taxa decai para 91%; de modo que o crescimento celular permanece crescente até as 10 horas e constante até as 24 horas após a inoculação. O consumo de glicose se faz em 8 horas, sendo que nas primeiras 2 horas este já tem redução de aproximadamente 50%; dado um inicial de 4% de glicose. A linhagem controle (FGSC 9721- vermelho) mostrou similaridades com a PE-2 nas três categorias.

Em termos de viabilidade, verifica-se que não houve alterações significativas entre as linhagem, permanecendo alta mesmo após as 24 horas. Ao analisar os índices de crescimento, verifica-se que, em comparação com a FGSC 9721 e PE-2, a linhagem transformada obteve uma taxa mais alta já a partir das cinco horas. O consumo de glicose, todavia permanece igual para todas as linhagens.

Esses dados demonstram não há diferenças significativas entre as linhagens, i.e, a linhagem transformada contendo a celulase β -1,4-endoglucanase não possui qualquer

alteração que acarreta na redução de crescimento, viabilidade ou consumo de glicose. Por isso podemos afirmar que os parâmetros permanecem muito semelhantes e que provavelmente a adição desta proteína heteróloga não afeta o potencial fermentativo da linhagem. Essas análises demonstram que nossa linhagem é altamente promissora perante os dados.

Além disso, a linhagem (transformada) garante que a produção proteica só seja ativada após o total consumo de glicose. Isto só é possibilitado devido à utilização de um plasmídeo de expressão que contém o promotor ADH2, cujo qual se trata de um promotor constitutivo que é regulado pela presença de glicose; i.e, ele só é ativado quando toda a glicose for convertida em etanol. Esse fator garante que a fermentação em momento algum seja interrompida até que toda a glicose presente seja consumida e convertida em álcool e, portanto, a produção proteica só passa a ser expressa após o processo fermentativo. De mais a mais, a sobrevivência celular mostra que a quantidade de etanol presente não afeta seu crescimento e índices de sobrevivência; acrescentando mais um fator positivo para a linhagem em teste.

5.12 Sequenciamento do *pYADE4+EG*

O DNA plasmidial foi sequenciado tal que, destes, obtivemos trechos do fragmento total. Para sequenciamento foram utilizados os *primers* ADH2F e pYADE4R e *primers* específicos. Esses trechos comparados (Figura 45); mostram uma sequência consenso de 1763 pb; de modo que as 591 pb iniciais condizem com o promotor ADH2 a partir do *primer* ADH2F.

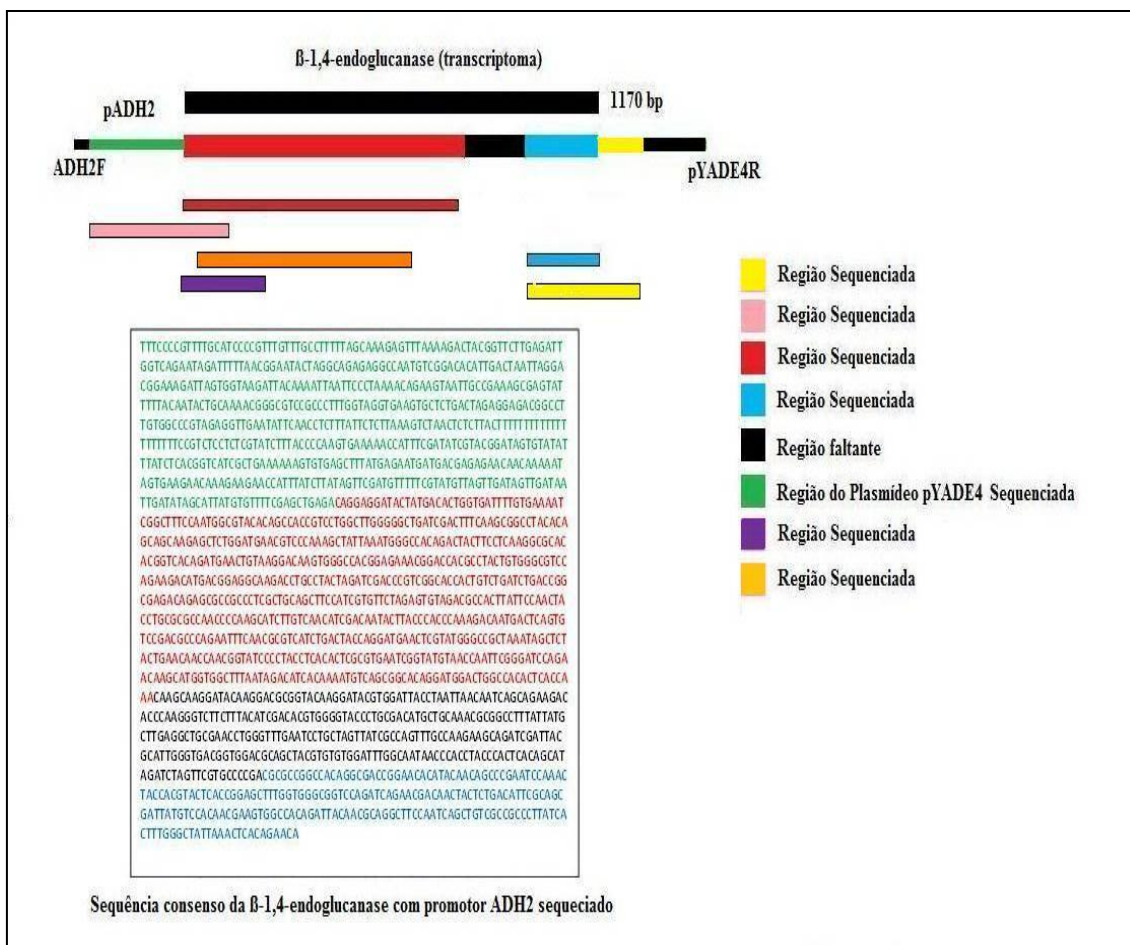


Figura 45: Sequência Consenso da β -1,4- endoglucanase de *V. heteropterus* encontrada a partir do alinhamento de trechos sequenciados a partir da construção yEG.

A partir dos trechos pudemos encontrar certa de 1763 pb; de modo que as 591 pb iniciais são referentes ao promotor ADH2. Ao alinhar os trechos encontrados a partir dos sequenciamentos, verificamos que este alinhamento confirma a ligação do fragmento de aproximadamente 1170 bp condizente com a β -1,4-endoglucanase de *V. heteropterus* com o plasmídeo de expressão pYADE4. Os trechos mostrados em vermelho, rosa e azul indicam as regiões obtidas com base no sequenciamento. Região preta não foi mostrada no sequenciamento. Região verde indica localização do ADH2 no plasmídeo pYADE4.

5.14 Predição da estrutura terciária

A predição do enovelamento ou topologia²¹¹ de uma proteína é atualmente um dos maiores problemas da Bioinformática Estrutural. O principal desafio é compreender

como a informação codificada na sequência linear de aminoácidos ou estrutura primária se traduz na estrutura 3D ou estrutura terciária de uma proteína; i.e, compreender sua configuração final. Muitas vezes essa predição pode ser realizada através da utilização de técnicas como cristalografia (por difração de raio X ou ressonância magnética nuclear); todavia, tal técnica requer um alto custo e tempo²¹². Nesse sentido, uma análise *in silico* previa se torna interessante, apesar de limitada. Isto porque através dessas metodologias pode-se obter uma predição de maneira mais rápida e correta das formas de enovelamento.

Existem diversos tipos de metodologias que podem ser empregadas nessa predição 3D *in silico*, de modo que são classificadas em dois grandes grupos: Comparativa e baseado em conhecimento. Na modelagem comparativa, uma sequência alvo é alinhada com uma sequência cuja estrutura terciária já é conhecida e armazenada no PDB. Trata-se, pois, da predição de maior acurácia nos resultados finais; uma vez que, faz uso da homologia para predizer estruturas similares ou idênticas²¹³⁻²¹⁵. Já a modelagem baseada em conhecimento faz uso de “potências estatísticas derivados da análise de padrões de enovelamento de proteínas com estruturas 3D conhecidas e armazenadas” (BUJNICKI, 2006 apud DORN, 2008)^{216,217}. E, assim, a sua estrutura pode ser predita mesmo não havendo homologia na base de dados²¹⁸, i.e, torna-se possível a obtenção de estruturas cujas formas de enovelamento não são ainda conhecidas²¹⁹.

Essa predição computacional, apesar de rápida, demanda grande complexidade e compreensão de estruturas, sejam elas primárias, secundárias ou terciárias.

Assim, a fim de se compreender melhor a função da β -1,4-endoglucanase, uma predição baseada na sequência do transcriptoma obtida a partir do cupim V. heteropterus foi realizada em comparação com outro cupim, o Nasutitermes

takasagoensis; cujo qual apresentou maior similaridade entre os nucleotídeos no alinhamento entre estes.

Inicialmente, portanto, foi realizado um alinhamento entre as sequências de aminoácidos para N. takasagoensis e V. heteropterus. (Figura 46). Neste alinhamento notamos uma identidade de 84% e *score* de 731 (máximo 1887) com 448 aminoácidos.

Sequence ID: |cl|10003 Length: 448 Number of Matches: 2

Range 1: 1 to 448 [Graphics](#) ▼ Next Match ▲ Previous Match

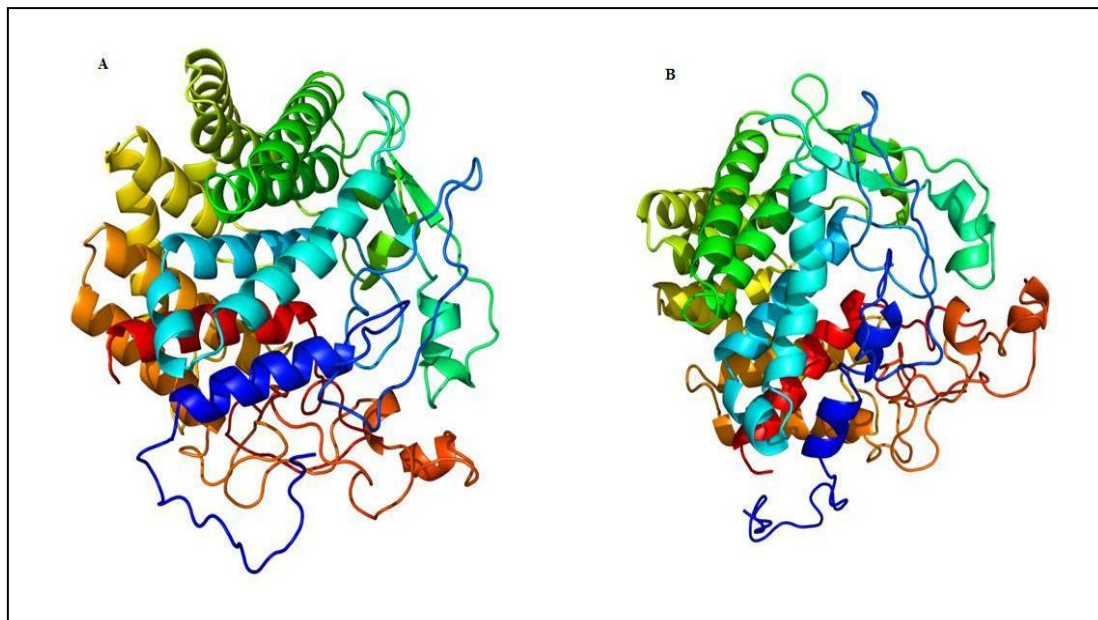
Score	Expect	Method	Identities	Positives	Gaps	Frame
731 bits(1887)	0.0	Compositional matrix adjust.	377/448(84%)	417/448(93%)	0/448(0%)	+2
Query 122		MRVFLc1l1salalcgaaYDYKQVLRDSSLFYEAQRSGRLPADQKVIWRKDSALNdgqdgq				301
Sbjct 1		M+VFLCLLSA ALCQAAAYDYKQVL +SLLFYEAQRSG LP+DQKVIWRK+SALND+G QG				60
Query 302		qdltggyYFDAGDFVKFGFPMAYTATVLAWGLIDFEAGYSSAGALDDGRKAVKWATDYFIK				481
Sbjct 61		QDLTGGY+DAGDFVKFGFPMAYTATVLAWGLIDF+A Y++AGALD+GRKA+KWATDYF+K				120
Query 482		AHTSQNEFYGVQGGDADHAFWGRPEDMTARPAYKIDTSRPGSDLAGETaaaalaaaSIV				661
Sbjct 121		AHT+ NE YGVQGG+ DHA+WGRPEDMT ARPAYKIDTS PGSDLAGETAALAAASIV				180
Query 662		FRNVDGTYSNLLTHARQLFDFANNYRGKYSDSITDARNFYASADYRDELVWAAAWLYRA				841
Sbjct 181		F++VD TYSNLLTHA+QLF+FAN+Y+GKYSDSI DA+NFIAS+DYRDELVWAAAW+Y A				240
Query 842		TNDNTYLNTAESLYDEFGLQNWGGGLNWDKSVGVQVLLAKLINKQAYKDTVQSYVNYLI				1021
Sbjct 241		T DNTYLNTAESLY++FGLQN GG NWD+K+SGVQVLLA+LINKQ YKD VQ YV+YLI				300
Query 1022		NNQKTPKGLLYIDMWGTLRHAANAFAFIMLEAAELGLSASSYRQFAQTQIDYALGDGGRS				1201
Sbjct 301		NNQKTPKGLLYID WTLRHAANAFAFIMLEAA LGL+ +SYRQFA+ QIDYALGDGGRS				360
Query 1202		FVCGFGSNPPTRPHRRSSSCPAPATCDWNTFNSPDPNYHVLVSGALVGGFPDQNDNYVDDR				1381
Sbjct 361		+VCGFG+NPPT HRSSSCP APATCDWNT+NSP+PNYHVL+GALVGGFPDQNDNY D R				420
Query 1382		SDYVHNEVATDYNAGFQSALAALVALGY	1465			
Sbjct 421		SDYVHNEVATDYNAGFQSA+AAL+ LGY	448			

Figura 46: Sequência consenso de aminoácidos alinhada com a β -1,4-endoglucanase de Nasutitermes takasagoensis.

Na imagem notamos identidade de 84 % com a β -1,4- endoglucanase da espécie Nasutitermes takasagoensis. O *Query* indica a sequência de N. takasagoensis e o *Subject*, a sequência consenso. O alinhamento se faz pelo *frame* +2.

Em seguida, uma predição 3D utilizando o programa Phyre2²⁰⁹, cujo qual faz uso de uma modelagem comparada, foi realizada, de modo que as estruturas terciárias de β -1,4-endoglucanase para ambos os cupins puderam ser comparadas (Figura 47).

Através desse programa, também, pudemos obter a estrutura secundária de cada uma (Anexo I e II).

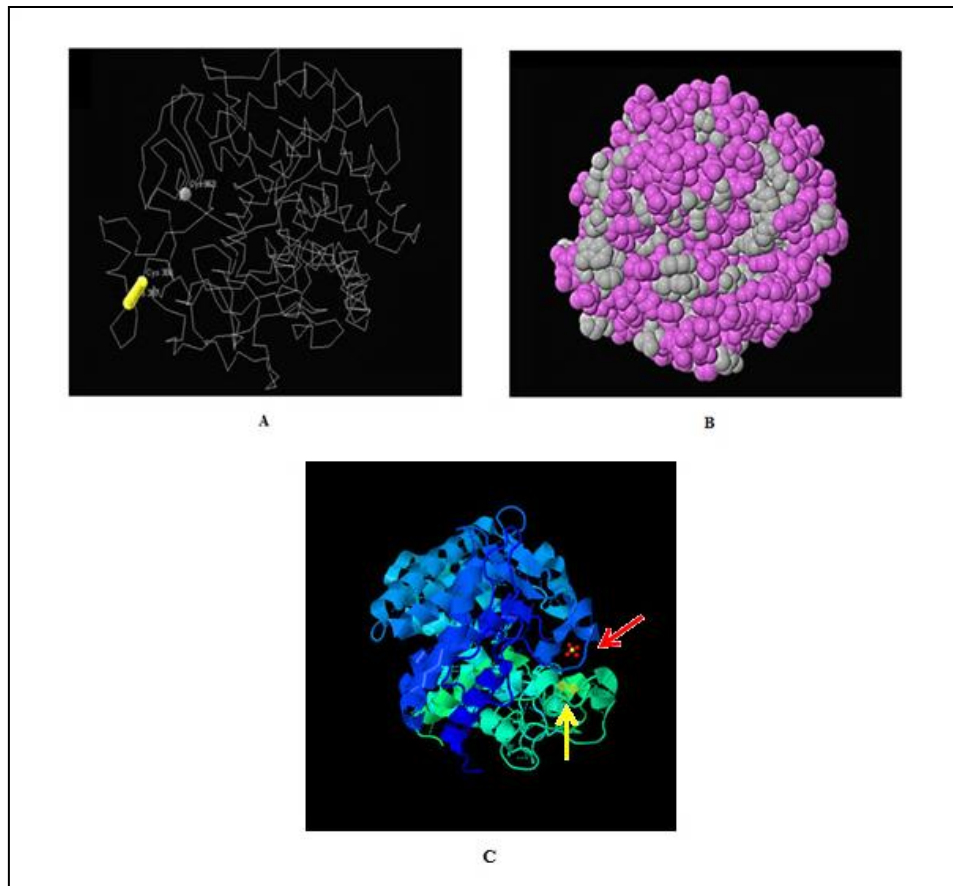


Fonte: <http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>²⁰⁹

Figura 47: Predição da estrutura terciária para β -1,4-endoglucanase.

A predição foi baseada na sequência para β -1,4-endoglucanase do transcriptoma de V. heteropterus (A) e sequência obtida em bando de dados para Nasutitermes takasagoensis (B). Predição feita pelo programa Phyre2; coloração arco-íris N→C terminal. N = vermelho; C = azul escuro.

Formada por α -hélices, cadeias β e regiões desordenadas; a configuração 3D tanto para N. takasagoensis quanto para a sequência de V. heteropterus refere-se a uma enzima com domínio catalítico de celulase; sendo formada por regiões polares e hidrofóbicas em uma proporção semelhante. O *template* d1Ks8a, com maior similaridade entre ambas (mostrado pelo Phyre2), indica a presença de oito átomos de sulfeto; uma ponte dissulfeto e três cisteínas (nas posições 363, 380 e 387) (Figura 48) e com resíduos catalíticos em mesma região (Anexo I e II); além de apresentar uma ligação sulfato.



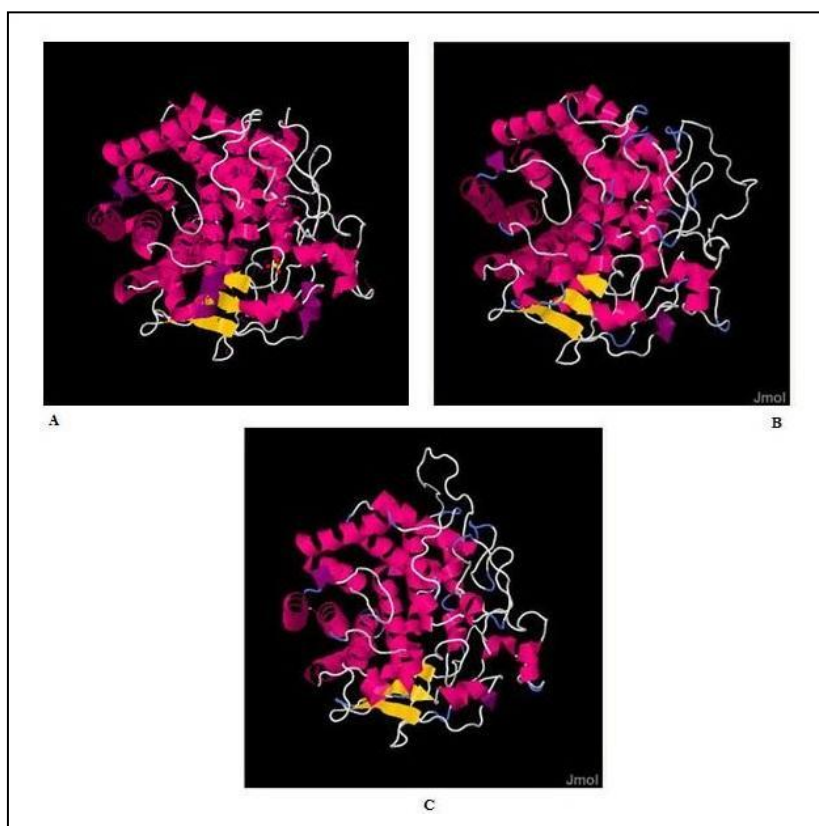
Adaptado de http://bioinformatics.org/firstglance/fgij/fg.htm?mol=http://www.sbg.bio.ic.ac.uk/phyre2/phyre2_output/46694633b85d761a/d1ks8a_1.pdb²²⁰ e <http://www.rcsb.org/pdb/explore/jmol.do?structureId=1KS8&opt=3>²²¹

Figura 48: Macromolécula de β -1,4-endoglucanase (template d1Ks8a).

(A) Conformação terciária da β -1,4-endoglucanase confere a macromolécula a presença de três cisteínas (sendo duas ligadas a ponte dissulfeto- em amarelo) e oito átomos de sulfeto. Além disso, (B), sua configuração permite a presença de regiões polares e hidrofóbicas em proporções próximas, mostradas em rosa e cinza respectivamente. Somente um lado da proteína é mostrado em B.(C) Há também a presença de uma ligação sulfato representada pela seta vermelha. A seta amarela indica a ponte dissulfeto.

Apesar da semelhança, as celulasas se diferem em relação a sua estrutura secundária e terciária (Figura 49), de modo que, comparada com o *template* d1Ks8a, a β -1,4-endoglucanase de *V. heteropterus* possui uma cobertura de 84 % contra os 97 % para *N. takasagoensis*. Essa diferença encontra-se relacionada principalmente ao

número de α -hélices presentes na estrutura, onde esta é maior em V. heteropterus estando mais próxima de d1Ks8a; e regiões desordenadas no final da molécula, presentes em formação diferenciada entre N. takasagoensis e V. heteropterus. Estes fatores conferem uma estrutura terciária com algumas divergências, vistas na figura 47 nas regiões representadas por azul marinho e laranja, principalmente (fim e meio da molécula, respetivamente).



Adaptado de <http://www.rcsb.org/pdb/explore/jmol.do?structureId=1Ks8&opt=3>²²⁰

Figura 49: Diferenças entre estrutura secundária e terciária para β -1,4-endoglucanase para V. heteropterus, N. takasagoensis e estrutura 1Ks8.

(A) *Template* 1Ks8. (B) β -1,4-endoglucanase de V. heteropterus. (C) β -1,4-endoglucanase de N. takasagoensis. As diferenças entre as três estruturas se faz, basicamente na composição de regiões desordenadas (representadas pelas linhas brancas) e α -hélices (aspirais rosas). Cadeias β estão representadas em amarelo. Essa estrutura está baseada na Estrutura de β -1,4-endoglucanase de Nasutitermes takasagoensis (1Ks8), pH 2,5 depositada no PDB. B e C não levam em consideram o pH.

Essas divergências, como já citado, podem vir a influenciar o pH e temperatura proteica, mas possivelmente, não afetam sua função dado sua similaridade e confiança de 100 %.

Considerações Finais

O etanol celulósico, ao lado do biodiesel, é uma promissora fonte de combustível sustentável e eficiente capaz de atender à demanda universal por combustíveis líquidos, tanto para veículos automotores quanto para alimentar as iminentes células de combustível. Diversos métodos para obtenção de etanol estão em experimentação e é imprescindível que o Brasil mantenha sua liderança natural neste campo para que o valor tecnológico agregado dos biocombustíveis possa ser revertido em nosso favor enquanto o produto avança para se tornar uma *commodities*. Neste caminho, a cana de açúcar desponta com larga vantagem como a planta sobre a qual devemos depositar nossos maiores esforços em curto prazo.

A bioenergia é uma medida de extrema importância para enfrentarmos os sérios desafios ambientais relacionados com os efeitos das mudanças climáticas globais. Ainda que esta não seja a única solução para este problema, certamente contribuirá para mitigar as emissões de combustíveis fósseis.

A produção de etanol celulósico com alta eficiência e sustentabilidade não será tarefa de poucos, mas resultará da integração entre diversos grupos de pesquisa especializados em diferentes áreas da fisiologia, ecologia, bioquímica, genética, enzimologia, física e engenharia, entre outras.

Esse desenvolvimento pode ser solucionado através do uso de celulases advindas de cupins. Isso porque, analogicamente, a produção de etanol pode ser comparada com o processo digestório de celulose em insetos.

O uso de transcriptomas, metabolomas e proteomas tem permitido um vasto aumento no conhecimento dos componentes dos sistemas celulósicos de insetos, bem como sua regulação. A expressão de celulases derivadas de insetos, como os cupins, em células de cultura tem promovido sucessos na característica específica e atividades dessas enzimas, bem como a identificação dos resíduos-chave para a sua atividade²²².

Devido ao baixo número de celulases de insetos serem caracterizadas quando comparadas com celulases de fungos e bactérias, algumas enzimas reportadas com novas características podem ser interessantes na produção de bioetanol.

Através da manipulação gênica podemos aumentar os níveis de expressão, atividade e termoestabilidade dessas enzimas²²³⁻²²⁶, sendo, portanto, de enorme interesse na produção de etanol de segunda geração através da biomassa de plantas, como a cana-de-açúcar.

Em adição, considerando a importância dessas celulases para sobrevivência desses insetos, estas podem vir a ser marcadores para o desenvolvimento de novas tecnológicas que visem combater estes insetos quando pragas^{227,228}.

É nesse sentido que o presente trabalho buscou a identificação, isolamento, clonagem e expressão de uma celulase (β -1,4-endoglucanase) do cupim da espécie V. heteropterus.

Testes utilizando CMC e gel de SDS mostraram que esta enzima inserida em levedura S. cerevisiae apresenta expressão considerável estando, portanto, ativa. Isso mostra indicativos iniciais para sua possível eficiência. Outros experimentos, contudo, se fazem ainda necessários. Os próximos passos (já em andamento) tem em vista a

purificação desta enzima, bem como verificação de sua expressão através de *Western blotting* e atividade biológica e enzimática (pH e temperatura ótimos, por exemplo). Experimentos visando à utilização desta enzima em linhagens industriais também vêm sendo cogitados.

A semelhança de 84 % entre a β -1,4-endoglucanase de *Nasutitermes takasagoensis* tanto em sua composição de nucleotídeos como aminoácidos e predição terciária, indicam grandes chances de sua atividade enzimática ser próxima ou igual; todavia temos que levar em consideração as alterações conformacionais na molécula final que podem vir a alterar esta atividade.

Outras enzimas, como a β -glucanase e a Exoglucanase também serão alvos de estudos futuros. Espera-se que, de posse das três enzimas, haja a formação de um complexo proteico inserido em *S. cerevisiae*, bem como análises de sua expressão mediante testes fermentativos em condições totalmente industriais e possível comercialização desta nova linhagem.

O Brasil está diante de mudança no modo de produção rara ou talvez inédita na história. Podemos desenvolver uma tecnologia de alto valor agregado e ao mesmo tempo usá-la também para recuperar a biodiversidade, integrando sustentabilidade e desenvolvimento tecnológico. Isso porque, com a implementação deste tipo de processo, algumas estimativas estabelecem que a produção de etanol no Brasil poderia ser duplicada, sem a necessidade de aumentar as áreas de cultivo da cana-de-açúcar. Sob esta panorâmica, faz-se evidente a importância do desenvolvimento de processos de produção de etanol a partir do bagaço e palha de cana, que envolvam baixos custos e altas produtividades.

Referências

1. WHEALS, A.E. *et al.* **Fuel ethanol after 25 years.** s.l.: Trends Biotechnol., 1999. v. 17, n. 12, p. 482-487.
2. WILLIAMS, P.R., *et al.* **Environmental and sustainability factors associated with next-generation biofuels in the U.S.: what do we really know?** s.l.: Environ. Sci. Technol., 2009. v. 43, n. 13, p. 4763-4775.
3. GURA, T. **Driving biofuels from field to fuel tank.** s.l.: Cell, 2009. v. 138, n.1, p. 9-12.
4. HIRAYAMA, K., H. *et al.* **Purification and characterization of termite endogenous beta-1,4-endoglucanases produced in Aspergillus oryzae.** Japão: Biosci. Biotechnol. Biochem., 2010. v. 74, n.8, p.1680-1686.
5. ROJO, F. **Biofuels from microbes: a comprehensive view.** s.l.: Microb. Biotechnol., 2008. v. 1, n. 3, p. 208-210.
6. UNICA – UNIÃO DA INDÚSTRIA DE CANA-DE-AÇÚCAR. DADOS E COTAÇÕES. 2010. Disponível em: < <http://www.unica.com.br/dadosCotacao/estatistica/> >. Acesso em: 20 abril 2011.
7. CONAB - Companhia Nacional de Abastecimento. **Acompanhamento da Safra Brasileira.** 2008. Disponível em: < <http://www.conab.gov.br> >. Acesso em janeiro de 2009.
8. MAPA – Ministério de Agricultura, Pecuária e Abastecimento. **Produção Brasileira de Cana, Açúcar e Etanol.** Brasil: Secretaria de Produção e Agroenergia. Departamento da cana-de-açúcar e Agroenergia, 2011. Disponível em: < http://www.agricultura.gov.br/arq_editor/file/Desenvolvimento_Sustentavel/Agroenergia/estatisticas/producao/ATUALIZACAO_16_03_2011/07_%20prod_cana_acucar_etanol.pdf >. Acesso em 3 abril 2011.
9. BASSO, L.C. *et al.* **Yeast selection for fuel ethanol production in Brazil.** Brasil: FEMS. Yeast. Res., 2008. v. 8, n. 7, p. 1155-1163.
10. BREMER, V.R., *et al.* **Emissions savings in the corn-ethanol life cycle from feeding coproducts to livestock.** s.l.: J Environ. Qual., 2010. v. 39, n. 2, p. 472-482.
11. LYND, L.R. *et al.* **Fuel ethanol from cellulosic biomass.** s.l.: Science, 1991. v. 151, n. 4999, p. 1318-1323.
12. LYND, L.R. *et al.* **How biotech can transform biofuels.** s.l.: Nat. Biotechnol., 2008. v. 26, n. 2, p. 169-172.
13. LEITE, R. C. C e LEAL, M. R. L. V. **O Biocombustível no Brasil.** s.l.: Cebrap, 2007. n. 78. p. 15-21.
14. MCLAUGHLIN, S.B. *et al.* **High-value renewable energy from prairie grasses.** s.l.: Environ. Sci. Technol., 2002. v. 36, n. 10, p. 2122-2129.
15. SCHMER, M.R. *et al.* **Net energy of cellulosic ethanol from switchgrass.** USA: Proc. Natl. Acad. Sci. U. S. A., 2008. v. 105, n. 2, p. 464-469.
16. SUN, Y. e CHENG, J.J. **Dilute acid pretreatment of rye straw and bermudagrass for ethanol production.** s.l.: Bioresour. Technol., 2005. v. 96, n. 14, p. 1599-1606.
17. WYMAN, C.E. **What is (and is not) vital to advancing cellulosic ethanol.** s.l.: Trends Biotechnol., 2007. v. 25, n.4, p. 153-157.

18. LEE, J. **Biological conversion of lignocellulosic biomass to ethanol**. s.l.: J. Biotechnol., 1997. v. 56, n.1, p. 1-24.
19. WYMAN, C.E. **Biomass Ethanol: Technical Progress, Opportunities, and Commercial Challenges**. s.l.: Annu. Rev. Energy. Environ., 1999. n. 24, p. 189-226.
20. HAHN-HAGERDAL, B., *et al.* **Bio-ethanol--the fuel of tomorrow from the residues of today**. s.l.: Trends Biotechnol., 2006. v. 24, n. 12, p. 549-556.
21. OLSSON, L., *et al.* **Separate and simultaneous enzymatic hydrolysis and fermentation of wheat hemicellulose with recombinant xylose utilizing *Saccharomyces cerevisiae***. s.l.: Appl. Biochem. Biotechnol., 2006. v. 129, n. 132, p. 117-129.
22. CHERRY, J.R e FIDANTEF, A. L. **Directed evolution of industrial enzymes: an update**. California: Current Opinion in Biotechnology 2004. v. 14, p. 438-443.
23. MASIERO, G. e LOPES, H. **Etanol e biodiesel como recursos energéticos alternativos: perspectivas da América Latina e da Ásia**. Brasília: Rev. Bras. Polít. Int., 2008. v. 51, n.2 , p. 60-79.
24. de OLIVEIRA, F.M.V. **Avaliação de diferentes pré-tratamentos e deslignificação alcalina na sacarificação da celulose de palha de cana**. Lorena: Universidade de São Paulo, 2010. [Dissertação de mestrado].
25. SANTIAGO, A.D. *et al.* **Impulsionando a produtividade e a produção agrícola da cana-de-açúcar no Brasil**. s.l.: International Workshop on Tropical Agriculture Development, 2006.
26. FAO. **Sugar: Global market analysis**. Roma: Food and Agriculture Organization of the United Nations, 2007.
27. BNDES – Banco Nacional de Desenvolvimento Econômico e Social. **Bioetanol de cana-de-açúcar: Energia para o desenvolvimento sustentável**. Rio de Janeiro, s.e.: 2008. 316p.
28. DAI, D. *et al.* **Energy efficiency and potentials of cassava fuel ethanol in Guangxi region of China**. s.l.: Energy Conversion and Management, 2006. v. 47.
29. EBAMM. **ERG Biofuels Analysis Meta-Model**. Release 1.0. Berkeley: University of California – Energy and Resources Group, 2005.
30. IEA – INTERNATIONAL ENERGY AGENCY. **Biofuels for transport: an international perspective**. Paris: International Energy Agency, 2004
31. NGUYEN, T. L. T. *et al.* **Energy balance and GHG-abatement cost of cassava utilization for fuel ethanol in Thailand**. s.l.: Energy Policy, 2007. v. 35, n. 9.
32. VÁSQUEZ, M.P. *et al.* **Enzymatic hydrolysis optimization to ethanol production by Simultaneous Saccharification and Fermentation**. s.l.: Applied Biochemistry and Biotechnology, 2007. n. 12, p. 137-140.
33. MARQUES, Fabrício. **O alvo é o bagaço**. Brasil: Pesquisa online Fapesp, 2008. Disponível em: < <http://revistapesquisa.fapesp.br/?art=3934&bd=1&pg=1> >. Acesso em: 24 abril 2010.
34. SCHLITTLER, L. A. F. S. **Engenharia de um bioprocesso para a produção de etanol de cana-de-açúcar**. Rio de Janeiro: UFRJ- Escola de Química, 2006. [Tese de Mestrado].
35. PANDEY, A. *et al.* Biotechnological potential of agro-industrial residues. In: _____. **Sugarcane bagasse**. s.l.: Bioresource Technology, 2000. v. 74, n. 1, p. 69-80.

36. ANSELMO FILHO, P. e BADR, O. **Biomass resources for energy in North-Eastern Brazil**. s.l: Applied Energy, 2004. v. 77, n. 1, p. 51-67.
37. LAMONICA, H. M. Potencial de geração de excedentes de energia elétrica a partir da biodigestão da vinhaça. In: _____. **VI Congresso Internacional sobre Geração Distribuída e Energia no Meio Rural – Agrener 2006**. Brasil: Unicamp, 2006.
38. FROLLINI, F.; PIMENTA, M.S.A. **Lignin**: Utilization as a macromonomer in the synthesis of phenolic type resins. Brasil: Anais da Associação Brasileira de Química, 1997. v. 46, n. 1, p. 43-49.
39. LAVARACK, B.P.; GRIFFIN, G.J.; RODMAN, D. **The acid hydrolysis of sugarcane bagasse hemicellulose to produce xylose, arabinose, glucose and other products**. s.l.: Biomass and Bioenergy, 2002. v. 23, n. 5, p. 367-380.
40. MOSIER, N. *et al.* **Features of promising technologies for pretreatment of lignocellulosic biomass**. s.l: Bioresource Technology, 2005. v. 96, n. 6, p. 673-686.
41. CRAWFORD, D.L. *et al.* **Production of microbial protein from waste cellulose by Thermomonospora fusca, a thermophilic actinomycete**. s.l: Biotechnology and Bioengineering, 1973. v. 15, n. 5, p. 833-843.
42. JANSEN, N.B.; FLICKINGER, M.C. e TSAO, G.T. **Production of 2,3-butanediol from D-xylose by Klebsiella oxytoca ATCC 8724**. s.l.: Biotechnology and Bioengineering, 1984. v. 26, n. 4, p. 362-369.
43. RODRIGUES, G.R. **Prospecção de Genes de Celulase Presentes em Biblioteca Metagenômica**. Jaboticabal: Universidade Estadual Paulista, 2008. [tese de mestrado].
44. SARROUH, B.F. *et al.* **Technical/Economical Evaluation of Sugarcane Bagasse Hydrolysis for Bioethanol Production**. s.l.: Chemical Engineering & Technology, 2007. v. 30, n. 2, p. 270-275.
45. FAIRBANKS, M. **Álcool**: Tecnologia amplia oferta e reduz custo do álcool. Brasil: Res. Química e Derivados, 2003. n. 417.
46. MUSTAFA, B.; HAVVA, B. e CAHIDE, O. **Progress in bioethanol processing**. Turkey: Progress Energy Combustion Sci., 2008. n. 34, p. 551-573.
47. POPPE, M. K.; PERICO, A. C. S. **Bioetanol combustível: uma oportunidade para o Brasil**. Brasília: CGEE, 2009. p. 129-136.
48. ANDERSON, W. F. e AKIN, D. E. **Structural and chemical properties of grass lignocelluloses related to conversion for biofuels**. s.l.: J Ind Microbiol Biotechnol., 2008. n. 35, p.355–366.
49. FURLANI NETO, V.L.; RIPOLI, T.C.; VILA NOVA, N.A. **Biomassa de cana-de-açúcar: energia contida no palhicho remanescente de colheita mecânica**. s.l: Stab - Açúcar, Álcool e Subprodutos, 1997.v.15, p.24-27.
50. ROSSELL, C. E. V. **Hidrólise de bagaço de cana-de-açúcar para obtenção de etanol**. Campinas: UNICAMP, s.d. Disponível em: < http://www.fiesp.com.br/agencianoticias/2007/05/15/carlos_eduardo_rossell.pdf >. Acesso em: 29 abril 2011.
51. MACEDO, I.C.; SEABRA, J.E.A.; SILVA, J.E.A.R, S. **Greenhouse gases emissions in the production and use of ethanol from sugarcane in Brazil: The 2005/2006 averages and a prediction for 2020**. s.l.: Biomass and Bioenergy, 2008. v. 32, p. 582-595.

52. AUDE, M.I.S. *et al.* **Manejo do palhico da cana-de-açúcar: efeito na produção de colmos industrializáveis e outras características agrônômicas.** Brasil: Ciência Rural, 1993. v.23, p.281-286.
53. TRIVELIN, P.C.O.; VICTORIA, R.L.; RODRIGUES, J.C.S. **Aproveitamento por soqueira de cana-de-açúcar de final de safra do nitrogênio da aquamônia-15N e uréia-15N aplicado ao solo em complemento à vinhaça.** Brasil: Pesquisa Agropecuária Brasileira, 1995. v.30, p.1375-1385.
54. PESQUISA FAPESP, 2008. **Aproveitamento total.** Brasil: Revista Pesquisa Fapesp 2009. n. 154. Disponível em: < <http://www.revistapesquisa.fapesp.br/arq/r/pt/> >.. Acesso em: 26 abril 2011.
55. ANDERSON, W. F. e AKIN, D. E. **Structural and chemical properties of grass lignocelluloses related to conversion for biofuels.** s.l.: J Ind Microbiol Biotechnol., 2008. n. 35, p.355–366.
56. RAGAUSKAS, A. *et al.* **The path forward for biofuels and biomaterials.** s.l.: Science, 2006. n. 311, p. 484–489.
57. SCHARF, M. E. e TARTAR, A. **Termite digestomes as sources for novel lignocellulases.** Florida: Biofuels, Bioprod. Bioref., 2008. n. 2; p. 540-552.
58. LANGE, J. P. **Lignocellulose conversion: an introduction to chemistry, process and economics.** s.l.: Biofuels Bioprod Bioref, 2007. n. 1, p. 39–48.
59. KUMAR, R.; SINGH, S. SINGH, O.V. **Bioconversion of lignocellulosic biomass: biochemical and molecular perspectives.** India: J Ind Microbiol Biotechnol., 2008. n. 35, p. 377-391.
60. ALBERTS, B. *et al.* **Junção Celular, Adesão Celular e Matriz Extracelular.** In: _____. **Biologia molecular da célula.** 4 ed. Porto Alegre: Artmed, 2002. cap. 19, p. 1118-1124.
61. NEUREITER, M.; *et al.* **Dilute-acid hydrolysis of sugarcane bagasse at varying conditions.** s.l.: Applied Biochemistry and Biotechnology, 2002. v. 98, p. 49-58.
62. GOLDSTEIN, I. S. **Organics Chemicals from Biomass.** Florida.: Boca Raton, 1981. Cap. 1, p. 10-15.
63. FENGEL, D. e WEGENER, G. **Wood: Chemistry, ultrastructure, reactions.** New York: Walter de Gruyter, 1989.
64. RAVEN, P. H.; EVERT, R. F.; EICHCHORN, S. E. **Biologia Vegetal.** 6 ed. Rio de Janeiro: Guanabara Koogan, 2001.
65. PITARELO, A.P. **Avaliação da Susceptibilidade do bagaço e da palha de cana-de-açúcar à bi conversão via pré-tratamento a vapor e hidrólise enzimática.** Curitiba: Universidade Federal do Paraná, 2007. [Tese de Mestrado].
66. SAAD, M.B.W. **Avaliação Técnica e Econômica Preliminar da Produção de Etanol via Hidrólise Enzimática de Bagaço de Cana-de-açúcar.** Lorena: Universidade de São Paulo, 2010. [Tese de Mestrado]
67. LYND, L.R. **Overview and evaluation of fuel ethanol from cellulosic biomass: Technology, Economics, the Environment and Policy.** New Hampshire: Annu. Rev. Energy. Environ., 1996. n.21, p. 403-465.

68. CANETTIERI, E. V.; SILVA, J. B. A. E.; FELIPE, M. G. A. **Application of factorial design to the study of xylitol production from eucalyptus hemicellulosic hydrolysate**. s.l.: Applied Biochemistry and Biotechnology, 2001. v. 90, p. 159-168.
69. ARANTES, V. e SADDLER, J.N. **Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis**. s.l.: Biotechnology for Biofuels, 2010. v. 3, n. 4. Disponível em: < www.biotechnologyforbiofuels.com/content/3/1/4 >. Acesso em: 30 abril 2011.
70. TOKUDA, G. *et al.* **Cellulose Digestion in the Wood-Eating Higher Termite, *Nasutitermes takasagoensis* (Shiraki)**: Distribution of Cellulases and Properties of Endo- β -1,4-glucanase. Japão: Zoological Science, 1997. n. 14, p. 83-93.
71. ALBERTS, B. *et al.* A parede celular das plantas. In: _____. **Biologia Molecular da Célula**. 5 ed. Porte Alegre: Artmed, 2010. P. 1197-1204.
72. MORAIS, S.A.L.; NASCIMENTO, E.A.; MELO, D.C. Chemical analysis of *Pinus oocarpa* Wood. In: _____. **Quantification of macromolecular components and volatile extractives**. s.l.: Revista Árvore, 2005. v. 29, n. 3, p. 461-470.
73. FAN, L.T.; GHARPURAY, M.M.; LEE, Y.H. **Cellulose hydrolysis**. New York: Springer-Verlag, 1987.
74. WYMAN, C. E. *et al.* **Coordinated development of leading biomass pretreatment technologies**. s.l.: Bioresource Technology, 2005. v. 96, p. 1959-1966.
75. WATANABE, H. e TOKUDA, G. **Cellulolytic Systems in Insects**. Japão: Annu. Rev. Entomol., 2010. n. 55, p. 609-632.
76. KOOTSTRA, A.M.J. *et al.* **Optimization of the dilute maleic acid pretreatment of wheat straw**. s.l.: Biotechnology for Biofuels, 2009. v. 2, n. 31. Disponível em: < www.biotechnologyforbiofuels.com/content/2/1 >. Acesso em: jan de 2010.
77. RAMOS, L.P. **The chemistry involved in the pretreatment of lignocellulosic materials**. s.l.: Química Nova, 2003. v. 26, p. 863-871.
78. MUSSATTO, S.I.; SANTOS, J.C. e ROBERTO, I.C. **Effect of pH and activated charcoal adsorption on hemicellulosic hydrolysate detoxification for xylitol production**. s.l.: Journal of Chemical Technology & Biotechnology, 2004. v. 79, n. 6, p. 590-596.
79. SENDELIUS, J. **Steam Pretreatment Optimisation for Sugarcane Bagasse in Bioethanol Production**. Sweden: Master of Science - Department of Chemical Engineering, Lund University, 2005.
80. MARX, D. E. **Ativadores Oxidantes TAMLTM: Agentes Branqueadores Verdes para a Indústria de Papel**. Scranton: Green Chemistry, s.d. Disponível em: < <http://academic.scranton.edu/faculty/cannm1/inorganic/inorganicmoduleport.html> >. Acesso em: 1 maio 2011.
81. CLAASSEN, P. A. M. *et al.* **Utilization of biomass for the supply of energy carriers**. s.l.: Applied Microbiology and Biotechnology, 1999. v. 52, p. 741-755.
82. CHANG, V. S. e HOLTZAPPLE, M. T. **Fundamental factors affecting enzymatic reactivity**. s.l.: Applied Biochemistry Biotechnology, 2000. p. 5-37.

83. CHENG, K. *et al.* **Sugarcane bagasse hemicellulose hydrolysate for ethanol production by acid recovery process.** s.l.: Biochemical Engineering Journal, 2008. v. 38, p. 105-109.
84. HU, T.Q. **Chemical modification, properties, and usage of lignin.** New York: Kluwer Academic/Plenum Publishers, 2002.
85. SÁNCHEZ, O.J. e CARDONA, C.A. **Trends in biotechnological production of fuel ethanol from different feedstocks.** s.l.: Bioresource Technology, 2008. v. 99, p. 5270-5295
86. ANDERSON, W. F. e AKIN, D. E. **Structural and chemical properties of grass lignocelluloses related to conversion for biofuels.** s.l.: J Ind Microbiol Biotechnol., 2008. n. 35, p.355–366.
87. PERLACK, R.D. *et al.* **Biomass as feedstock for bioenergy and bioproducts industry: The technical feasibility of a Billion-Ton Annual Supply.** Oak Ridge/TN: Oak Ridge National Laboratory, 2005.
88. BUCKERIDGE, M.S.; dos SANTOS, W.D. e de SOUZA, A.P. **A rota para o etanol celulósico no Brasil.** São Paulo: Universidade de São Paulo, 2010. Disponível em: < <http://mundodacana.blogspot.com/2010/08/as-rotas-para-o-etanol-celulosico-no.html> >. Acesso em 9 maio 2011.
89. DIONISIO, G.F. *et al.* **Aplicação de complexo enzimático lignocelulósico para a hidrólise de bagaço de cana-de-açúcar.** Minas Gerais: VIII Congresso Brasil. Eng. Quim. em IC, 2009.
90. da SILVA, V.F.N. **Estudos de pré-tratamento e sacarificação enzimática de resíduos agroindustriais como etapas no processo de obtenção de etanol celulósico.** Lorena: Universidade de São Paulo, 2009. [Tese de Mestrado]
91. DIEN, B. S.; COTTA, M. A.; JEFFRIES, T. W. **Bacteria engineered for fuel ethanol production: current status.** s.l.: Applied Microbiology and Biotechnology, 2003. v.63, p.258–266.
92. JEFFRIES, T. W. **Engineering yeasts for xylose metabolism.** s.l.: Current Opinion in biotechnology, 2006. v.17, p.1–7.
93. HAHN-HÄGERDAL, B.; *et al.* **Bioethanol: the fuel of tomorrow from the residues of today.** s.l.: Trends in Biotechnology, 2006.v.24.
94. **Biomass Research and Development Technical Advisory Committee.** Estados Unidos: Roadmap for Biomass Technologies in the United States, 2002. Disponível em: < <http://www.brdisolutions.com/pdfs/FinalBiomassRoadmap.pdf> >. Acesso em 28 jan 2009.
95. EERE (Office of Energy Efficiency and Renewable Energy and Office of Science). **Biomass Research and Development Board.** 2006. Disponível em: < <http://www.eere.energy.gov/> >. Acesso em 15 fev 2009.
96. ZHANG, Y. H. P. **Reviving the carbohydrate economy via multi-product lignocellulose biorefineries.** s.l.: Journal of Industrial Microbiology and Biotechnology, 2008. v. 35, p. 367-375.
97. ZHANG, Y. H. P. *et al.* **Fractionating recalcitrant lignocellulose at modest reaction conditions.** s.l.: Biotechnology and Bioengineering, 2007. v. 97, p. 214-223.

98. CHANG, V. S. e HOLTZAPPLE, M. T. **Fundamental factors affecting enzymatic reactivity**. s.l.: Applied Biochemistry Biotechnology, 2000. p. 5-37.
99. LAUREANO-PEREZ, L. *et al.* **Understanding factors that limit enzymatic hydrolysis of biomass**. s.l.: Applied Biochemistry and Biotechnology, 2005. p. 1081-1099.
100. HENDRIKS, A. T. W. M. e ZEEMAN, G. **Pretreatments to enhance the digestibility of lignocellulosic biomass**. s.l.: Bioresource Technology, 2009. v. 100, p. 10-18.
101. KIM, S. e HOLTZAPPLE, M. T. **Effect of structural features on enzyme digestibility of corn stover**. s.l.: Bioresource Technology, 2006. v. 97, p. 583-591.
102. HIMMEL, M. E. *et al.* **Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production**. s.l.: Science, 2007. v. 315, p. 804-807.
103. SUN, Y. e CHENG, J. **Hydrolysis of lignocellulosic materials for ethanol production: a review**. S.l.: Bioresource Technology, 2002. n. 83, p. 1-11.
104. FERRAZ, A. *et al.* **The use of white-rot decaying as a pretreatment for organosolv delignification of *Eucalyptus grandis* wood**. Canada: Workshop on Applications of Biotechnology in Bioenergy Systems, Anais...,1994.
105. CARRASCO, F. **Thermo-mechano-chemical pretreatment of wood in a process development unit**. s.l.: Wood Science and Technology, 1992. v. 26, p. 413-428.
106. GALBE, M. e ZACCHI, G. **A review of the production of ethanol from softwood**. s.l.: Applied Microbiology Biotechnology, 2002. p. 618-628.
107. ZHOU, X., *et al.* **Production and characterization of a recombinant beta-1,4-endoglucanase (glycohydrolase family 9) from the termite *Reticulitermes flavipes***. s.l.: Arch Insect Biochem Physiol., 2010. v. 74, n. 3, p. 147-62.
108. KLYOSOV, A.A. **Trends in biochemistry and enzymology of cellulose degradation**. S.l.: Biochemistry, 1990. V. 29, p. 10577-10585.
109. WILLIS, J. D; OPPERT, C. e JURAT-FUENTES, J. L. **Methods for discovery and characterization of cellulosic enzymes from insects**. Tennessee: Insect Science, 2010. p. 1-15.
110. TOMME, P., WARREN, R. A. e GILKES, N. R. **Cellulose hydrolysis by bacteria and fungi**. s.l.: Adv. Microb. Physiol 1995. n. 37, p. 1-81.
111. HENRISSAT, B. **A classification of glycosyl hydrolases based on amino acid sequence similarities**. s.l.: Biochem. J., 1991. n. 280, p.309-316.
112. HENRISSAT, B. & BAIROCH, A. **New families in the classification of glycosyl hydrolases based on amino acid sequence similarities**.s.l.: Biochem. J. ,1993. n. 293, p. 781-788.
113. LO, N. WATANABE, H. e SUGIMURA, M. **Evidence for the presence of a cellulase gene in the last common ancestor of bilaterian animals**. s.l.: Proc. R Soc. Lond. B. Biol. Sci., 2003. n.270, p. 69-72.
114. DAVISON, A. e BLAXTER, M. **Ancient origin Of glycosyl hydrolase family 9 cellulase genes**. s.l: Mol. Biol. Evol., 2005. n. 22, p. 1273-1284.

115. SMANT, G., J. *et al.* **Endogenous cellulases in animals: isolation of beta-1,4-endoglucanase genes from two species of plant-parasitic cyst nematodes.** USA: Proc. Natl. Acad. Sci., 1998. n. 95, p. 4906–4911.
116. ROSSO, M. N., *et al.* **Isolation of a cDNA encoding a beta-1,4-endoglucanase in the root-knot nematode Meloidogyne incognita and expression analysis during plant parasitism.** s.l.: Mol. Pl. Microbe Interact., 1999. n. 192, p. 585–591.
117. WATANABE, H. *et al.* **A cellulose gene of termite origin.** s.l.: Nature, 1998. n. 394, p. 330–331.
118. TOKUDA, G., *et al.* **Metazoan cellulase genes from termites: intron/exon structure and sites of expression.** s.l.: Biochem. Biophys. Acta, 1999. N. 1447, p. 146–159.
119. LO, N., *et al.* **Evidence from multiple gene sequences indicates that termites evolved from wood-feeding cockroaches.** s.l.: Curr. Biol., 2000. n. 10, p. 801–804.
120. BYRNE, K. A., *et al.* **Isolation of a cDNA encoding a putative cellulase in the red claw crayfish Cherax quadricarinatus.** s.l.: Gene, 1999. n. 239, p. 317–324.
121. GIRARD, C. E JOUANIN, L. **Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, Phaedon cochleariae.** s.l.: Insect Biochem. Mol. Biol., 1999. n. 29, p. 1129–1142.
122. XU, B., JANSON, J. C. e SELLOS, D. **Cloning and sequencing of a molluscan endo-beta-1,4-glucanase gene from the blue mussel, Mytilus edulis.** s.l.: Eur. J. Biochem., 2001. n. 268, p. 3718–3727.
123. KHAMBHAMPATI, S. **A phylogeny of cockroaches and related insects based on DNA sequence of mitochondrial ribosomal RNA genes.** USA: Proc. Natl. Acad. Sci, 1995. n. 92, p. 2017-2020.
124. CLEVELAND, L. R. **Symbiosis between termites and their intestinal protozoa.** USA: Proc. Natl. Acad. Sci, 1923. n. 9, p. 424-428.
125. CLEVELAND, L. R. **The physiological and symbiotic relationships between the intestinal protozoa of termites and their host, with special reference to Reticulitermes flavipes Kollar.** s.l: Biol. Bull. Mar. Biol. Lab., 1924. n. 36., p. 117-227.
126. CLEVELAND, L. R. **The ability of termites to live perhaps indefinitely on a diet of pure cellulose.** s.l: Biol. Bull., 1925. n. 48, p. 289-293.
127. CLEVELAND, L. R. **The wood-feeding roach Cryptocercus, its protozoa, and symbiosis between protozoa and roach.** s.l: Mem. Am. Acad. Artd. Sci., 1934. n. 17, p. 185-342.
128. PRINKS, R. A. e KREULEN, D. A. **Comparative aspects of plant cell wall digestion in insects.** s.l: Anim. Feed Sci. Technol., 1991. n.32, p. 101-118.
129. KUKOR, J. J. e MARTIN, M. M. **Cellulose digestion in Monochamus marmorator Kby (Coleoptera: Cerambycidae).** s.l: Science, 1991. n. 220, p. 1161-1163.
130. MARTIN, M. M. **Invertebrate-Microbial Interaction: Ingested Fungal Enzymes in Arthropod Biology.** Nova York: Comstock, 1987.

131. MARTIN, M. M. **The evolution of insect-fungus associations**: from contact to stable symbiosis. s.l.: Am. Zool., 1992. n. 32, p. 593-605.
132. WATANABE H. e TOKUDA, G. **Animal cellulases**. Basel: CMLS, Cell. Mol. Life Sci., 2001. n.58, p. 1167-1178.
133. DEHAL, P *et al.* **The draft genome of Ciona intestianlis**: insights into chordate and vertebrate origins. s.l.: Science, 2002. n. 298, p. 2157-2167.
134. PENNISI, E. **Tunicate genome shows a little backbone**. s.l.: Science, 2002. n. 298, p. 2111-2112.
135. SCHOLL, E. H. *et al.* **Horizontally transferred genes in plant-parasitic nematodes**: a high-throughput genomic approach. s.l.: Gemone Biol., 2003. n. 4, p. 39.
136. YAN, Y., *et al.* **Genomic organization of four b-1,4-endoglucanase genes in plant-parasitic cyst nematodes and its evolutionary implications**.s.l.: Gene, 1998. n. 220, p. 61-70.
137. PARNAS, I. **The cellulolytic activity in the snail, Levantina hierosolyma Boiss**. s.l.: J. Cell. Comp. Physiol., 1961. n.58, p.195-201.
138. STRADINE, G. A. e WHITAKER, D. R. **On the origin of the cellulase and chitinase of Helix pomatia**. s.l.: Can. J. Biochem.Physiol., 1963. n. 41, p. 1621-1626.
139. MARSHALL, J. J. **Purification of a b-1,4-glucan hydrolase (cellulase) from the snail, Helix pomatia**. s.l; Comp. Biochem. Physiol., 1973. n. 44, p. 981-988.
140. MARSHALL, J. J. e GRAND, R. J. A. **Characterization of a b-1,4-glucan hydrolase from the snail, Helix pomatia**. s.l.: Comp. Biochem. Physiol., 1976. n. 53, p. 231-237.
141. MAEDA, I. *et al.* **Purification and characterization of a cellulase from the giant snail Achatina fulica**. s.l.: Biosci. Biotech. Biochem., 1996. n. 60, p.122-124.
142. ELYAKOVA, L. A., SOVA, V. V. e VASKOVSHY, V. E. **Cellulase of marine mollusc Littorina sp.** Acta: Biochim. Biophys., 1968. n. 167, p. 462-464.
143. YOKOE, Y. e YASUMASU, I. **The distribution of cellulose in invertebrates**. s.l.: Comp. Biochem. Physiol., 1964. n. 63, p. 323-338.
144. GREENE, R. V., GRIFFIN, H. L. e FREER, S. N. **Purification and characterization of an extracellular endoglucanase from the marine shipworm bacterium**. s.l.: Arch. Biochem. Biophys., 1988. n. 267, p. 334-341.
145. GRIFFIN, H. L.; FREER, S. N. e GREENE, R. V. **Extracellular endoglucanase activity by a novel bacterium isolated from marine shipworm**. s.l: Biochem. Biophys. Res. Commun., 1987. n. 144, p. 143-151.
146. IMAN, S. H.; GREENE, R. V. e HOCKRIDGE, M. E. **Zymographic analyses of carboxymethylcellulases secreted by the bacterium from wood-boring marine shipworms**. s.l.: Biotechnol. Tech., 1993. n.7, p. 579-584.
147. DISTEL, D. L. e ROBERTS, S. J. **Bacterial endosymbionts in the gills of the deep-sea wood-boring bivalves Xylophaga atlantica and Xylophaga washingtona**. s.l.: Biol. Bull., 1997. n. 192, p. 253-261.
148. XU, B. *et al.* **Purification, characterization and amino-acid sequence analysis of a thermostable, low molecular mass endo-b-1,4-glucanase from blue mussel, Mytilus edulis**. s.l.: Eur. J. Biochem., 2000. n. 267, p. 4970-4977.

149. WATERBURY, J. B. **A cellulolytic nitrogen-fixing bacterium cultured from the gland of Deshayes in shipworms (*Bivalvia: Teredinidae*)**. s.l.: Science, 1983. n. 221, p. 1401-1403.
150. NISHIDAA, Y. *et al.* **Isolation and primary structure of a cellulase from the Japanese sea urchin *Strongylocentrotus nudus***. Japão: Hokkaido University, 2007.
151. ANZAI, H., NISIZAWA, K. e MATSUDA, K. **Purification and characterization of a cellulase from *Dolabella auricularia***. Tokyo: J. Biochem., 1984. n.96, p. 1381-1390.
152. BRUSCA, R. C. e BRUSCA, G. J. **Invertebrates**. EUA: Sinauer, 1990.
153. NAKASHIMA, K *et al.* **The evolutionary origin of animal cellulose synthase**. s.l.: Dev. Genes Evol., 2004. n. 214, p. 81-88.
154. MATTHYSSE, A.G. *et al.* **A functional cellulose synthase from ascidian epidermis**. USA: Proc. Natl. Acad. Sci., 2004. n.101, p. 986-991.
155. LO, N. WATANABE, H. e SUGIMURA, M. **Evidence for the presence of a cellulase gene in the last common ancestor of bilaterian animals**. s.l.: Proc. R Soc. Lond. B. Biol. Sci., 2003. n.270, p. 69-72.
156. MILANO, S. e FONTES, L. R. **Online Termites Database**. Disponível em: < <http://www.cupim.net/> >. Acesso em: 23 ago 2009.
157. CONSTANTINO, R. **On-Line Termites Database**. Disponível em: < <http://www.unb.br/ib/zoo/docente/constant/catal/catnew.html> >. Acesso em: 28 jan 2010.
158. COSTA-LEONARDO, A.M. **Biological data for the Neotropical termite *Nasutitermes globiceps* (Isoptera: Termitidae: Nasutitermitinae)**. Brasil: Sociobiology, 2000. v. 36, n.2, p. 63-78.
159. **Cupins**. Disponível em: < <http://ciencia.hsw.uol.com.br/cupim2.htm> >. Acesso em: 14 out 2009.
160. MIURA, T. **Developmental regulation of caste-specific characters in social-insect polyphenism**. Indiana: Evolution & Development, 2005. vol. 7, n. 2, p. 122 – 129.
161. TSUCHIYA, M.; WATANABE, D. e MAEKAWA, K. **Effect on mandibular length of juvenile hormones and regulation of soldier differentiation in the termite *Reticulitermes speratus* (Isoptera: Rhinotermitidae)**. Japão: Appl. Entomol. Zool., 2008. vol. 43, n. 2, p. 307 – 314. Disponível em: < <http://odokon.org/> >. Acesso em: 14 out 2009.
162. HAYASHI, Y.; LO, N.; MIYATA, H. e KITADE, O. **Sex-Linked Genetic Influence on Caste Determination in a Termite**. Science, 2007. vol. 318, p. 985 – 987. Disponível em : < <http://www.sciencemag.org/cgi/collection/genetics> >. Acesso em : 14 out 2009.
163. CROZIER, R.H e SCHULÜNS, H. **Genetic caste determination in termites: out of the shade but not from Mars**. Australia: Wiley Periodicals, Inc., 2008. BioEssays. vol. 30, p. 299 – 302. Disponível em: < www.interscience.wiley.com >. Acesso em: 14 out 2009.
164. ROUX, E.A. e KORB, J. **Evolution of eusociality and the soldier caste in termites: a validation of the intrinsic benefit hypothesis**. Germany: J . Evol. Biol., 2004. v.17, p. 869-875.

165. THORNE, B.L.; BREISCH, N.L. e MUSCEDERE, M.L. **Evolution of eusociality and the soldier caste in termites**: Influence of intraspecific competition and accelerated inheritance. Maryland: Department of Entomology, University of Maryland, 2003. vol. 100, n. 22, p. 12808 – 12813. Disponível em : < www.pnas.org/cgi/doi/10.1073/pnas.2133530100 >. Acesso em: 14 out 2009.
166. NOIROT, C. e PASTEELS, J. M. **Ontogenetic development and the evolution of the worker caste in termites**. *Experientia*, 1987. vol. 43, p. 851 – 860.
167. NOIROT, C. Sexual castes and reproductive strategies in termites. In: W. ENGELS (ed.). **Social Insects: An Evolutionary Approach to Castes and Reproduction**. Berlin: Springer-Verlag, 1990. p. 5 – 35.
168. KORB, J. e KATRANTZIS, S. **Influence of environmental conditions on the expression of the sexual dispersal phenotype in a lower termite**: implications for the evolution of workers in termites. Regensburg, Alemanha: University of Regensburg, 2004. *Evolution & Development*. vol. 6, n.5, p. 342 – 352.
169. ABE, T. Evolution of life types in termites. In: KAWANO, S.; CONNELL, J. e HIDAKA, T. **Evolution and coadaptation in biotic communities**. Tokyo: University of Tokio, 1987. p. 125 – 148.
170. HENDEE, E.C. **The association of the termites, Kaloterme minor, Reticuliterme hesperus, and Zootermopsis angusticollis with fungi**. s.l.: Univ Calif Publ Zool, 1933. n.39, p.111–134.
171. HENDEE, E.C. The association of termites and fungi. In: Kofoid, C.A. (ed) **Termites and termite control**. Berkeley: Berkeley University, 1934. p 105–116.
172. BLACKWELL, M. e ROSSI, W. **Biogeography of fungal ectoparasites of termites**. s.l.: Mycotaxon, 1986. n. 25, p. 581–601.
173. CRUSE, A. **Termite defences against microbial pathogens**. Australia: s.e., 1998.
174. ROSENGAUS, R.B. *et al.* **Nesting ecology and cuticular microbial loads in dampwood (Zootermopsis angusticollis) and drywood termites (Incisiterme minor, I. schwarzi, Cryptoterme cavifrons)**. s.l.: J Insect Sci, 2003. n. 3, p.31.
175. ROOSE-AMSALEG, C.; BRYGOO, Y.; HARRY, M. **Ascomycete diversity in soil-feeding termite nests and soils from a tropical rainforest**. s.l.: Environ Microbiol, 2004. n. 6, p.462–469.
176. HUGHES, D.P.; PIERCE, N.E. e BOOMSMA, J.J. **Social insect symbionts: evolution in homeostatic fortresses**. s.l.: Trends Ecol Evol, 2008. n. 22, p.672–677.
177. JACKSON, D.E. e HART, A.G. **Does sanitation facilitate sociality?** s.l.: Anim Behav, 2009. n. 77, p. 1-5.
178. SANDS, W.L. The association of termites and fungi. In: Krishna, K. e Weesner, F.M. (eds) **Biology of termites**. New York: Academic Press, 1969. v. 1, p 495–524.
179. BIGNELL, D.E. e EGGLETON, P. Termites in ecosystems. In: ABE, T.; BIGNELL, D.E. e HIGASHI, M. (eds) **Termites: evolution, sociality, symbiosis, ecology**. Dordrecht: Kluwer Academic Publishers, 2000. p 363–387.

180. HOGAN, M. E. *et al.* **The site of cellulose breakdown in higher termites (Nasutitermes walker and Nasutitermes exitiosus)**. s.l.: Insect Physiol., 1988. n.34, p. 891-899.
181. OHTOKO, K. *et al.* **Diverse genes of cellulose homologues of glycosyl hydrolase family 45 from the symbiotic protists in the hindgut of the termite Reticulitermes speratus**. s.l.: Entremophiles, 2000. n.4, p. 343-349.
182. COSTA-LEONARDO, A.M. **Cupins-praga: Morfologia, biologia e controle**. Divisa:Rio Claro, 2002. 128p.
183. WARNECKE, F. *et al.* **Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite**. Nature Publishing Group, 2007. vol. 450, p. 560 – 570.
184. SHARF, M. E. e TARTAR, A. **Termite digestomes as sources for novel lignocellulases**. Florida: Biofuels, Bioprod. Bioref, 2008. vol 2, p. 540 – 552. Disponível em: < (www.interscience.wiley.com >. Acesso em: 23 out 2009.
185. ANKLIN-MÜHLEMANN, R. *et al.* **Morphological, microbiological and biochemical studies of the gut flora in fungus-growing termite Macrotermes subhyalinus**. s.l.: J. Insect Physiol., 1995. n. 41, p. 929-940.
186. BIGNELL, D. E. *et al.* **Association of actinomycete-like bacteria with soil-feeding termites (Termitidae, termitinae)**. s.l.: Appl. Environ. Microbiol., 1979. n. 37, p. 339-342.
187. BIGNELL, D. E. *et al.* **Specialization of the hindgut wall for the attachment of symbiotic micro-organisms in a termite Procupitermes aburiensis (Isoptera, Termitidae, Termitinae)**. s.l.: Zoomorphology, 1980. n.96, p. 103-112.
188. CZOLIJ, R.; SLAYTOR, M. e O'BRIEN, R. W. **Bacterial flora of the mixed segment and the hind gut of the higher termite Nasutitermes exitiosus Hill (Termitidae, Nasutitermitinae)**. s.l.: Appl. Environ. Microbiol., 1985. n.49, p. 1226-1236.
189. NAKASHIMA, K. *et al.* **Dual cellulose-digesting system of the wood-feeding termites, Coptotermes formosanus Shikari**. s.l.: Insect Biochem. Mol. Biol., 2002. n.32, p. 777-784.
190. INWARD, D.; BECCALONI, G. e EGGLETON, P. **Death of an order: a comprehensive molecular phylogenetic study confirms that termites are eusocial cockroaches**. s.l.: Biol. Lett., 2007. n.3, p. 331-335.
191. OHKUMA, M. **Termite symbiotic systems: efficient bio-recycling of lignocellulose**. s.l.: Appl. Microbiol. Biotechnol., 2003. n. 61, p. 1-9.
192. BONNER, J. **Termites' enzyme anomaly**. 2007. Disponível em: < <http://www.rsc.org/chemistryworld/News/2007/March/26030702.asp> >. Acesso em: 20 maio 2011.
193. WOOD, T.G. e JOHNSON, R.A. The biology, physiology and ecology of termites. In: _____. **Economic Impact and Control of Social Insects**. New York: Praegar, 1986. p.1-68.
194. INOUE, T., *et al.* Symbiotic associations with protists. In: ABE, T., BIGNELL, D.E. e HIGASHI, M. (eds). **Termites: evolution, sociality, symbiosis, ecology**. Dordrecht: Kluwer Academic Publishers, 2000. P. 275–288.
195. FRANCO CAIRO J.P.*et al.* **Functional characterization and target discovery of glycoside hydrolases from the digestome of the lower termite Coptotermes gestroi**. Brasil: Biotechnol Biofuels. 2011; vol. 4, n. 50.

196. COSTANTINO, R. **Chave ilustrada para identificação dos gêneros de cupins (insecta: isoptera) que ocorrem no brasil**. São Paulo: Papéis Avulsos de Zool.,1999. v. 40, n. 25, p. 387-448.
197. SILVESTRI, F. **Contribuzione alla conoscenza dei Termiti e Termitofili dell'America Meridionale**. s.l.: Redia, 1903. n.1, p. 1-234.
198. GRASSÉ, P.P. **Termitologia**. Paris: Masson,1986. 3v.
199. KRISHNA, K. e WEESNER, F.M. **Biology of Termites**. New York: Academic Press, 1969. 2v.
200. SNODGRASS, R.E. **Principles of Insect Morphology**. New York: MacGraw-Hill, 1935.
201. THORNE, B.L. **Termite terminology**. s.l.: Sociobiology, 1998. v. 28, n. 3, p.253-263.
202. CRISTINO, L.G. e ALBIN, L. **O reino dos cupins**. Brasil: UNESPCiência, 2013. Disponível em: < http://www.unesp.br/aci_ses/revista_unesp-ciencia/acervo/42/como-se-faz >. Acesso em: 30 ago 2013.
203. WINSTON, F.; DOLLARD, C. RICUPERO-HOVASSE, S.L. **Construction of a set of convenient Saccharomyces cerevisiae strains that are isogenic to S288C**. Boston: Yeast, 1995. v. 11, p. 53-55.
204. SAMBROOK, J. *et al.* **Molecular Cloning: A Laboratory Manual**. 3 ed. Australia: CSH press, 2000.
205. MACROGEN. Disponível em: < <http://www.macrogen.com/>>. Acesso em: 6 ago 2013.
206. **GeneBank**. Disponível em: < <http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucleotide> >. Acesso em: 8 set 2010.
207. FITCH, W.M. e MARGOLIASH E. **Construction of phylogenetic trees**. s.l.: Science, 1967. n. 155, p. 279-284.
208. **MEGA 5.05**. Disponível em: < <http://www.megasoftware.net/> >. Acesso em: 24 jun 2011.
209. KELLEY, L.A. e STERNBERG, M.J.E. **Protein structure prediction on the web: a case study using the Phyre server**. s.l.: Nature Protocols, 2009. n. 4, p. 363-371. Disponível em: <http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index> >. Acesso em: 23 ago 2013.
210. GENEROSO, W.C. et al. **Recombinant expression and characterization of an endoglucanase III (cel12a) from Trichoderma harzianum (Hypocreaceae) in the yeast Pichia pastoris**. Brasil: Genetics and Molecular Research, 2012. vol. 11, n. 2, p. 1544-1557.
211. CREIGHTON, T. E. **Protein folding**.s.l.: Biochemical Journal, 1990. n.270, p. 1-16.
212. BAXEVANIS, A. D. e OUELLETTE, B. F. F. **Bioinformatics: a practical guide to the analysis of genes and proteins**. 3 ed. EUA, New Jersey: John Wiley and Sons, Inc., New Jersey, 2005.

213. JONES, D. T. ; TAYLOR, W. R. e THORNTON, J. M. **A new approach to protein fold recognition**. s.l.: Nature, 1992. vol. 358, n. 6381, p. :86–89.
214. MARTIM-REMON, M. A. et al. **Comparative protein structure modelling of genes and genomes**. a.l.: Annual Review of Biophysics and Biomolecular Structure, 2000. vol.29, n. 16, p. 291–235.
215. TRAMONTANO, A. e LESK A. M. **Protein structure prediction**. Germany: John Wiley and Sons, 2006.
216. BUJNICKI, J. M. **Protein structure prediction by recombination of fragments**. s.l.: ChemBiochem: a European Journal of Chemical Biology, 2006. vol.7, p. 19-27.
217. DORN, M. **Uma proposta para a predição computacional da estrutura 3D aproximada de polipeptídeos com redução do espaço conformacional utilizando análise de intervalos**. Porto Alegre: Pontifícia Universidade Católica do Rio Grande do Sul -Faculdade de Informática, 2008. [Tese de Mestrado].
218. KOLINSKI, A. **Protein modelling and structure prediction with a reduced representation**. s.l.: Acta Biochimica Polonica, 2004. n. 51, p.349–371.
219. NGO, J. T.; MARKS, J. e KARPLUS, M. Computational complexity, protein structure prediction and the Levinthal Paradox. In: _____. MERZ JR, K. e LE GRAND, S., editors. **The Protein Folding Problem and Tertiary Structure Prediction**. Boston: BIRKHÄUSER, 1994. p. 435–508.
220. **d1Ks8a**. Disponível em: < http://bioinformatics.org/firstglance/fgij/fg.htm?mol=http://www.sbg.bio.ic.ac.uk/phyre2/phyre2_output/46694633b85d761a/d1ks8a_1.pdb >. Acesso em: 26 ago 2013.
221. **1Ks8**: The structure of Endoglucanase from termite, *Nasutitermes takasagoensis*, at pH 2.5. PDB (Protein Data Bank). Disponível em: < <http://www.rcsb.org/pdb/explore/jmol.do?structureId=1KS8&opt=3> >. Acesso em: 29 ago 2013.
222. MARANA, S.R. *et al.* **Investigation of the substrate specificity of β -glucosidase from *Spodoptera fungipeda* using site-directed mutagens and bioenergetics analysis**. s.l.: European Journal of Bioch., 2004. n.271, p. 4169-4177.
223. SASAGURI, S. **Codon optimization prevents premature polyadenylation of heterologously-expressed cellulases from termite-gut symbionts in *Aspergillus oryzae***. s.l.: J. of General and Applied Microb., 2008. n. 54, p. 343-351.
224. NI, J. *et al.* **Random exchanges of non-conserved amino acids residues among from parental termite cellulases by family shuffling improved thermostability**. s.l.: Protein Engineering, Design and Selection, 2007a. n. 20, p. 532-542.
225. NI, J. *et al.* **Heterologous expression and enzymatic characterization of β -glucosidase from the drywood-eating termite, *Neotermes koshunensis***. s.l.: Applied Entomology and Zoology, 2007b. n.42, p. 447-463.
226. Ni, J; TAKEHARA, M e WATANABE, H. **Heterologous overexpression of a mutant termite cellulase gene in *Escherichia coli* by DNA shuffling of four orthologous parental cDNAs**. s.l.:Bioscience, Biotech and Biochem., 2005. n.69, p. 1711-1720.

227. SAMI, A.J. e SHAKARI, A.R. **Biochemical characterization of endo-1,4-D-glucanase activity of a green insect pest *Aulacoptera foveicollis* (Lucas)**. s.l.: Sciences J., 2008, n.5, p. 30-36.
228. ZHOU, X *et al.* **Inhibition of termite cellulases by carbohydrate-based cellulase inhibitors: Evidence from *in vitro* biochemistry and *in vivo* feeding studies**. USA: Pesticide Biochemistry & Physiology, 2008. N.90, P. 31-41.
229. WANG, Z. et al. **RNA-seq: A revolutionary tool for transcriptomics**. S.L.: Nature, 2009. vol. 10, p.57-63. Disponível em : <<http://www.nature.com/nrg/journal/v10/n1/abs/nrg2484.html>>. Acesso em: 5 jun. 2009.
230. CARVALHO, M.C.C.G. e da SILVA, D.C.G. **Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas**. Santa Maria: Ciência Rural, 2010. vol. 40, n. 3, p. 735-744.
231. ILLUMINA, Inc. **TruSeq™ RNA and DNA Sample Preparation Kits v2**. USA: Illumina, 2011. Disponível Em: < http://www.illumina.com/documents/products/datasheets/datasheet_truseq_sample_prep_kits.pdf >. Acesso em: 12 out 2012.
232. SHENDURE, J.e JI, H. **Next-generation DNA sequencing**. s.l: Nature biotechnology, 2008. v.26, n.10, p.1135-1145. Disponível em: <<http://www.nature.com/nbt/journal/v26/n10/abs/nbt1486.html>>. Acesso em: 5 jun. 2009.
233. TURCATTI, G. et al. **A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis**. s.l.: Nucleic acids research, 2008. v.36, e25. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2275100/?tool=pubmed>>. Acesso em: 5 jun. 2009.
234. FEDURCO, M. et al. **BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies**. s.l.: Nucleic acids research, 2006. v.34, n.3, p.e22. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1363783/?tool=pubmed>>. Acesso em: 5 jun. 2009.
235. **Trinity**. Disponível em: < <http://trinityrnaseq.sourceforge.net> >. Acesso em: 29 mar 2013.
236. GRABHERR, M.G. **Full-length transcriptome assembly from RNA-Seq data without a reference genome**. s.l.: Nature Biotechnology, 2011.
237. **CAZY**. Disponível em: < <http://afmb.cnrs-mrs.fr/~cazy/> >. Acesso em: 3 fev 2013.

APÊNDICE

TRANSCRIPTOMA

Definição: O transcriptoma consiste no conjunto (completo ou parcial) de transcritos de uma célula e suas quantidades, em uma condição fisiológica ou estado específico de desenvolvimento. Dessa forma, incluem os RNAs codificantes (mRNAs) e os RNA não-codificantes (tRNA, rRNA, RNA regulatório, RNA estrutural, dentre outros)²²⁹.

As novas tecnologias de sequenciamento, denominadas de tecnologias de sequenciamento de nova geração (*Next Generation*) veem sendo uma poderosa alternativa para estudos ligados a genômica estrutural e funcional; com capacidade de permitir um sequenciamento representativo, com dados altamente informativos e reprodutíveis tanto do genoma como do transcriptoma (onde há uma grande precisão na quantificação dos transcritos) em um único passo, reduzindo custos e tempo²³⁰.

Portanto, este é de grande interesse para compreender melhor todos os mecanismos envolvidos nos processos de transcrição (alteração nos níveis de expressão) ocorrentes nas células. Para tanto, o sequenciamento em alta escala, fazendo uso da tecnologia de RNASeq, é imprescindível; já que, através desta pode-se obter um transcriptoma de alta resolução; sendo mais eficaz que as tecnologias utilizadas em *Microarrays* ou sequenciamentos de cDNAs convencionais, que demandam altos custos, tempo e que normalmente são complexos; dificultando as análises comparativas²³⁰.

A busca por potenciais genes biotecnológicos, a verificação dos níveis de expressão diferencial entre castas e entre espécies, busca por mutações em regiões

conservadas, são de grande interesse para aumentar o conhecimento científico e ampliar o estudo da biologia de qualquer organismo. Assim sendo, a utilização de espécies semelhantes, pertencentes a um mesmo táxon, mas com hábitos e ecologia diferenciados é uma proposta válida para estudos afins.

Material e Métodos

O serviço de sequenciamento dos RNAs foi realizado junto a ESALQ – USP, em Piracicaba; através do Departamento de Zootecnia, sob responsabilidade do Prof. Dr. Luiz Lehmann Coutinho (Professor Titular e Chefe do Departamento). Cerca de 4 µg de cada amostra de cupins (Velocitermes sp, Heterotermes tenuis e Cornitermes cumulans de castas soldado e operário) foram utilizados para sequenciamento dos RNAs totais e construção das bibliotecas RNASeqs. No presente trabalho, somente análises de Velocitermes sp foram realizadas, com enfoque em celulases.

No caso, a ênfase se perfaz na β-1,4-endoglucanase por ser a principal celulase na degradação de materiais lignocelulose, cuja função se relaciona a quebra inicial da celulose em fragmentos de celobiose e polissacarídeos menores. Trata-se, pois, da celulase mais bem caracterizada e já bem descrita em literatura, e alvo deste estudo.

Para tanto, segue abaixo uma explanação simplificada de como foi realizado a construção da biblioteca de RNA Seqs, bem como a do transcriptoma em si e a busca por esta celulase.

1. Construção de Bibliotecas de RNASeq

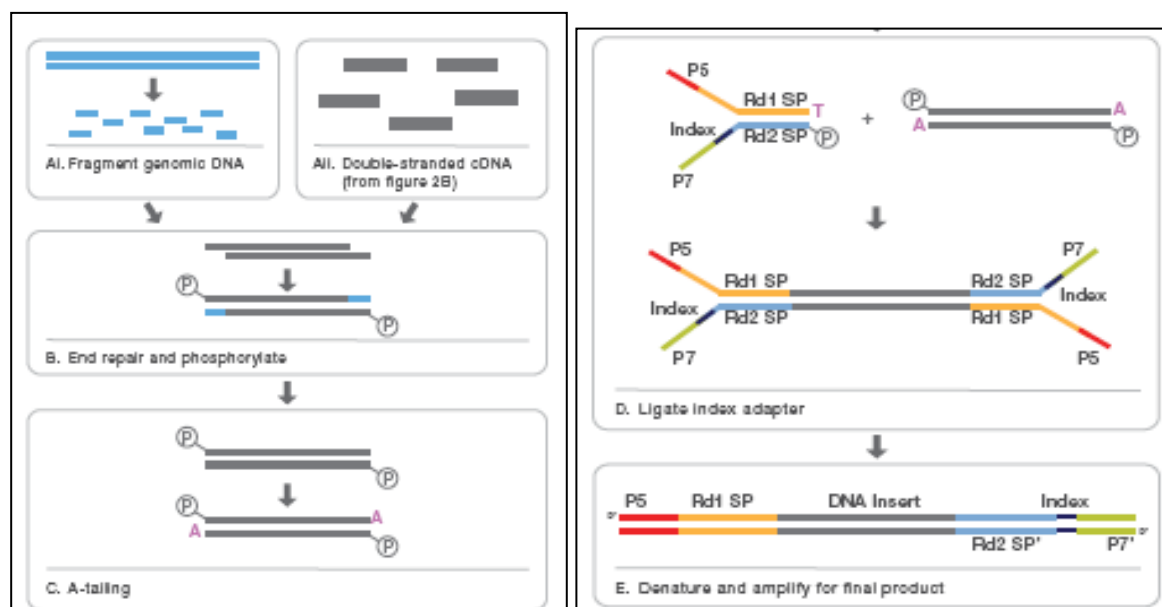
O RNA total é preparado através do uso do kit “*TruSeq RNA Sample Prep kit*” (Illumina®), cujo qual permite uma preparação rápida do RNA, reduzindo etapas de purificação, pipetagens e transferência de amostra; além de facilitar a conversão de RNA à cDNA²³¹.

Para isso, começando com o RNA total, o RNA mensageiro é, inicialmente, purificado através da seleção poli-A (que reduz passos de transferência, precipitação, eluição e fragmentação em uma única etapa) e, em seguida, é quimicamente fracionado em fragmentos e convertido em fitas-simples de cDNA através do uso de *random hexamer priming*. Esta estratégia permite uma cobertura mais uniforme em toda a transcrição, proporcionando ajustes definidos pelo usuário para comprimentos de inserção mais curtos ou mais longos (Figura 50A)²³¹.

Posteriormente, faz-se necessário a obtenção de fitas-duplas de cDNA que é, então, utilizada para a construção das bibliotecas. Nessa etapa, cDNA fita-dupla com final *blunt* (cego) são gerados usando uma combinação de reação de preenchimento e atividade exonuclease; passando por reparação e fosforilação (Figura 50B). Ao final deste cDNA fita-dupla *blunt* uma base A é acrescida em cada fita, preparando-os para a ligação dos adaptadores de sequenciamento (Figura 50C). Tais adaptadores são específicos (P5, P7) possuidores de uma base T saliente no 3'-final que se liga a base A adicionada ao fragmento de dupla fita de cDNA (Figura 50D e 50E). Essa metodologia permite a fixação dos fragmentos ao suporte de sequenciamento por hibridização a um dos adaptadores fixados²³². Essa etapa é denominada Preparação da Biblioteca.

Dessa forma, ao passo seguinte, são realizados ciclos de desnaturação e amplificação utilizando-se uma plataforma sólida de vidro (PCR de fase sólida); onde os adaptadores são fixados pela sua extremidade 5', deixando a extremidade 3' livre

onde a reação se iniciará para a formação dos *clusters*, que nada mais são do que um conjunto de muitas cópias de cada fragmento amplificados^{230,233,234}. Estes *clusters* são possíveis devido a alta densidade de adaptadores no suporte, cuja qual facilita a hibridização dos adaptadores livres dos fragmentos imobilizados a sua sequência complementar fixa próximo ao clone inicial durante este processo de anelamento. Após este processo, o fragmento forma uma estrutura em “ponte”, onde ocorre a extensão, formando a fita complementar, também em “ponte” e, posteriormente sua separação e linearização. Esses ciclos repetidos 35 vezes permite a formação de cerca de 1000 cópias de cada fragmento. Essa etapa é denominada Geração de *Clusters* (Figura 51).



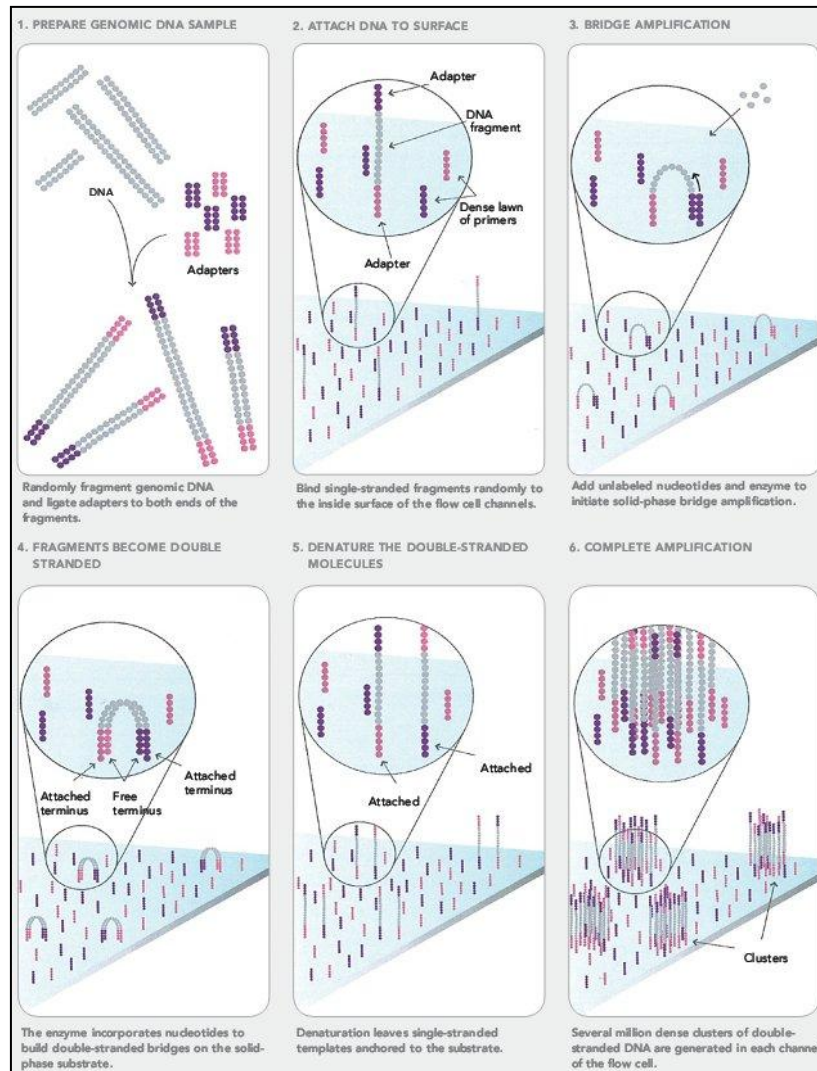
Fonte: http://www.illumina.com/documents/products/datasheets/datasheet_truseq_sample_prep_kits.pdf²³¹

Figura 50: Ligação de adaptadores na Construção de sequência sem PCR.

A construção da biblioteca começa com o DNA genômico ou com o cDNA dupla-fita produzido a partir do RNA total (A). Fragmentos *blunt-ends* são criados (B) e uma base A é adicionado ao final (C) para a ligação dos adaptadores (D). O produto final é gerado (E) e está pronto para a amplificação na *Cluster Station*.

A formação dos *clusters* de sequenciamento possibilita que o sinal de fluorescência gerado com a incorporação dos nucleotídeos aos fragmentos que estão sendo sequenciados tenha uma intensidade suficiente para garantir sua detecção pelo

sistema de *lasers* do equipamento. Cerca de 50 milhões de *clusters* podem ser produzidos por linha.

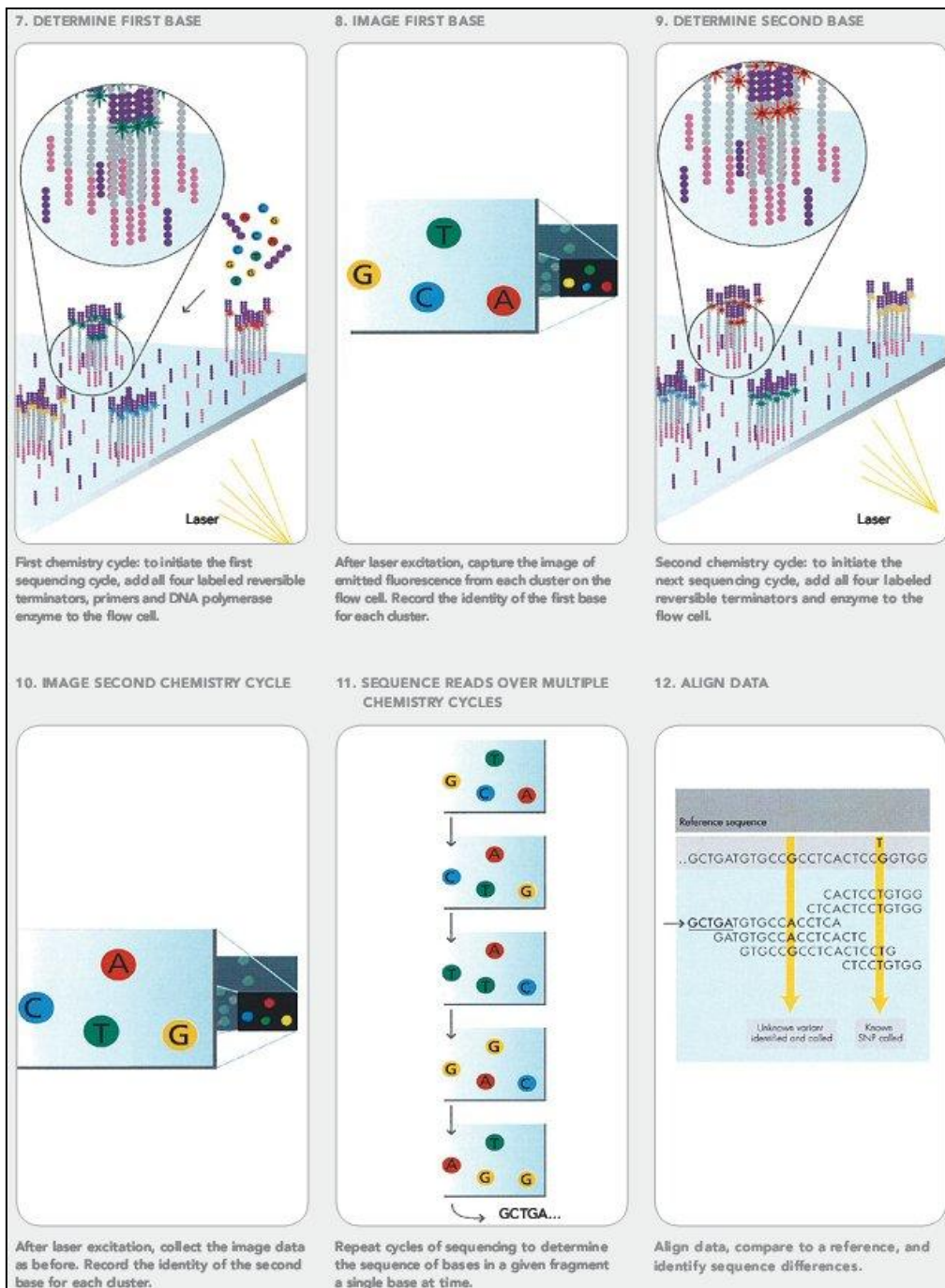


Fonte: http://www.illumina.com/documents/products/datasheets/datasheet_truseq_sample_prep_kits.pdf²³¹

Figura 51: Geração de Clusters.

Basicamente o processo de formação de clusters consiste na hibridização da região 5' terminal do fragmento à plataforma sólida onde estão presente adaptadores complementares. Isso permite que a região 3' terminal fique livre para anelamento e extensão da fita, graças a imobilização da fita na plataforma pela formação da estrutura em "ponte". Após este processo, ocorre uma desnaturação onde as fitas são separadas e linearizadas. Essa etapa de anelamento, extensão e desnaturação ocorre por 35 ciclos, permitindo a formação dos *clusters*, que contem cerca de 1000 cópias de cada fragmento.

Uma vez comparadas, a *flowcell* é colocada no equipamento HiScanSQ, no qual, através do *TruSeq SBS kit V3, 200 cycles (HS)*, os nucleotídeos marcados são incorporados em cada uma das fitas dos *clusters*. A incorporação desses nucleotídeos se dá de modo que, em seguida, ocorre uma etapa de lavagem para a remoção de reagentes excedentes e do 3' terminal bloqueado, bem como o fluoróforo do nucleotídeo incorporado no ciclo anterior. Após a incorporação, o equipamento faz a leitura da superfície da *flowcell* através de imagens e reconhece o nucleotídeo incorporado pelo comprimento de onda que o mesmo emite quando excitado pelos *lasers* do sequenciador (Figura 52). Essa etapa é denominada Sequenciamento^{230,232}. No presente caso, esse sequenciamento foi do tipo *Paired End*, de modo que os *READS* foram de 100 pb.



Fonte: http://www.illumina.com/documents/products/datasheets/datasheet_truseq_sample_prep_kits.pdf²³¹

Figura 52: Sequenciamento.

O processo de sequenciamento consiste, basicamente na leitura dos nucleotídeos marcadas incorporados. Esses nucleotídeos marcados são inseridos a partir da região 3' terminal dos fragmentos dos *clusters*, de modo que a cada inserção uma etapa de lavagem é realizada para a retirada dos fluoróforos, reagentes excedentes e 3' terminal bloqueado. A leitura é feita em todos os *clusters* ao mesmo tempo através da imagem e reconhecimento por comprimento de onda gerado.

2. Análise do Sequenciamento e construção do transcriptoma

A Bioinformática e análise de dados foi realizada em colaboração com o Laboratório de Genômica e Expressão (LGE) da UNICAMP, sob coordenação do Prof. Dr. Gonçalo Amarante Guimarães Pereira, através da utilização do programa Trinity²³⁵ (<http://trinityrnaseq.sourceforge.net>). Este programa consiste em três etapas principais²³⁶:

- a. Tratamento de dados e montagem - INCHWORM: Nessa etapa, os READs foram analisados pela quantidade e qualidade. READs com erros, de baixa complexidade ou que aparecem apenas uma vez são excluídos.
- b. Montagem - CHRYSALIS: Corre a construção dos *contigs* (alinhamento dos READs). Nessa etapa são determinados os tamanhos e verificados a possibilidade de *splicing* ocorrentes nos *contigs*. Os *contigs* são agrupados se houver sobreposição perfeita.
- c. Alinhamento – BUTTERFLY: *Contigs* semelhantes e agrupados são analisados e alinhados para a verificação de *splicing* alternativos, de modo a fazer reconstruções plausíveis dos transcritos distintos para isoformas; resolvendo ambiguidades e possíveis erros.

Posteriormente a essas análises, são realizadas buscas para a identificação de proteínas através de programas como CAZY, rpsblast e blastx. A busca por possíveis celulasas foi efetuada usando o CAZY²³⁷ (<http://afmb.cnrs-mrs.fr/~cazy/>). A partir destes, os *contigs* foram selecionados mediante as famílias GHFs e alinhadas contra banco de dados utilizando o BlastX. O *contig* possuidor da ORF da proteína foi selecionado com base na comparação com outras proteínas do banco de dados.

A partir desta sequência (*contig*) os *primers* específicos foram desenhados, como citado na seção 4.9. (Figura 53)

```

TCCACGTGACTTTCTGATCA GAGTACATCGACTTGAGTGTCTCTCCAGTTGTGCAGCTGATCA GGTAA
CAGCAGCTTCTGTGTAATTACTTATAAGTCGGCAGGACCGTCCTACCAAAA CACTT CACGGACAGCAA
CCAATACCAGCCAGCATGAAGGTCTTCTTGCCTTCTTTCTGCGTTCGCGCTT GCCAA GCTGCTTAG
ACTACAAGCAAGTACTCAGCAATCCCTACTGTCTACGAGGCTCAGCGATCGGGAAATTTGCCCTCTG
ATCAGAAAGTCACGTGGAGGAAGAAATCCGCCTTAAACGACA AAGGAGCCAGGGCCAGGACCTAAC
AGGAGGATACTATGACGCTGGTGATTTTGTGAAATT CGGCTT TCCAATGGCGTACACAGCCACCGTCCT
GGCTTGGGGGCTGATCGACTTTCAAGCGGCCCTACACTGCA GCA GGAGCTCTGGATGAAGGTGCAAA
GCTATTA AATGGCCACAGACTACTTCTCAAGGCGCACACGGCCACAAATGAACTGTACGGACAAGT
GGGCCAGGGAGAAACGGACCA CGCTACTGGGGACGTCCAGAA GACATGACGGAGGCAAGACCTGC
CTACAAGATCGACACGTCCGCCACCGGCTCGGATCTGGCCGGCAGACAGCCGCCGCCCTCGCTGCAG
CTTCCATCGTGTCCAGAGTGTGACGCCACTTATTCCAACAACCTGCTCACACACGCCAAGCAGCTTTT
CAACTTCGCCAACAGTTATCAAGGCAAAATACAGTACTCAATCGCCGACGCCAAGAAATTTCTACGGTC
CTCTGACTACAGGGATGAGCTAGTATGGGCAGCTGCATGGATGTACAATGCAACCAAGGATAACACAT
ACCTCAACACCGCTGAATCGCTATATAACCAATTCGGTCTCCAGAAATGAATGGTGGCTTAACTGGG
ATAACA AATGTCCGGTGTACAGTTCTACTGGCCAGACTCACA AACAAGCAAGGATACAAGGACGG
GTA CAAGGATACGTGGATTACCTAATTACAATCAGCAGAAAGACACCCAA GGGTCTTCTTACATCGAC
ACGTGGGGTACCCTGCGACATGCTGCAACGCGGCCCTTATTATGCTTGAAGCTGCGAACCTGGGTTT
GAATCCTGTAGTTATCGCCAGTTGCCAA GAAGCAGATCGATTACGCATTGGGTGACGGTGGACGCA
GCTACGTGTGTGGATTTGGCAATAACCACTACCCACTCACAGCATAGATCTAGTTCTGCCCCGACG
CGCCGGCCACATGTGACTGGAACACATACAACA GCCCGAATCCAACTACCACGTACTCACCGGAGCTT
TGGTGGGCGGTCCAGATCAGAA CGAACACTCTCGGACATTGCGAGCGATTATGTCCACAACGAAGTG
GCCACAGATTAACAACGCAGGCTTCCAA TCAGCTGTCCGCCCTCATCACTTTGGGCTATTAAACTCAC
AGAACATTCACTTTGTCTTCTGCTGTGAGTGTTCACAGAA TGTAACATTGTTCTGAGCAGTGA AATTA
ATCTTTGTTTGCTGTTAGGTCTGATTCATTTGTAAAATAATA TCA TTA CTTCATACAGGGCTGAAATAA
CAATGCATAATCTTAGAACGCGA AATGGATGTTGAAATAACTACAATATTACAACA AAGTACACAGTA
AAACTAAGGGCATTATTCA GAAA CGT GTTTCTTTCAATTAATAATTCTATGA GAAAAAAAATAAGAATG
TTCCCTGT AATATTG TAA CAAA AAGGACTCTAAAATTTATAGTGA GTCTCAGTGT TTAAGGAAAGTGA
TAACAGAGA GTGAAAAA ACTGAAATGA ACTGTCCA GTAAA CATATTGGGAGT AAGA AATTATCAATTG
AAAAATTA TCACTTAAAATGAAGACTACAATAAGATTGGA CAGATCAAGTAAGCTACACA CAATGAA
GGTGGATACCGTAGA CAGTGGAAATGCAAGACA CTGGCGA AATACAGAAAA GATGTA GAGACCACTAA
CCACACACAGAAA CA CTGGCACTGATTA ACTTTTCTCCATGGGCGGAA ACCAGGAAAAATCTTAAAA
CACACTTAACTTTTAAAAATGATTTTAAATCTCTTGTACCACTATTTAACA CACTTA ACTAACACTTGG
AT

```

Figura 53: Sequência da β -1,4-endoglucanase de V. heteropterus a partir de seu transcriptoma.

A sequência utilizada a partir do transcriptoma de *Velocitermes* sp foi selecionada de modo que a ORF estivesse completa para posterior desenho dos *primers* específicos. Dessa maneira, selecionou-se a região abrangente entre o ATG e *stop códon*, de modo que a sequência estivesse em *frame*.

ANEXO I

Predição da Estrutura secundária de β -1,4-endoglucanase de *Nasutitermes takasagoensis*.

A predição da estrutura secundária para a sequência de Aminoácidos de *Nasutitermes takasagoensis* para a celulase β -1,4-endoglucanase (Figura 54) mostrou similaridade com a sequência denominada d1ks8a do programa Phyre2; caracterizada pela presença de um domínio catalítico de celulase. Essa comparação mostra uma similaridade e confiança de 100% e cobertura de 97 % (Figura 55). Basicamente essa celulase é formada por 49 % de α -hélices; 11 % de cadeias β e 3 % de regiões desordenadas.

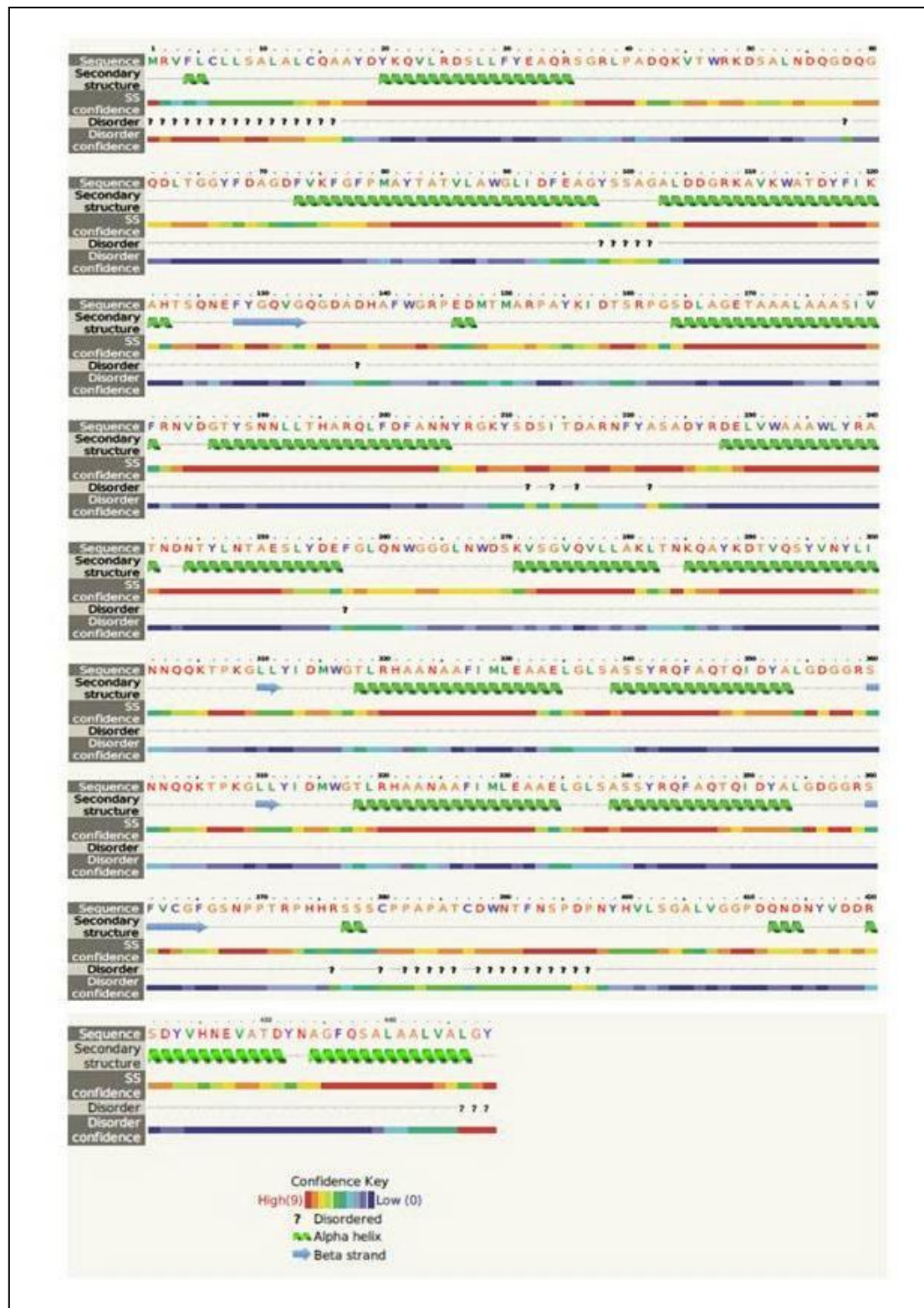


Figura 54: Estrutura Secundária de β -1,4-endoglucanase de *N. takasagoensis*.

A estrutura secundária de β -1,4-endoglucanase de *N. takasagoensis* é formada básica por 49 % de α -hélice; 11 % de cadeias β e 3 % de regiões desordenadas, representadas pelas aspirais verde, setas azuis e ponto de interrogações, respectivamente. Regiões em vermelho possuem maior confiabilidade segundo o programa.

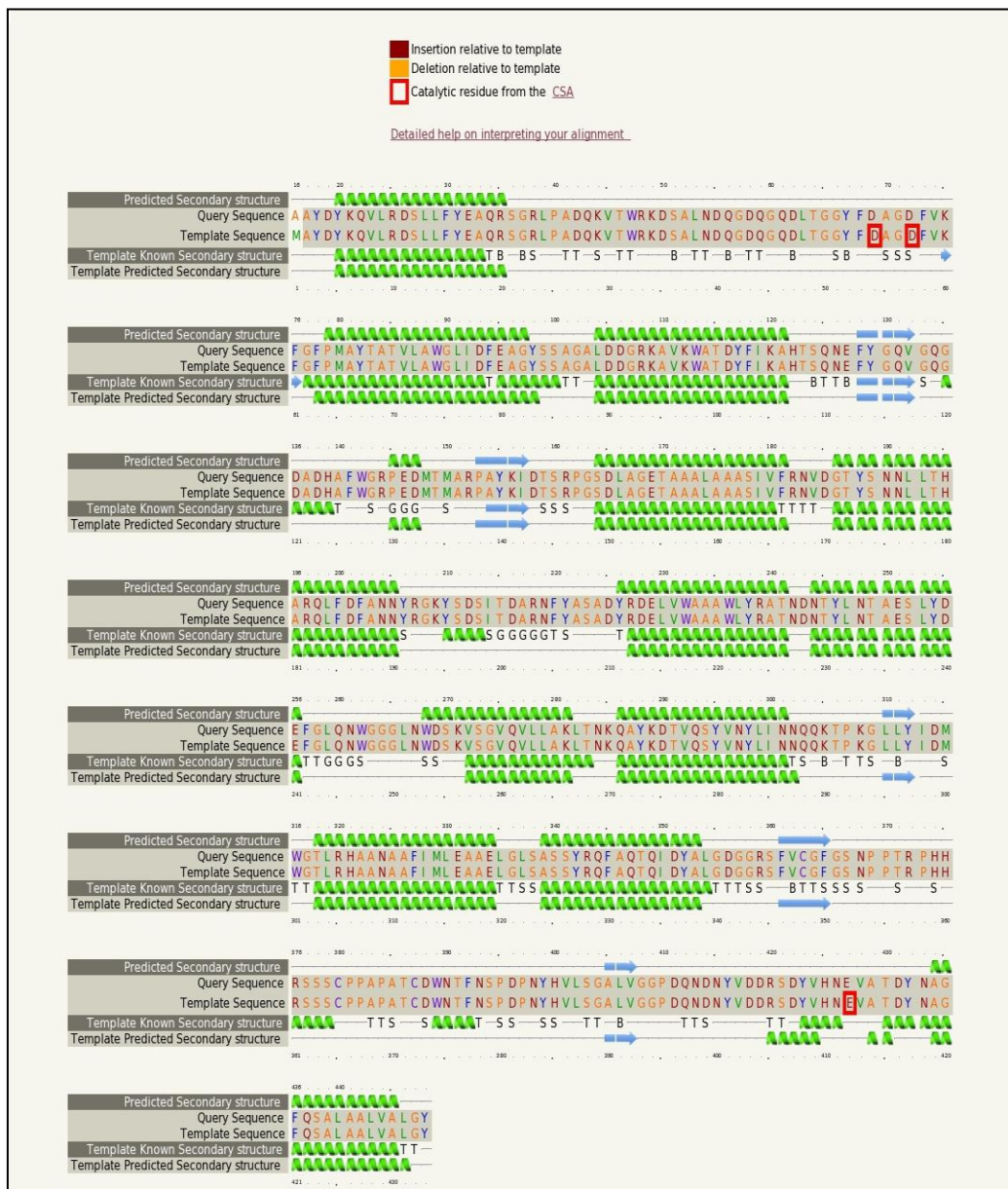


Figura 55: Alinhamento da β -1,4-endoglucanase de *Nasutitermes takasagoensis* com *template* d1Ks8a.

A maior similaridade foi resultante da comparação da β -1,4-endoglucanase de *N. takasagoensis* com o *template* d1Ks8a obtido através do programa Phyre2. Essa comparação mostra uma confiança e similaridade de 100 % e cobertura de 97% ; sendo o que mais se aproxima da configuração secundária da celulase em questão. Na imagem as aspirais verdes indicam regiões em α -hélice e as setas azuis indicam as regiões de cadeias β . Os quadrados em vermelhos nos resíduos de aminoácidos D e E indicam resíduos catalíticos. T = Ligação por Hidrogênio; S= curva ; B = resíduos em pontes β isolados ; G = 3 voltas de hélices ; ? = regiões desordenadas. *Query* indica a sequência de *N. takasagoensis* e o *template*, a estrutura d1Ks8a.

ANEXO II

Predição da Estrutura secundária de β -1,4-endoglucanase de V. heteropterus.

A predição da estrutura secundária para a sequência de Aminoácidos de V. heteropterus para a celulase β -1,4-endoglucanase (Figura 56) mostrou similaridade com a sequência denominada d1ks8a do programa Phyre2; caracterizada pela presença de um domínio catalítico de celulase. Essa comparação mostra uma similaridade e confiança de 100 % e cobertura de 84% (Figura 57). Basicamente essa celulase é formada por 50 % de α -hélices; 4 % de cadeias β e 8 % de regiões desordenadas.

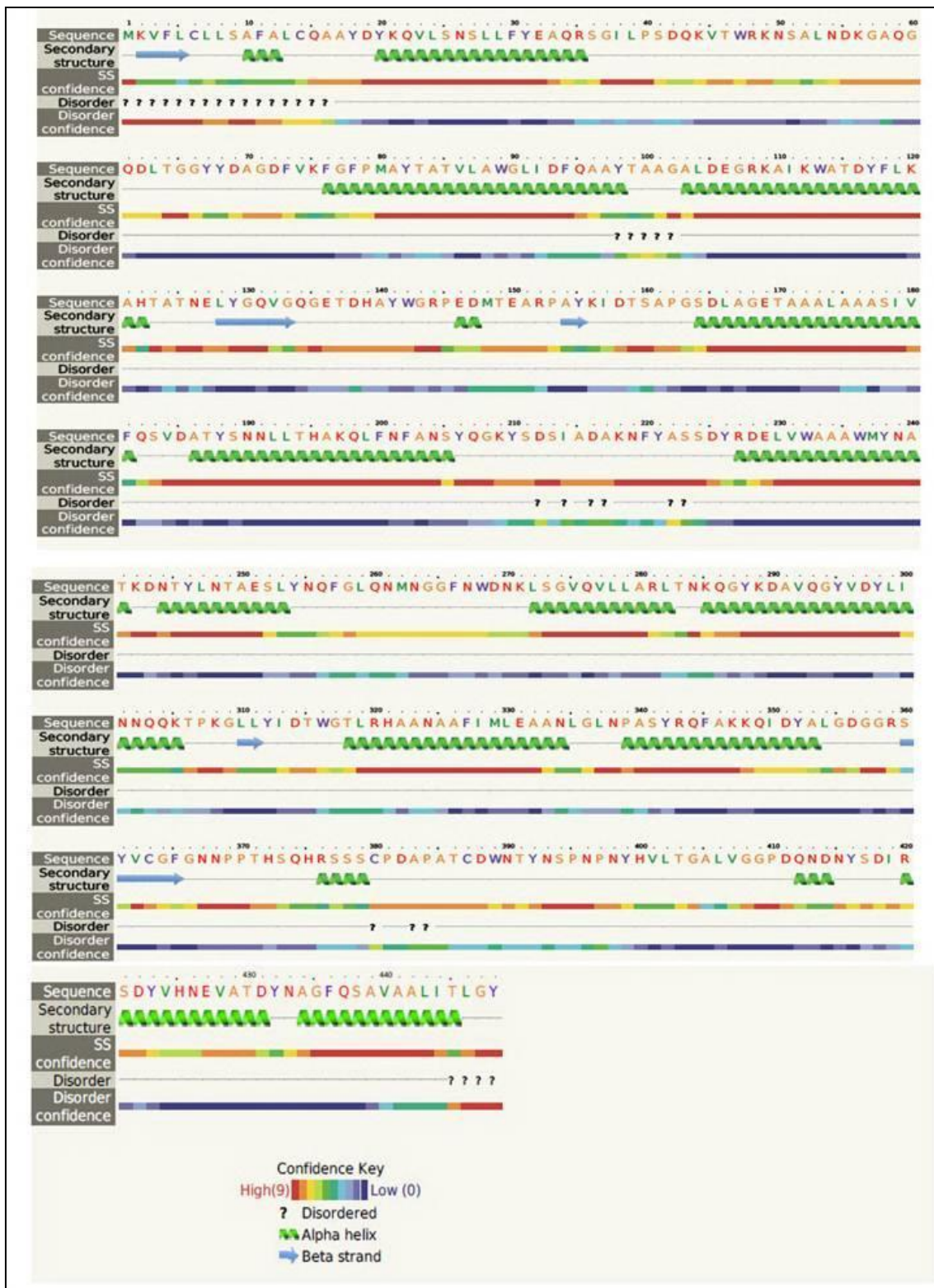


Figura 56: Estrutura Secundária de β -1,4-endoglucanase de *V. heteropterus*.

A estrutura secundária de β -1,4-endoglucanase de *V. heteropterus* é formada básica por 50 % de α -hélice; 4 % de cadeias β e 8 % de regiões desordenadas, representadas pelas aspirais verde, setas azuis e ponto de interrogações, respectivamente. Regiões em vermelho possuir maior confiabilidade segundo o programa.

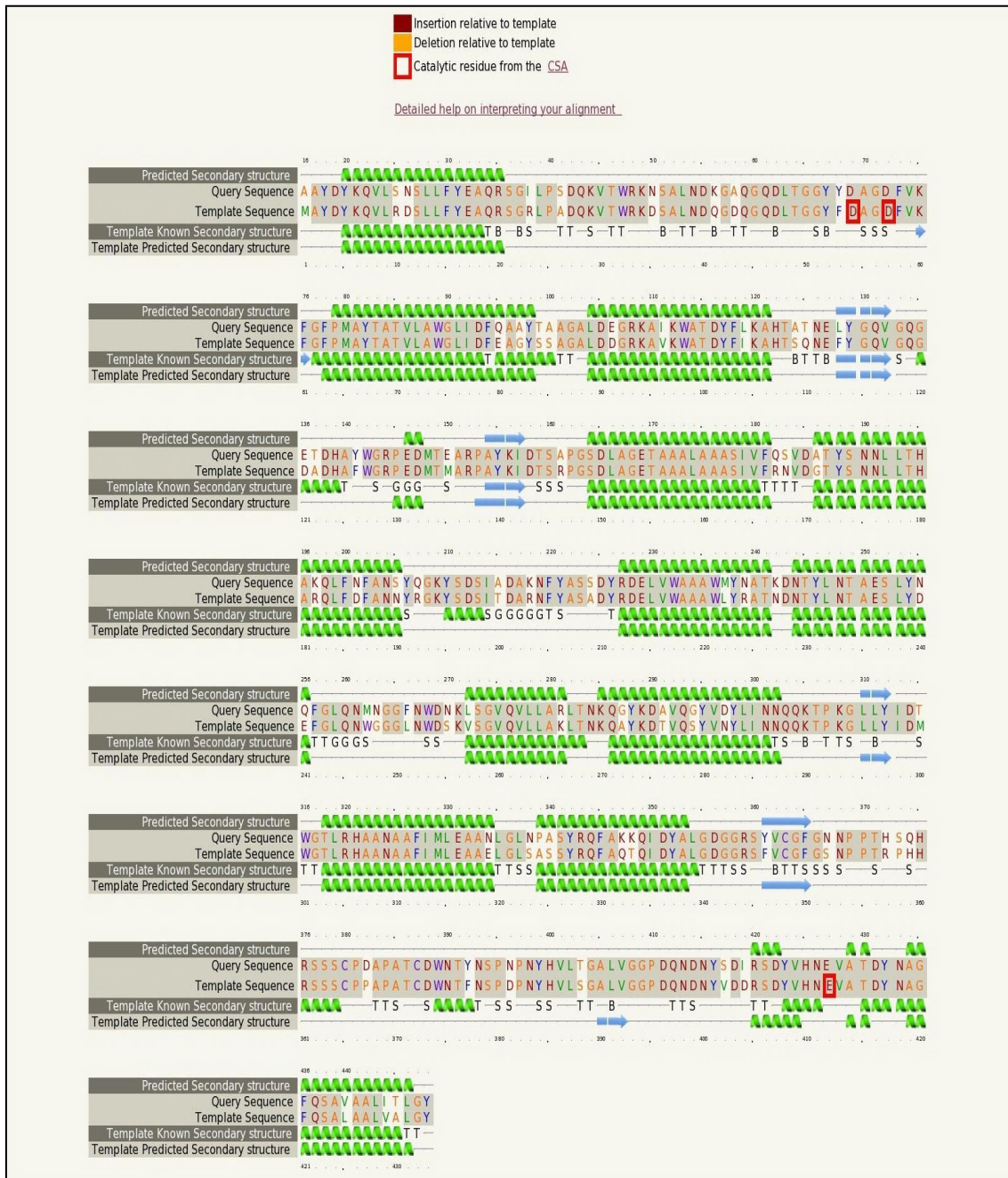


Figura 57: Alinhamento da β -1,4-endoglucanase de *V. heteropterus* com *template* d1Ks8a.

A maior similaridade foi resultante da comparação da β -1,4-endoglucanase de *V. heteropterus* com o *template* d1Ks8a obtido através do programa Phyre2. Essa comparação mostra uma confiança e similaridade de 100 % e cobertura de 84%; sendo o que mais se aproxima da configuração secundária da celulase em questão. Na imagem as aspirais verdes indicam regiões em α -hélice e as setas azuis indicam as regiões de cadeias β . Os quadrados em vermelhos nos resíduos de aminoácidos D e E indicam resíduos catalíticos. T = Ligação por Hidrogênio; S= curva ; B = resíduos em pontes β isolados ; G = 3 voltas de hélices; ? = regiões desordenadas. *Query* indica a sequência de *V. heteropterus* e o *template*, a estrutura d1Ks8a.