

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

ROCHANE REGINA NEVES BAPTISTA TORRES

ANÁLISE DO GENE CODIFICANTE DA ENZIMA BETA-GLUCOSIDASE NA
LEVEDURA *Dekkera bruxellensis*.

RECIFE – PE

2012

ROCHANE REGINA NEVES BAPTISTA TORRES

ANÁLISE DO GENE CODIFICANTE DA ENZIMA BETA-GLUCOSIDASE NA
LEVEDURA *Dekkera bruxellensis*.

Dissertação apresentada ao Programa de Pós - Graduação em Genética da Universidade Federal de Pernambuco como parte dos requisitos exigidos para a obtenção do título de Mestre em Genética.

Orientador: Marcos Antônio de Moraes Jr

Recife, PE

2012

Catálogo na Fonte:
Bibliotecária Teresa Cristina Moreira de Lucena, CRB-4/1419

T693a Torres, Rochane Regina Neves Baptista

Análise do gene codificante da enzima Beta-Glucosidase na Levedura *Dekkera bruxellensis* / Rochane Regina Neves Baptista Torres. – Recife: O Autor, 2012.

71 folhas: il., fig., tab.

Orientador: Marcos Antônio de Moraes Júnior

Dissertação (mestrado) – Universidade Federal de Pernambuco. Centro de Ciências Biológicas. Pós-graduação em Genética, 2012.

Inclui bibliografia e anexos

1. Bagaço de cana 2. Celulose – Celobiose 3. Cana-de-açúcar I Moraes Júnior, Marcos Antônio de (orientador) II. Título.

579.135

CDD (22.ed.)

UFPE/CCB-2012-164

ROCHANE REGINA NEVES BAPTISTA TORRES

ANÁLISE DO GENE CODIFICANTE DA ENZIMA BETA-GLUCOSIDASE NA
LEVEDURA *Dekkera bruxellensis*..

APROVADA EM

BANCA EXAMINADORA:

Presidente: Marcos Antônio de Moraes Jr.

1° Examinador: Valdir de Queiroz Balbino.

2° Examinados: Brígida Thaís Luckwu de Lucena.

Recife, PE

2012

À milha filha Larissa, por me mostrar um
amor incondicional.

AGRADECIMENTOS

Ao meu esposo Marcello, pela paciência, incentivo e compreensão pelos anos morando em casas separadas para a conclusão desse trabalho.

À minha família, por acreditar em mim e torcer pelo meu sucesso.

À Usina Santa Teresa, por investir nesse trabalho e acreditar no meu potencial.

Ao meu Orientador Marcos Morais, pela paciência, pela oportunidade e por acreditar no meu trabalho.

Aos meus amigos, Theresa, Brígida, Billy, Fernanda e Rute, pelos momentos de descontração, conselhos e contribuição à execução desse trabalho.

Ao Professor Antônio Carlos, por todas as vezes que o perturbei com as minhas dúvidas.

Ao Professor Valdir e ao seu aluno Marcus Sergipe, por sempre me ajudar em quando precisei.

Ao Professor Roberto, por ajudar na identificação dos resíduos do sítio ativo da beta-glucosidase.

Aos amigos do Grupo Labem, Luciana, Ribas, Rodrigo, Carol, Anna Carolina, Esteban, José Roberto e Felipe pelo convívio, por dividir a capela, centrífuga, bancada, pelas conversas, por contribuir de alguma forma para a realização desse trabalho.

Aos amigos do Grupo Papiloma Vírus, André, Eliane, Felipe, Marcelo, Janaina, pelo convívio, pelas conversas e por ajudar quando precisei.

RESUMO

O bagaço da cana-de-açúcar é constituído por celulose, hemicelulose e lignina. A celulose, quando hidrolisada, é transformada em celobiose que por sua vez pode ser hidrolisada em duas moléculas de glicose. A enzima β -glucosidase é a proteína responsável pela hidrólise de celobiose em glicose. Entretanto, o bagaço é, na maioria das vezes, queimado para a produção de energia elétrica para ser utilizada no processo industrial não sendo utilizado para a produção de etanol. Por ter um crescimento mais rápido do que a levedura *Saccharomyces cerevisiae* em condições industriais de fermentação alcoólica apresentando uma maior resistências a altas concentrações de etanol, altas temperaturas e baixo pH têm feito da *Dekkera bruxellensis* alvo de pesquisas e melhoramentos genéticos. O presente trabalho teve como objetivo identificar e analisar *in silico* o gene codificante da β -glucosidase no genoma da levedura *D. bruxellensis* através das ferramentas de bioinformática. Os resultados mostraram que a partir do alinhamento BLASTp, uma seqüência curta da levedura *D. bruxellensis* apresentou similaridade com seqüências codificantes para a enzima β -glucosidase de espécies da divisão ascomiceto. A seqüência da espécie *Kluyveromyces marxianus* obteve o maior percentual de similaridade com a seqüência de *D. bruxellensis*. Foram identificados no alinhamento múltiplo, através de conhecimento prévio da literatura, os resíduos do sitio ativo da enzima β -glucosidase presentes na seqüência de *D. bruxellensis* mostrando ser sítios bastante conservados.

Palavras-Chaves: Bagaço, Cana-de-Açúcar, Celulose, Celobiose, microrganismo e Análise *in silico*.

ABSTRACT

Bagasse of cane sugar is composed of cellulose, hemicellulose and lignin. Cellulose, when hydrolyzed, is transformed into cellobiose which can be hydrolyzed to two glucose molecules. The enzyme β -glucosidase is the protein responsible for the hydrolysis of cellobiose to glucose. However, the bagasse is, in most cases, burned to produce electricity to be used in the industrial process and it is not used for ethanol production. Due to have a faster growth than *Saccharomyces cerevisiae* yeast in industrial fermentation conditions and presenting a greater resistance to high concentrations of ethanol, low pH and high temperatures have made the *Dekkera bruxellensis* subject of research and genetic manipulations. This study aimed to identify and analyze *in silico* gene encoding β -glucosidase in the *D. bruxellensis* yeast genome through the computational tools of bioinformatics. The results showed that in the blastp alignment, a short sequence of *D. bruxellensis* yeast showed similarity with sequences coding for the β -glucosidase enzyme in the ascomycete division species. The *Kluyveromyces marxianus* sequence had the highest percentage in similarity with *D. bruxellensis* sequence. The multiple alignment tool was able to identify, using prior knowledge of the literature, the active site residues of the β -glucosidase enzyme in the sequence of *D. bruxellensis*. Sites showed to be very conserved.

Key Words: Cane sugar, bagasse, cellulose, cellobiose, microorganism and *in silico* analysis.

LISTA DE ILUSTRAÇÕES

Figura 1. Esquema da hidrólise da celulose. Na primeira reação a celulose é hidrolisada enzimaticamente em celobiose que é hidrolisado em glicose pela beta-glicosidase. O polímero da celulose, a enzima possui especificidade pelas ligações β -1,4-glicosídicos podendo ser classificadas como exo-glucanase ou celobiohidrolase ou endoglucanase dependendo do sítio ativo. 27

Figura 2. Estrutura da beta-glicosidase de *Kluyveromyces marxianus*. Esquerda. Estrutura na forma de tetrâmero. Direita. Estrutura do monômero. Em azul, domínio GH3 N-terminal. Em verde, domínio α/β – sanduíche. Em amarelo, domínio PA14. Em laranja, domínio GH3 C-terminal. Yoshida *et al*, 2010. 27

Figura 3. Estrutura detalhada do sítio ativo da enzima beta-glicosidase da bactéria *Thermotoga neapolitana*. Glicose está mostrada com uma superfície transparente no meio da figura (01). Pozzo *et al*, 2010. 28

Figura 4. Domínios N-terminal e C-terminal encontrados na seqüência do contig 145 de *Dekkera bruxellensis* através do programa on-line Swiss-model (<http://swissmodel.expasy.org>). 41

Figura 5. Representação dos domínios de Beta-glicosidase. (a) Beta-glicosidase de *Thermotoga neapolitana*, o retângulo representando o domínio 1 (resíduos 1-306), o octógono representando o domínio 2 (resíduos 321-536)

e o círculo representando o domínio 3 (resíduos 600-721). (b) Contig 145 *Dekkera bruxellensis*, retângulo representa o domínio 1 (resíduos 1-177) e o octógono representa o domínio 2 (resíduos 252-279). 42

Figura 6. Visualização do tetrâmero da enzima β -glucosidase de espécie *Th. neapolitana* através do programa Cn3. A seqüência da enzima β -glucosidase da espécie *A. marxianus* encontra-se destacada em amarelo. 42

Figura 7. Visualização do tetrâmero da enzima β -glucosidase de espécie *Th. neapolitana* através do programa Cn3. Em amarelo, a seqüência do contig 145 da espécie *D. bruxellensis*. 43

Figura 8. Visualizaçãodo monômero da enzima β -glucosidase de espécie *K. marxianus* através do programa Cn3. Em amarelo, os resíduos do sítio ativo presentes na seqüência da espécie *D. bruxellensis*. 43

Figura 9. Alinhamento múltiplo das seqüências de beta-glucosidase evidenciando os resíduos do sitio ativo visualizados no programa Bioedit versão 7.0. 44

Figura 10. Alinhamento múltiplo das seqüências de beta-glucosidase evidenciando os resíduos D58 (Esq.) e E458 (Dir.) do sitio ativo encontrados na bactéria *Thermotoga neapolitana* visualizados no programa Bioedit versão 7.0. 45

Figura 11. Árvore filogenética construída a partir das seqüências do gene da beta-glucosidase usando o agrupamento Neighbor-Joining do programa Mega 5. 46

Figura 12. Árvore filogenética construída a partir das seqüências do gene da beta-glucosidase das espécies da divisão Ascomycetes e do grupo externo, bactéria *Thermotoga neapolitana*, usando o agrupamento Máximo Verossimilhança do programa PhymI. 47

Figura 13. Topologias mais freqüentes encontradas a partir de seqüências protéicas de genes ortólogos de *Dekkera bruxellensis*. (B) Topologia 1 apresentou uma freqüência de 34%. (C) Topologia 2 apresentou uma freqüência de 27%. (D) Topologia apresentou uma freqüência de 26%. Woolfit *et al*, 2007. 54

Figura 14. Esquema taxonômico do grupo dos ascomicetos. (A) Divisão. (B) Sub-divisão. (C) Classe. (D) Família. (E) Gênero. 56

LISTA DE TABELAS

- Tabela 1. Identificação das seqüências do gene Beta-glucosidase dos microrganismos selecionados. 39
- Tabela 2. Similaridades das seqüências codificantes da enzima beta-glucosidade dos microrganismos selecionados. 40
- Tabela 3. Crescimento das linhagens do grupo *Brettanomyces/Dekkera* em meio de cultura contendo variadas fontes de carbono. Galafassi *et al*, 2010. 49

LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

A. clavatus - *Aspergillus clavatus*

A. flavus - *Aspergillus flavus*

A. fumigatus - *Aspergillus fumigatus*

A. niger - *Aspergillus niger*

A. oryzae - *Aspergillus oryzae*

C. tropicalis - *Candida tropicalis*

D. bruxellensis - *Dekkera bruxellensis*

K. marxianus - *Kluyveromyces marxianus*

N. fischeri - *Neosartotya fischeri*

P. angusta - *Pichia angusta*

S. cerevisiae - *Saccharomyces cerevisiae*

S. angusta - *Scheffersomyces angusta*

S. etchellsii - *Schwanniomyces etchellsii*

T. stipitatus - *Talaromyces stipitatus*

Th. neapolitana - *Thermotoga neapolitana*

V. albo-atrum – *Verticillium albo-atrum*

DBDM – *Dekkera/Brettanomyces* meio diferencial

DDBJ – Banco de dados do Japão

EBI – Instituto europeu de bioinformática

KEGG – Enciclopédia de genes e genoma Kyoto

NCBI – Centro de informação de biotecnologia

PDB – Banco de dados de proteína

ÚNICA – União da Indústria de Cana-de-açúcar

SUMÁRIO

RESUMO.	VI
ABSTRACT.	VII
LISTA DE ILUSTRAÇÕES.	VIII
LISTA DE TABELAS.	XI
LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS.	XII
1. INTRODUÇÃO.	15
2. REVISÃO DA LITERATURA.	17
Produção Alcoólica no Brasil.	17
<i>Dekkera bruxelensis</i> .	20
Bagaço, Celulose e Beta-glucosidase.	23
Técnicas de Bioinformática.	28
3. OBJETIVOS.	33
3.1 Objetivo Geral.	33
3.2 Objetivos Específicos.	33
4. MATERIAL E MÉTODO.	34
4.1 Seleção e Análise de Seqüências.	34
4.2 Alinhamento Múltiplo.	34
4.3 Resíduos do Sítio Ativo.	35
4.4 Análise Filogenética	35
5. RESULTADOS	37

5.1 Seleção das Seqüências.	37
5.2 Alinhamento Múltiplo.	38
5.3 Predição dos Sítios Ativos.	41
5.4 Análise Filogenética.	45
6. DISCUSSÃO.	48
7. CONCLUSÕES.	57
8. REFERÊNCIAS.	58
9. ANEXO.	69
10. MEMORIAL.	71

1. INTRODUÇÃO

O estado de Pernambuco é o segundo maior produtor de álcool da região Norte/Nordeste do país. Entretanto, a tendência de crescimento do preço mundial do petróleo, aliada à necessidade de redução de emissão de poluentes na atmosfera e ao grande volume de vendas de automóveis bicomustíveis, tende a impulsionar as demandas internas e externas por álcool, sendo necessário, por isso, grande volume de investimento. A produção de etanol ocorre em larga escala através da fermentação da sacarose extraído da gramínea *Saccharum officinarum*, também conhecida como cana-de-açúcar. Atualmente, o bagaço da cana-de-açúcar é na maioria das vezes queimado para a produção de energia elétrica para ser utilizada no processo industrial. O bagaço é constituído por celulose, hemicelulose e pela lignina. A celulose, quando hidrolisada, é transformada em celobiose que, por sua vez, pode ser hidrolisada em duas moléculas de glicose. A enzima beta-glucosidase é a proteína responsável pela hidrólise de celobiose em glicose.

A espécie *Dekkera bruxellensis* tem sido associada a episódios de contaminação nas indústrias de vinho, cerveja e sucroalcooleira comprometendo o rendimento de produção. Diante desses acontecimentos, vários trabalhos têm sido realizados a fim de identificar, entender e minimizar os efeitos causados por essa levedura. Análises fisiológicas de isolados dessa espécie apontaram uma reduzida capacidade fermentativa em comparação à levedura *Saccharomyces cerevisiae*, entretanto a dinâmica populacional de *D. bruxellensis* tem mostrado que essa levedura possui um crescimento mais rápido do que *S. cerevisiae* em condições industriais de

fermentação alcoólica e uma vez iniciado tende a dominar todo o processo. Resistências a teor alcoólico, altas temperaturas e pH baixo são características que têm feito da *D. bruxellensis* alvo em pesquisas em melhoramentos genéticos do grupo do Núcleo de Engenharia Metabólica.

O presente trabalho teve como objetivo identificar e analisar *in silico* o gene codificante da β -glucosidase no genoma da levedura *D. bruxellensis* através das ferramentas computacionais de bioinformática.

2. REVISÃO DA LITERATURA

Produção Alcoólica no Brasil

O setor sucroalcooleiro do Brasil é o mais competitivo do mundo possuindo os maiores níveis de produtividade, rendimento e menores custos de produção (Vidal *et al*, 2006). O estado de Pernambuco é o segundo maior produtor de álcool da região Norte-Nordeste com uma produção de mais de 372.000m³ na safra 2010/2011. Esse resultado equivale a 19% da produção da região de um total de 1.990.178m³ de etanol. Entretanto, a produção da região Norte/Nordeste representou apenas 7% da produção total do Brasil no mesmo período (Sindaçúcar, 2011). Dados fornecidos pela União da Indústria de Cana-de-açúcar (Única) da safra 2008/2009 mostraram um faturamento de 45 bilhões de reais com a produção de 560 milhões de toneladas de cana-de-açúcar em 450 usinas em todo país. Nesse período foram produzidos cerca de 27,5 bilhões de litros de álcool, distribuídos em mais de 18,2 bilhões de litros de álcool hidratado e 9,3 bilhões de litros de álcool anidro (Única, 2011).

A tendência de crescimento do preço mundial do petróleo, aliada à necessidade de redução de emissão de poluentes na atmosfera e ao grande volume de vendas de automóveis bicombustíveis tende a impulsionar as demandas internas e externas por álcool, sendo necessário, por isso, grande volume de investimento (Vidal *et al*, 2006). De acordo com dados da UNICA, em 2012, 50% da frota circulante de automóveis no Brasil serão bicombustíveis, e esse percentual subirá para 65% em 2015. Ainda segundo esta fonte, em 2009 eram mais de 7 milhões dos automóveis bicombustíveis em circulação no Brasil e o etanol já havia substituído 50% do consumo da

gasolina (Dal Poggetto, 2009). Atualmente a capacidade de produção de etanol só tem condições de garantir o abastecimento de 45% da frota de veículos flex, a previsão de crescimento para a safra 2011/2012 será de apenas 0,52% e sem novos investimentos a capacidade de abastecimento poderá atingir 37% em 2020 (Alvarenga, 2011).

O interesse por essa fonte de energia se deve, dentre outros aspectos, ao fato de que o álcool combustível é um produto obtido de fonte renovável. Desta forma, contribui para redução do efeito estufa e não agride a camada de ozônio, sendo o seu diferencial na preservação ambiental (Única, 2005). Mesmo sendo considerado um biocombustível por ter origem biológica não fóssil utilizando carbono da atmosfera para produzir energia, o balanço de CO₂ não é neutro pois é preciso levar em conta a energia necessária para a produção dos adubos, transporte e irrigação. A utilização do etanol aliado a uso de combustível biodiesel oferece grandes vantagens para o meio ambiente, tendo em vista que a poluição é menor que a do óleo diesel comum (Costa-Neto *et al*, 2000; Araújo *et al*, 2002).

A produção de etanol ocorre em larga escala através da fermentação da sacarose extraído da gramínea *Saccharum officinarum*, também conhecida como cana-de-açúcar. A fermentação ocorre nos meses de agosto a janeiro após o período de maturação, tempo necessário para a planta acumular açúcar (Cheavegatti-Gianotto *et al*, 2011). A quantidade acumulada de açúcar geralmente está entre 12 e 17% do peso seco da planta sendo utilizado na forma de caldo e/ou melaço (Wheals *et al*, 1999) por células de levedura *Saccharomyces cerevisiae*. As leveduras utilizadas no processo podem ser de panificação prensada, bem como linhagens isoladas do próprio processo

industrial ao longo de toda a safra (Stroppa *et al*, 2000). Após cada batelada de fermentação, bem como durante o processo de fermentação contínua, as células de levedura são separadas do meio fermentado por centrifugação e reutilizadas no processo, após serem tratadas com ácido sulfúrico para redução da contaminação bacteriana. A alta concentração de açúcar no meio de fermentação, temperaturas de 33°C a 35°C e o ambiente anaeróbico contribuem para a maior eficiência na produção de etanol (Wheals *et al*, 1999).

No Brasil a fermentação alcoólica ocorre em condições sépticas, embora em algumas indústrias, o caldo de cana pode passar por um tipo de pasteurização, tornando assim, inevitável à presença de microrganismos indesejáveis (Guerra & Angelis, 1998). O inóculo de fermentação fica vulnerável à contaminação microbiana, principalmente por bactérias e leveduras selvagens, que acompanham a matéria-prima utilizada no processo (Parazzi *et al*, 2002) ou que são provenientes de pontos na linha de produção (Ceccato-Antonini & Silva, 2000). A contaminação pode se instalar como foco contínuo resultando em reduções significativas na eficiência industrial (Ceccato-Antonini & Silva, 2000).

Leveduras *S. cerevisiae* dominam o processo de fermentação alcoólica e já existem até linhagens comerciais específicas para esta fermentação (Lallemand Inc., Canadá; AEB Group, Brasil). Entretanto, episódios de contaminação ao longo da safra podem fazer com que espécies que não sejam a levedura selecionada para a condução da produção de álcool, substituam o inóculo inicial (Cabrini & Gallo, 1999; Granchi *et al*, 1999; Silva-Filho *et al*, 2005^A). O primeiro trabalho o qual identificou espécies associadas à cana-de-açúcar e ao caldo pertenciam aos gêneros *Candida*, *Endomyces*, *Hansenula*,

Kloeckera, *Pichia*, *Saccharomycoides*, *Schizosaccharomyces* and *Torulopsis* e somente mais tarde os gêneros de *Debaromyces*, *Rhodotorula* e *Cryptococcus* foram identificados (Azeredo *et al*, 1998). A levedura *Dekkera bruxellensis*, normalmente documentada como a principal contaminante na produção de vinho (Cocolin *et al*, 2004), parece despontar como um importante contaminante na fermentação de álcool combustível (Guerra & Angelis, 1998; Araújo *et al*, 2005; Basílio *et al*, 2005; Liberal *et al*, 2007).

Dekkera bruxellensis

Leveduras do gênero *Brettanomyces*, ou seu teleomorfo (estágio sexual) *Dekkera*, foram descritos pela primeira vez por Claussen em 1904, na produção de cerveja (*apud* Rodrigues *et al*, 2001). Desde então, várias publicações têm evidenciado o isolamento desse gênero nas produções de vinhos, cervejas e outras bebidas em países como Itália, França, Austrália, Nova Zelândia e Estados Unidos (Rodrigues *et al*, 2001; Abbott *et al*, 2005)

Esses microorganismos são classificados como pertencendo ao reino Fungi, divisão Ascomycota (Saccaromycota), sub-divisão Ascomycotina (Saccharomycotina), classe Saccharomycetes, ordem Saccharomycetales, família Saccharomycetacea, gênero *Dekkera* (Barnett *et al*, 1990).

Muitas variedades de espécies de leveduras têm sido associadas a estragos na indústria de vinho. Destas, as espécies pertencente ao gênero *Dekkera* representam ser o maior dos problemas (Comitini *et al*, 2004; Cocolin *et al*, 2003). As leveduras de *D. bruxellensis* podem produzir compostos voláteis fenólicos, ácidos acéticos e tetrahidropiridinas. Outros aspectos, como

a fisiologia e a ecologia desse gênero, ainda permanecem incertos (Cocolin *et al*, 2003; Rodrigues *et al*, 2001).

Dias *et al* (2002) tinham como objetivo isolar linhagens capazes de crescer em meio de cultura DBDM (*Dekkera/Brettanomyces* diferencial medium) e pela habilidade de produzir 4-etilfenol em meio sintético contendo ácido. Entretanto espécies *P. guilliermondii*, *C. cantarelli*, *C. wickerhamii*, *D. hansenii* e *K. lactis* apresentaram as mesmas características que *D. bruxellensis* e *D. anomala* em meio DBDM, apresentado apenas diferenças na taxa de produção de 4-etilfenol.

Vários métodos moleculares de detecção de leveduras contaminantes têm sido desenvolvidos, valendo-se principalmente da facilidade de execução e da rapidez e confiabilidade dos resultados. Reações de PCR utilizando os primers específicos DB90F/DB394R foram desenvolvidas para a identificação de *Brettanomyces/Dekkera bruxellensis* e *Brettanomyces/Dekkera anomalus* em diversas amostras (Cocolin *et al*, 2003). O trabalho de Basílio *et al*. (2008) mostrou que a técnica de PCR fingerprinting com o primer (GTG)₅ foi capaz de diferenciar 29 padrões de bandas distintos para leveduras contaminantes. Estes isolados não-*Saccharomyces cerevisiae* foram identificados pelo seqüenciamento dos domínios D1/D2 da região 26S do rDNA. A levedura contaminante identificada como sendo da espécie *D. bruxellensis* foi o padrão de amplificação não-*S. cerevisiae* mais freqüentemente encontrado no monitoramento, estando presente em 52% de todas as amostras com contaminação. A levedura *D. bruxellensis* foi responsável também pela maior parte dos episódios de contaminação severa nas diferentes destilarias estudadas (Basílio *et al*, 2008), bem como em destilarias da América do Norte

(Abbott *et al*, 2005). Esta espécie apresenta posição de destaque nos problemas de qualidade do vinho e de contaminações dos processos fermentativos em geral (Mitrakul *et al*, 1999). Análises fisiológicas de isolados desta espécie mostraram uma reduzida capacidade fermentativa em relação à levedura do processo *S. cerevisiae*. Estas características fazem com que destilarias com o mosto fermentativo contaminados com esta levedura apresentem tanto queda no rendimento em etanol quanto aumento no tempo de fermentação (Araújo *et al*, 2005).

Liberal *et al* (2007) analisaram colônias isoladas do processo fermentativo alcoólico através das técnicas de PCR fingerprinting utilizando o primer de microssatélite GTG5, seqüenciamento da região D1/D2 26S rDNA e análise de PCR espécie específico com os primers DB90F e DB394R, mostrando que a espécie não *S. cerevisiae* mais freqüente correspondia à espécie *D. bruxellensis*. A dinâmica populacional dessa levedura foi acompanhada durante todo o período de safra, mostrando o potencial da *D. bruxellensis* de crescer mais rápido do que a *S. cerevisiae* em condições industriais causando freqüentemente severos episódios de contaminação.

Um levantamento preliminar do seqüenciamento do genoma de uma linhagem de *D. bruxellensis* foi realizado por pesquisadores da Suécia obtendo uma cobertura inicial de 40% do genoma total. Os pesquisadores identificaram aproximadamente 3,000 genes. Uma análise comparativa de sítio a sítio mostrou que 49% dos aminoácidos possuíam similaridade a ortólogos de *Saccharomyces*. Análises filogenéticas sugerem que *D. bruxellensis*, *S. cerevisiae* e *Candida albicans* possuem relações bem próximas formando um agrupamento filogenético com três ramos equivalentes. (Woolfit *et al*, 2007).

2.3 Bagaço, Celulose e Beta- glucosidase

O bagaço da cana-de-açúcar é constituído de 32 a 44% de celulose, um polímero da glicose formado por seis carbonos (hexoses), de 27 a 32% de hemicelulose, polímero de açúcar composto por cinco carbonos (pentose) e de 19 a 24% de lignina (Marques, 2009; Soccol *et al*, 2010). O composto formado pela lignocelulose pode conter em seu peso seco de 5-20% dos açúcares xilose e arabinose, entretanto a xilose é a mais abundante, a qual não pode ser fermentada naturalmente pela levedura *Saccharomyces cerevisiae* (Hahn-Hagerdal *et al*, 2006). Essas fontes de carbono, celulose e hemicelulose, correspondem por dois terço da energia da planta que não são convertidos a bioetanol (Marques, 2009).

Com a tecnologia de hidrólise poderia ser gerado um adicional de 150 mil litros de etanol por dia e com o aperfeiçoamento da técnica poderia chegar a 400 mil litros em 2025 (Marques, 2009). A produção de etanol combustível a partir da biomassa de lignocelulose inclui as etapas de pré-tratamento, hidrólise, fermentação e separação. Para que a biomassa possa ser utilizada por microorganismos é necessário ser submetidos a um pré-tratamento capaz de desorganizar o complexo. Essa etapa é um ponto chave da operação para uma conversão eficiente de lignocelulose em etanol devido à lignina ser um grande obstáculo capaz de inibir que a fermentação ocorra por dificultar a acessibilidade dos carboidratos da planta ao processo biológico como a hidrolise enzimática . O papel principal do método de pré-tratamento é para diminuir as interações dos compostos da parede celular fazendo com que a celulose e a hemicelulose fiquem mais disponível para a hidrolise enzimática e

assim tornar a fermentação susceptível. Um dos métodos para o pré-tratamento é a explosão de vapor, na qual os carboidratos são expostos através da alta pressão do vapor sob condições específicas (Soccol *et al*, 2010).

Alem da acessibilidade dos carboidratos, existem outros gargalos na produção de etanol a partir bagaço, como a hidrólise da celulose, aproveitamento dos açúcares de cinco carbonos e os resíduos da destilação. Desenvolvimentos de tecnologias para o pré-tratamento do bagaço, enzimas para aumentar a sacarificação de celulose/hemicelulose e acessibilidade para a fermentação de açúcares C6 e C5 estão sendo realizadas (Marques, 2009, Soccol *et al*, 2010, Hahn-Hagerdal *et al*, 2006).

Após a fermentação, à parte da lignina, a principal parte sólida da biomassa, pode ser queimada para a produção de energia elétrica para ser utilizada no processo industrial (Hahn-Hagerdal *et al*, 2006).

A celulose é o composto polimérico principal da parede celular das plantas, o polissacarídeo mais abundante na terra, e uma importante fonte renovável. A composição química desse composto é bastante simples, são formados por um polímero linear de mais de 1000 resíduos de D-glucose interligados por ligações β -1,4 glicosídicos (Baldrian *et al*, 2008).

Em um ambiente rico em celulose, os fungos filamentosos do grupo dos Basidiomicetos são os mais eficientes em degradar madeiras mortas e detritos. O sistema de degradação dos fungos difere do sistema das bactérias, entretanto essas diferenças são menos evidentes dentro do mesmo grupo taxonômico (Lynd *et al*, 2002).

A celulose é hidrolisada a celobiose pelas enzimas celobiohidrolase/ exo-glucanase e endoglucanase, dependendo do local do sítio de ação da enzima no polímero. Na segunda reação, a celobiose é hidrolisada em duas moléculas de glicose pela enzima beta-glucosidase (EC 3.2.1.21) também conhecida como celobiase, como mostra a figura 1 (Soccol *et al*, 2010). Como celobiose é um substrato bastante disponível, as enzimas β -glucosidases são produzidas pela maioria dos microrganismos. Foram isoladas ou detectadas atividades dessas enzimas nas leveduras *Rhodotorula minuta* e *Sporobolomyces singulares* (Lynd *et al*, 2002; Onishi & Tanaka, 1996; Ishikawa *et al*, 2005).

Beta-glucosidases isoladas do grupo dos Basidiomicetos exibem alta variabilidade estrutural refletido da localização da enzima, se intracelular ou extracelular, com variações em massa que vão de 35 a 640kDa, sendo as enzimas menores em torno de 100kDa geralmente monômeros e extracelulares. A enzima β -glucosidase do fungo *Trametes versicolor* é uma glicoproteína grande de 300kDa, monomérica que apresenta cerca de 90% de glicosilação. A atividade enzimática da β -glucosidase é altamente inibida pela glucose, gluconolactone e celobionolactone. Seu pH ideal varia entre 3,5 e 5,5, com exceção da enzima intracelular da espécie *Phanerochaete chrysosporium* que possui um pH neutro. A temperatura ideal é entre 45 e 75°C (Baldrian *et al*, 2008). As enzimas produzidas pelas leveduras *Rhodotorula minuta* e *Sporobolomyces singulares* são capazes também de clivar os monossacarídeos fucose e galactose e o dissacarídeo lactose (Onishi & Tanaka, 1996; Ishikawa *et al*, 2005).

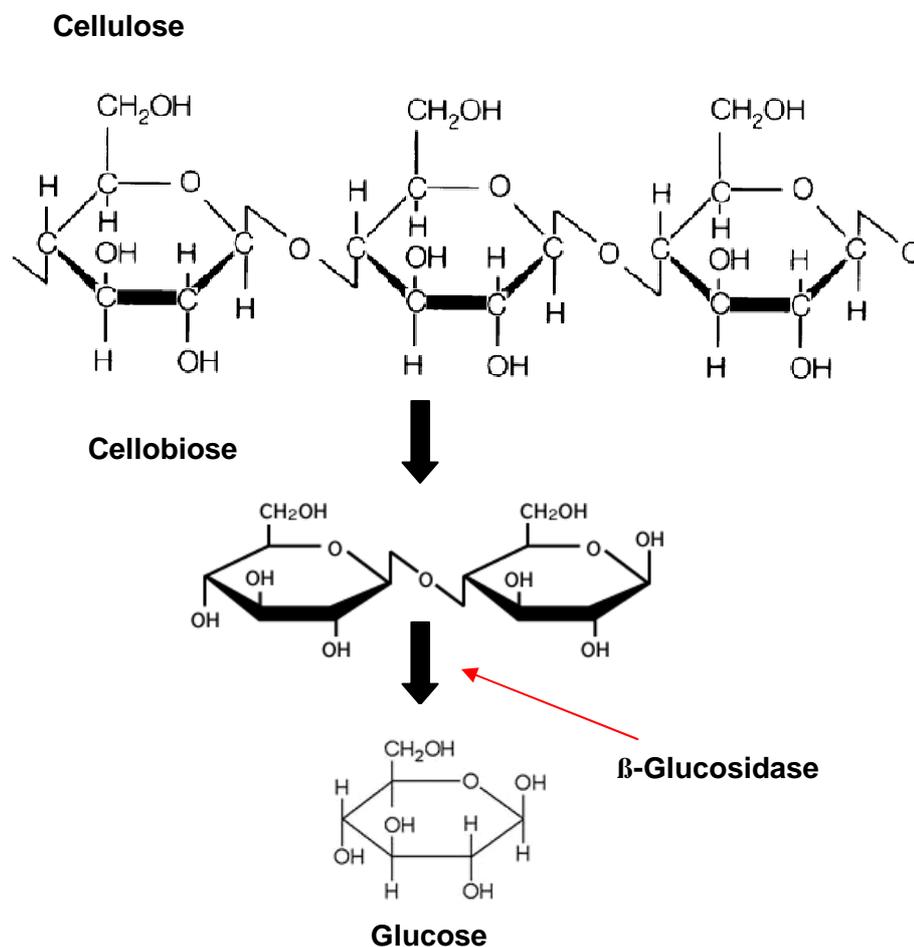


Figura 1. Esquema da hidrólise da celulose. Na primeira reação à celulose é hidrolisada enzimaticamente em celobiose que é hidrolisado em glucose pela beta-glicosidase. O polímero da celulose, a enzima possui especificidade pelas ligações β -1,4-glicosídicos podendo ser classificadas como exo-glucanase ou celobiohidrolase ou endoglucanase dependendo do sítio ativo.

O cristal da enzima de β -glucosidase da levedura *K. marxianus* tem a estrutura de um tetrâmero (Figura 2) e é composto por 721 aminoácidos distribuídos em três domínios conectados por duas estruturas de ligação (linkers). O sítio ativo é localizado na interface dos domínios 1 e 2 numa abertura de aproximadamente 7Å de largura e 18Å de profundidade e tem o sítio -1 bem definido que é o local onde a molécula de glicose se liga. Os resíduos do sítio ativo da enzima estão posicionados nos sítios D58, R130, K163, H164, R174, Y210 e E458 (Figura 3). O sítio D58 é o resíduo chave envolvendo a acomodação do substrato no sítio -1 formando pontas de hidrogênio com a hidroxila do açúcar e é uma região bastante conservada. O resíduo E458 no domínio 2, na extremidade oposta do substrato, é considerado o catalítico ácido/base (Pozzo *et al*, 2010).

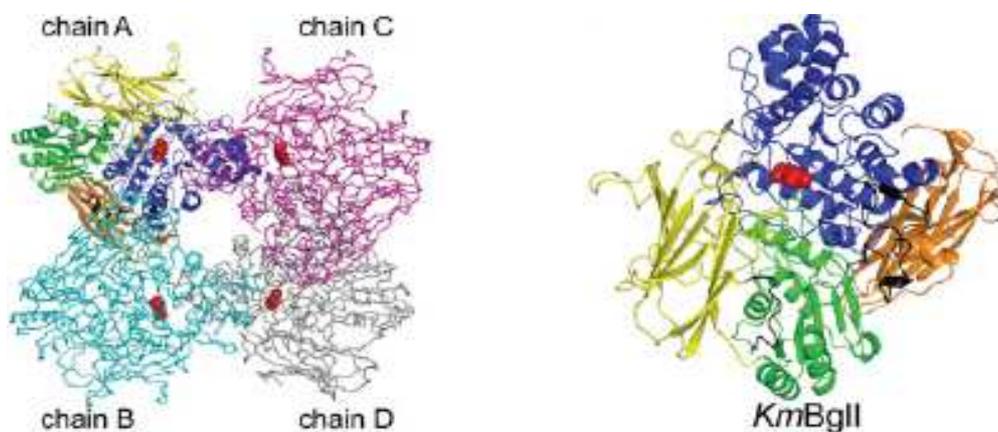


Figura 2. Estrutura da beta-glucosidase de *K. marxianus*. Esquerda. Estrutura na forma de tetrâmero. Direita. Estrutura do monômero. Em azul, domínio GH3 N-terminal. Em verde, domínio α/β – sanduíche. Em amarelo, domínio PA14. Em laranja, domínio GH3 C-terminal. Yoshida *et al*, 2010.

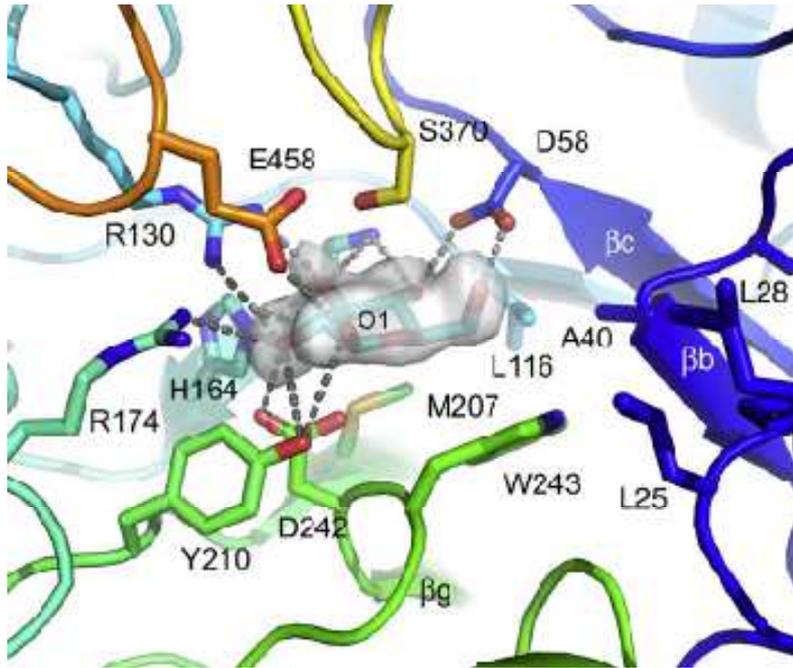


Figura 3. Estrutura detalhada do sítio ativo da enzima beta-glucosidase da bactéria *Thermotoga neapolitana*. Glicose está mostrada com uma superfície transparente no meio da figura (O1). Pozzo *et al*, 2010.

2.4 Técnicas de Bioinformática

De acordo com a teoria da evolução, todos os organismos desenvolveram de um único ancestral comum e através de mecanismos de variações adquiriram a diversidade atual. Para se investigar a evolução e o parentesco entre genes e organismos, diferentes tipos de dados podem ser usados, o modo clássico, por comparação entre espécies por caracteres morfológicos, ou pelo modo molecular utilizando informações genéticas com os nucleotídeos e as seqüências de aminoácidos. Comparações taxonômicas mostram que genes de espécies com parentescos próximos geralmente diferem apenas por números limitados de mutações, enquanto que genes

menos relacionados diferem por números maiores de substituições. Ainda existem alguns genes que são mais conservados que outros, especialmente aqueles que codificam sítios catalíticos (Lemey *et al*, 2009).

Com o surgimento dos seqüenciadores automáticos de DNA, houve uma explosão na quantidade de seqüências a serem armazenadas, exigindo recursos computacionais cada vez mais eficientes. Além do armazenamento ocorria, paralelamente, a necessidade de análise desses dados, o que tornava indispensável à utilização de plataformas eficientes para a interpretação dos resultados obtidos. As ferramentas computacionais permitiram converter dados genômicos em conhecimentos para ser usado pelos pesquisadores unindo a ciência da computação e a biologia molecular (Rouzé *et al*, 1999; Prosdocimi *et al*, 2002).

Em conseqüência da grande quantidade de seqüências de nucleotídeos e aminoácidos produzidos em projetos genômicos, proteômicos e transcriptômicos foram criados depositórios internacionais de dados públicos e privados. Os bancos de dados mais utilizados são: o GenBank o qual possui seqüências de DNA e proteínas, o European Bioinformatics Institute (EBI) com seqüências de DNA; o DNA Data Bank of Japan (DDBJ), o Protein Data Bank (PDB) o qual armazena estruturas tridimensionais de proteínas, o Kyoto Encyclopedia of Genes and Genomes (KEGG) com dados de seqüências de genomas de vários organismos diferentes e informações relacionadas as suas vias metabólicas (Prosdocimi *et al*, 2002). Os bancos de dados GenBank, DDBJ e EMBL compõem o International Nucleotide Sequence Databases (INSD – Banco de dados de seqüências de nucleotídeos internacional), os quais permitem a permuta dados entre eles diariamente para servir a

comunidade científica com a mesma qualidade e quantidade de dados (Tateno *et al*, 2002).

O GenBank é um banco de dados público de seqüências de nucleotídeos e proteínas com suportes bibliográficos e anotações biológicas construídas e distribuídas pela National Center for Biotechnology Information (NCBI). Além das seqüências, o NCBI disponibiliza informações sobre taxonomia, mapeamento, estrutura, domínio de proteínas e ainda fornece a ferramenta BLAST (Basic local Alignment Search Tool) que proporciona buscas de seqüências similares no GenBank ou em outros bancos de dados (Benson *et al*, 2002).

O Projeto *Dekkera bruxellensis* (PDB) foi elaborado pelo Laboratório Genômica e Expressão (LGE) da Universidade de Campinas (UNICAMP) a partir de seqüências disponibilizadas pelo Professor Jure Piskur da Universidade de Lund na Suécia. O banco de dados PDB possui depositadas 40% do genoma de *D. bruxellensis* e tem como objetivo permitir anotações e análises das seqüências depositadas (Carazzolle *et al*, 2007). A biblioteca genômica foi realizada e seqüenciada por Agowa (Berlin, Alemanha). As seqüências foram conseguidas através de digestão enzimática e clonagem em vetor pMCL210. A leitura das seqüências foi realizada através do programa Phred e montadas em contigs utilizando o programa Phrap. A identificação dos genes foi realizada comparando as seqüências com *Saccharomyces* Genome Database e contigs maiores que 100 nucleotídeos foram anotados como ortólogos de *Saccharomyces cerevisiae* (Woolfit *et al*, 2007).

Na área da bioinformática, o alinhamento de seqüências possui uma diversidade de aplicações sendo considerada uma das operações mais

importante desta área. Esse método de comparação procura determinar o grau de similaridade entre duas ou mais seqüências (Prosdocimi *et al*, 2002), seqüências homólogas sob investigação se alinham a fim de que sítios homólogos formem colunas (Lemey *et al*, 2009). O alinhamento global é freqüentemente utilizado para determinar regiões mais conservadas de seqüências ortólogas. O programa mais utilizado para esse tipo de tarefa é o Clustal X (Thompson *et al*, 1997) e a versão web (Clustal W - <http://www.ebi.ac.uk/clustalw/index.html>) (Prosdocimi *et al*, 2002). O princípio geral do Clustal W é o alinhamento múltiplo de seqüências de DNA ou proteínas produzindo significados biológicos para seqüências divergentes. O programa determina a melhor pontuação obtida por cada unidade pareada (match) nas seqüências selecionadas e as alinham de forma que as identidades, similaridades e diferenças sejam visualizadas (Clustal W, 2007). O programa Bioedit é um editor de seqüências biológicas que foi planejado para procedimentos básicos na edição, alinhamento, manipulação e análise de seqüências de nucleotídeos e aminoácidos, não sendo, entretanto, um programa poderoso de análises, mas oferece muitas funções rápidas e de fácil manuseio para anotações e manipulação das seqüências (Bioedit, 2007).

O programa Molecular Evolutionary Genome Analysis (MEGA) tem sido utilizado no alinhamento de seqüências, estimação de divergência, e visualização de árvores filogenéticas. A versão MEGA 4 faz inferências filogenéticas com métodos de distância evolutivas e parcimônia testando hipóteses evolucionárias (Prosdocimi *et al*, 2002, Tamura *et al*, 2007). O programa disponibiliza os algoritmos mais utilizados que são o UPGMA (Unweighted Pair Group Method with Arithmetic Means) e o Neighbor-Joining

(Agrupamento de Vizinhança) que tem o propósito de construir árvores filogenéticas através de dados de distância. O princípio desse método é encontrar pares das unidades taxonômicas funcionais que minimize a extensão das ramificações totais para cada estágio de agrupamento começando a partir de uma árvore no formato de estrela (Tamura *et al*, 2007; Saitou *et al*, 1987; Lemey *et al*, 2009). O programa PhymI utiliza o método máxima verossimilhança (ML), o qual examina diferentes topologias de árvores e avalia os diferentes parâmetros somando todas as possibilidades nucleotídica ou de aminoácidos nas posições das seqüências. A árvore que tiver a maior semelhança, ou máxima probabilidade, será a escolhida. Esse algoritmo requer bastante tempo para estimar todos os parâmetros, entretanto, o programa PhymI reduz o tempo de processamento por utilizar o algoritmo BioNJ para estimar a árvore inicial, otimiza a árvore através do FastNNI e por fim escolhe a árvore através da máxima probabilidade com os valores ML (Lemey *et al*,2009)

A cada vez que um programa de filogenia molecular é utilizado para gerar uma árvore sobre um conjunto de dados escolhidos, o resultado pode ser diferente. Por isso, para validar uma árvore filogenética se faz inúmeras repetições com o programa escolhido e se testa cada ramo, estatisticamente, para escolher um a um aquela com a maior probabilidade de ocorrência para a composição final da árvore. Assim funciona o Bootstrap gerando conjuntos modificados de dados, obtidos aleatoriamente a partir dos dados do alinhamento. Probabilidades altas (> 50%) indicam que, mesmo com algumas alterações, os dados suportam o ramo ao qual essa probabilidade se refere, já probabilidades baixas (< 50%) significam que não se pode ter certeza de que determinado ramo seja correto (Prosdocimi *et al*, 2002).

3. OBJETIVOS

Objetivo Geral

Este trabalho tem como objetivo identificar e analisar o gene beta-glucosidase da levedura contaminante *Dekkera bruxellensis* na fermentação alcoólica.

Objetivos Específicos

- a) Identificar os resíduos do sítio ativo da seqüência de *Dekkera bruxellensis*;
- b) Identificar a presença de domínios específicos da proteína beta-glucosidase na seqüência de *Dekkera bruxellensis*.
- c) Identificar as relações filogenéticas do gene beta-glucosidase de *Dekkera bruxellensis*.
- d) Predição da estrutura da proteína beta-glucosidase de *Dekkera bruxellensis*.

4. MATERIAIS E MÉTODOS

4.1 Seleção e Análise de Seqüências.

Seqüências de *Dekkera bruxellensis* foram retiradas do banco de dados do Laboratório de Genômica e Expressão - Projeto Dekkera Bruxellensis (PDb) (<http://www.lge.ibi.unicamp.br/dekkera/>). As buscas foram realizadas através das palavras-chave hidrolase, amylase, glucosidase e amyloglucosidase. Os contigs encontrados foram comparados contra o banco de dados NCBI usando a ferramenta tBLASTn. Apenas as seqüências codificantes completas que apresentaram alinhamentos com e-value -20 ou menor foram resgatadas. A seqüência do contig 145 de *D. bruxellensis* foi selecionada e analisada usando a ferramenta do ORF Finder do NCBI (<http://www.ncbi.nlm.nih.gov/>) e a seqüência de aminoácido no frame -2 foi selecionada. A seqüência de aminoácido foi comparada no banco de dado NCBI usando a ferramenta BLASTp e seqüências identificadas sendo codificante para a enzima beta-glucosidase com um e-value -20 ou menor foram selecionadas. O banco também permitiu identificar os domínios presentes na seqüência.

4.2 Alinhamento Múltiplo.

As seqüências encontradas foram agrupadas originado um arquivo multi-fasta. O alinhamento múltiplo foi feito utilizando o programa CLUSTALw e editado manualmente no programa BIOEDIT 7.0 e MEGA 5. Os programas Mega 5 e PHYLM foram utilizados para a construção das árvores filogenéticas

aplicando os algoritmos neighbor-joining e máximo verossimilhança com testes estatísticos bootstrap com 1000 replicações.

4.3 Resíduos de Sítio Ativo.

A identificação dos resíduos do sítio ativo do contig 145 de *D. bruxellensis* foi realizada a partir de comparação através do trabalho de Pozza *et al* (2010) utilizando alinhamento múltiplo entre as seqüências da bactéria *Thermotoga neapolitana* e da levedura *Kluyveromyces marxianus*. O programa de visualização de estrutura protéica Cn3 (<http://www.ncbi.nlm.nih.gov/>) foi utilizado para a visualização dos resíduos do sítio ativo da proteína beta-glucosidade da bactéria *Thermotoga neapolitana* e da levedura *Kluyveromyces marxianus* presentes na seqüência do contig 145 da levedura *D. bruxellensis*. O programa on-line Swiss-model foi utilizado para modelamento protéico a partir da seqüência do contig 145 de *D. bruxellensis*.

4.4 Análise Filogenética.

As análises filogenéticas foram realizadas utilizando os programas Mega 4 e Phylml. No programa Mega 5 foram utilizados os recursos alinhamento, construção e visualização das árvores filogenéticas aplicando o algoritmo Neighbor-joining com testes estáticos bootstrap com 1000 replicações. O programa Phylml foi utilizado para a construção de árvores filogenética usando o algoritmo Máximo Verossimilhança com testes estáticos bootstrap com 1000 replicações. Para identificar o melhor modelo evolutivo a ser utilizado nas

nossas seqüências de aminoácido, foi rodado o programa Prot Test, o qual sugeriu o modelo LG+G.

5. RESULTADOS

5.1 Seleção das Seqüências.

A partir da ferramenta de busca do banco PDb foram identificados nove contigs (1001, 816, 145, 1662, 1684, 3105, 3173, 470 e 936) através das palavras-chave hydrolise, amylase, glucosidase e amyloglucosidase. Um contig em particular, contig 145, foi selecionado devido análise preliminar realizada utilizando a ferramenta de tBlastx do banco de dados Genolevures inferia como sendo similar a seqüência codificante para o gene da β -glucosidase. A seqüência nucleotídica do contig 145 foi submetida a ferramenta de análise gráfica ORF Finder do banco de dados NCBI (<http://www.ncbi.nlm.nih.gov/projects/gorf/>) a fim de identificar a região de quadro de leitura. A ferramenta inferiu a região codificante da seqüência no frame -2 (Anexo).

A seqüência codificante da levedura *D. bruxellensis* foi, então, comparada nos bancos de dados Genolevures e NCBI. As seqüências ortológicas completas que indicavam sendo codificantes da enzima β -glucosidase e que apresentaram e-value ≤ 20 ou menor foram selecionadas no formato fasta gerando um arquivo multi-fasta (Tabela 1). Foram selecionadas uma espécie de bactéria da família Thermatogaceae e 13 espécies do grupo ascomiceto pertencentes as sub-divisões Saccharomycotina e Pezizomycotina. Dentro da sub-divisão Saccharomycotina foram identificados às famílias Debaryomycetaceae e Saccharomycetaceae através dos gêneros *Scheffersomyces*, *Schwanniomyces* e *Kluyveromyces*, *Pichia* e *Candida*,

respectivamente. Dentro da sub-divisão Pezizomycotina foi identificada a família Trichocomaceae representada pelos gêneros *Aspergillus*, *Talaromyces* e *Neosartorya*.

5.2 Alinhamento Múltiplo

O alinhamento múltiplo possibilitou determinar as regiões similaridades, divergências e inclusões de gaps nas seqüências codificantes da enzima β -glucosidase das espécies selecionadas. O programa Bioedit, possibilitou ainda, a construção de uma matriz de identidade a qual fornece o percentual de similaridade entre as seqüências através da ferramenta sequence identity matrix. A matriz de identidade evidenciou uma similaridade de aproximadamente 61% entre o contíg 145 de *D. bruxellensis* com o gene da enzima β -glucosidase de *K. marxianus*, e de aproximadamente 36%, com o gene da bactéria *Th. neapolitana* (Tabela 2).

Tabela 1. Identificação das seqüências do gene Beta-glucosidase dos microrganismos selecionados.

Espécie	N° de acesso do GenBank	Lenght (aa)	score	e-value	Identities	Positivities	Gaps
<i>Aspergillus clavatus</i>	XP_001274045.1	838	310	3E-97	53%	71%	2%
<i>Aspergillus flavus</i>	XP_002378447.1	856	313	4E-98	51%	71%	2%
<i>Aspergillus fumigatus</i>	XP_753915.1	838	313	2,00E-98	53%	71%	2%
<i>Aspergillus niger</i>	XP_001398281.2	839	315	6E-99	53%	73%	2%
<i>Aspergillus oryzae</i>	XP_001823113.1	839	315	2E-98	51%	71%	2%
<i>Candida tropicalis</i>	XP_002547835.1	837	340	2E-108	57%	76%	2%
<i>Kluyveromyces marxianus</i>	ACY95404.1	845	374	1E-101	61%	79%	1%
<i>Neosartorya fischeri</i>	XP_001259962.1	838	314	1E-98	53%	71%	2%
<i>Pichia angusta</i>	EFW97210.1	1702	330	1E-100	55%	72%	2%
<i>Scheffersomyces angusta</i>	XP_001385159.1	839	350	2E-112	57%	77%	2%
<i>Schwanniomyces etchellsii</i>	ACF93471.1	847	363	1E-98	59%	79%	1%
<i>Talaromyces stipitatus</i>	XP_002481927.1	842	307	5E-96	52%	70%	4%
<i>Thermotoga neapolitana</i>	ABI29899.1	721	208	3E-58	40%	60%	4%
<i>Verticillium albo-atrum</i>	XP_003009697.1	1226	329	1E-101	54%	72%	2%

Tabela 2. Similaridades das seqüências codificantes da enzima beta-glucosidase dos microrganismos selecionados. Bioedit 7.

Percentual de similaridade entre as Seqüências de Beta-glucosidade															
	Db	Km	Se	Sa	Ct	Va	Pa	An	Nf	Af	Ao	Ts	Afl	Ac	Tn
<i>Db_Contig145</i>	100	60.9	58.5	56.2	55.3	53.5	54.5	51.8	51.8	52.1	51.5	51.5	51.5	52.1	35.9
<i>K. marxianus</i>		100	84.0	65.0	58.7	55.5	57.2	53.8	54.9	55.2	54.9	51.1	54.9	54.5	35.5
<i>S. etchellsii</i>			100	64.7	58.7	52.8	58.6	50.1	51.1	51.5	52.8	50.1	52.8	51.8	35.5
<i>S. stipitis</i>				100	70.8	56.7	66.5	56.2	55.8	56.5	57.9	51.7	57.9	56.5	37.3
<i>C. tropicalis</i>					100	54.2	61.6	52.3	52.7	52.7	53.4	53.0	53.4	53.7	35.0
<i>V. albo-atrum</i>						100	53.2	61.2	63.3	62.6	59.1	61.9	59.1	62.2	39.2
<i>P. angusta</i>							100	54.8	55.5	55.8	56.5	55.8	56.5	56.2	38.6
<i>A. niger</i>								100	88.8	87.8	86.4	71.8	86.4	85.7	36.6
<i>N. fischeri</i>									100	97.5	87.1	72.5	87.1	92.0	37.6
<i>A. fumigatus</i>										100	87.1	71.8	87.1	91.6	37.3
<i>A. oryzae</i>											100	72.2	100	86.8	38.3
<i>T. stipitatus</i>												100	72.2	70.8	39.0
<i>A. flavus</i>													100	86.8	38.3
<i>A. clavatus</i>														100	37.3
<i>Th. neapolitana</i>															100

Db- *Dekkera bruxellensis*; Km- *Kluyveromyces marxianus*; Se- *Schwanniomyces etchellsii*; Sa- *Scheffersomyces angusta*; Ct- *Candida tropicalis*; Va- *Verticillium albo-atrum*; Pa- *Pichia angusta*; An- *Aspergillus niger*; Nf- *Neosartya fischeri*; Af- *Aspergillus fumigatus*; Ao- *Aspergillus oryzae*; Ts- *Talaromyces stipitatus*; Afl- *Aspergillus flavus*; Ac- *Aspergillus clavatus*; Tn- *Thermotoga neapolitana*.

5.3 Resíduos de Sítios Ativos.

O contig 145 da levedura *Dekkera bruxellensis* possui 279 resíduos de aminoácidos arranjados em dois domínios (Figuras 4,5). O primeiro foi o domínio N 3 terminal da família glicosil-hidrolase (resíduos 1-177) com um e-value 8.10e-60. O segundo domínio foi o C 3 terminal da família glicosil-hidrolase (resíduos 252-279), o qual apresentou um e-value de 3.90e-05. A representação da estrutura terciária da enzima β -glucosidase do contig 145 de *D. bruxellensis* também foi elaborada a partir do programa on-line Swiss-model, o qual selecionou a seqüência ortóloga do cristal de *K. marxianus* como sendo a seqüência mais similar. As seqüências das espécies *K. marxianus* e *D. bruxellensis* foram visualizados pelo programa Cn3 (figuras 6,7).

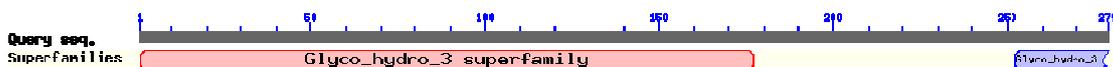


Figura 4. Domínios N-terminal e C-terminal encontrados na seqüência do contig 145 de *Dekkera bruxellensis* através do programa on-line Swiss-model (<http://swissmodel.expasy.org>).

A partir de dados da literatura foram identificados sete resíduos do sítio ativo que são D58 (ácido aspártico), R130 (arginina), K163 (lisina), H164 (histidina), R174 (arginina), Y210 (tirosina) e E458 (ácido glutâmico). Desses resíduos, foram identificados cinco presentes na seqüência da levedura *D. bruxellensis*, (Figura 8) os aminoácidos arginina, lisina, histidina e tirosina, apenas os aminoácidos polares carregados negativamente ácido aspártico e ácido glutâmico não puderam ser identificados devido a seqüências ser incompleta (Figura 9,10).

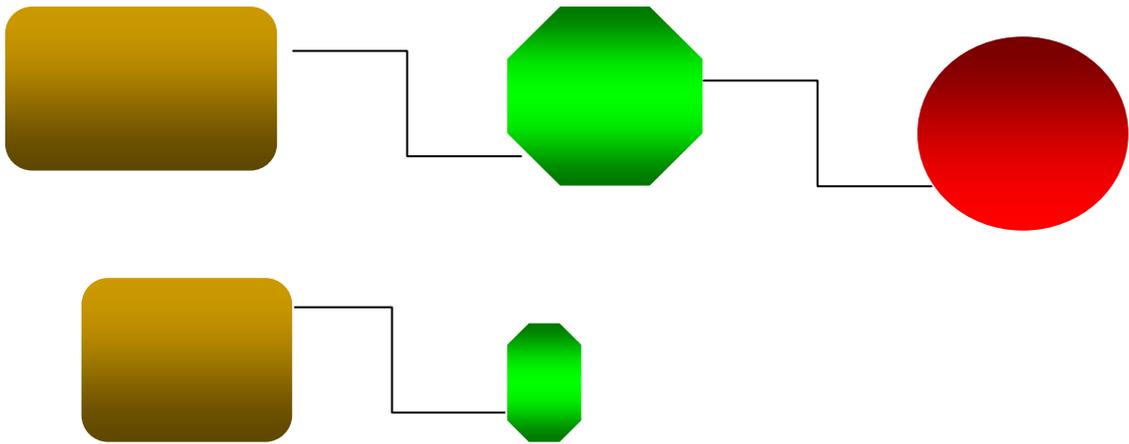


Figura 5. Representação dos domínios de Beta-glucosidase. (a) Beta-glucosidase de *Thermotoga neapolitana*, o retângulo representando o domínio 1 (resíduos 1-306), o octógono representando o domínio 2 (resíduos 321-536) e o círculo representando o domínio 3 (resíduos 600-721). (b) Contig 145 *D. bruxellensis*, retângulo representa o domínio 1 (resíduos 1-177) e o octógono representa o domínio 2 (resíduos 252-279).

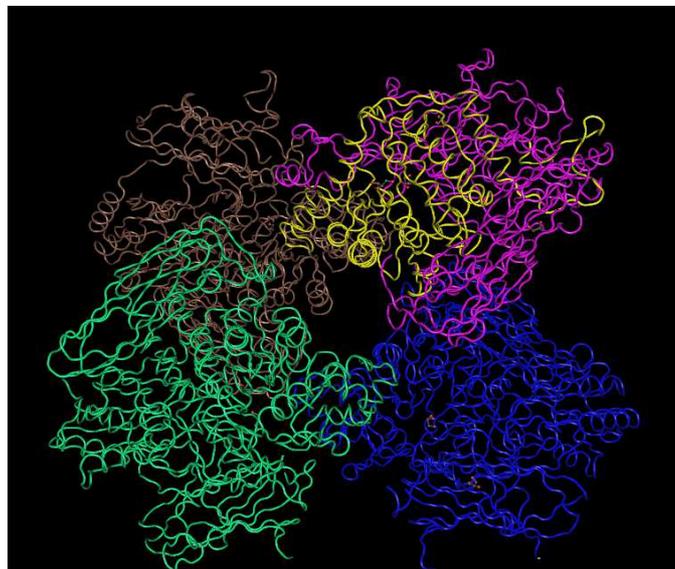


Figura 6. Visualização do tetrâmero da enzima β -glucosidase de espécie *Th. neapolitana* através do programa Cn3. A seqüência da enzima β -glucosidase da espécie *A. marxianus* encontra-se destacada em amarelo.

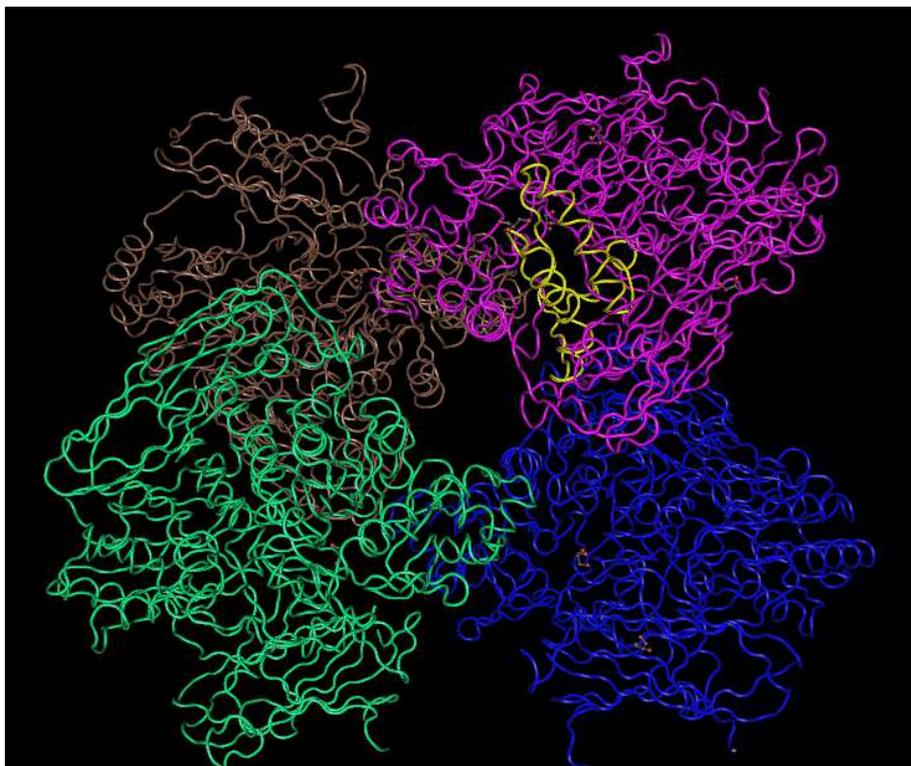


Figura 7. Visualização do tetrâmero da enzima β -glucosidase de espécie *Th. neapolitana* através do programa Cn3. Em amarelo, a seqüência do contig 145 da espécie *D. bruxellensis*.



Figura 8. Visualizaçãodo monômero da enzima β -glucosidase de espécie *K. marxianus* através do programa Cn3. Em amarelo, os resíduos do sítio ativo presentes na seqüência da espécie *D. bruxellensis*.

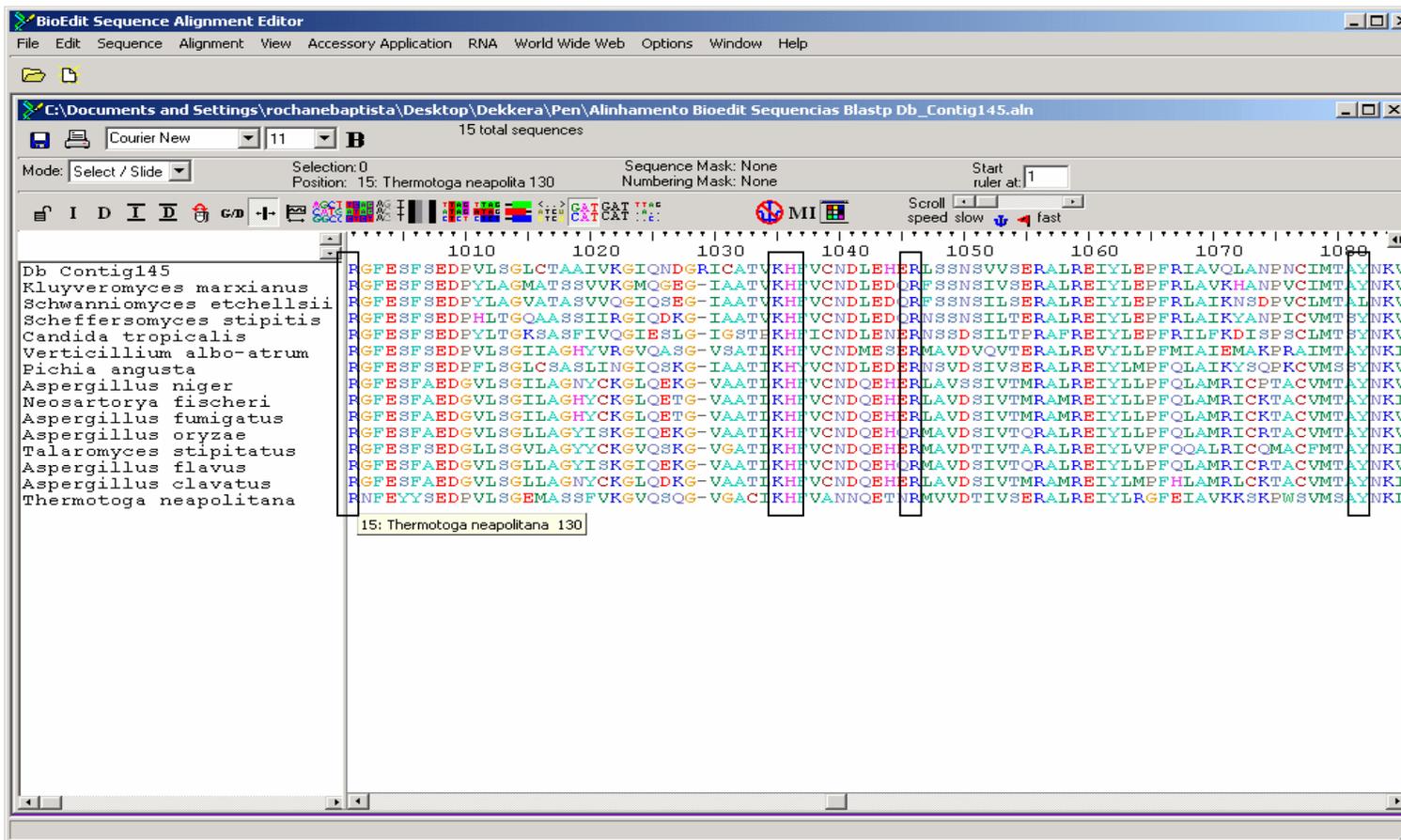


Figura 9. Alinhamento múltiplo das seqüências de beta-glucosidade evidenciando os resíduos do sitio ativo visualizados no programa Bioedit versão 7.0.

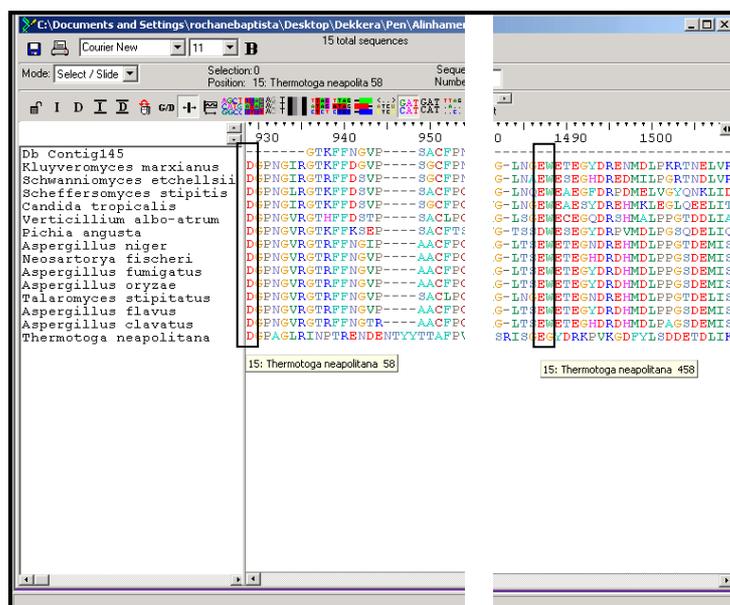


Figura 10. Alinhamento múltiplo das seqüências de beta-glucosidase evidenciando os resíduos D58 (Esq.) e E458 (Dir.) do sítio ativo encontrados na bactéria *Thermotoga neapolitana* visualizados no programa Bioedit versão 7.0.

5.4 Análise Filogenética.

Para verificar a evolução do gene da β -glucosidase dentro do grupo dos fungos, foram construídas árvores filogenéticas baseadas na seqüência de aminoácidos. Verificamos ambas as árvores construídas utilizando tanto o método neighbor-joining no programa mega 5 (Figura 11) e o método Máximo verossimilhança (Figura 12) no programa Phylm apresentaram clados bastantes semelhantes. A seqüência da espécie *D. bruxellensis* formou um clado com o clado feito pelas as espécies *K. marxianus* e *S. etchellsii* em 74% e 89% das repetições testadas, respectivamente. A seqüência do gene da enzima β -glucosidase da bactéria *Th. neapolitana* se posicionou na porção basal do dendrograma, ficando como um grupo externo para esse agrupamento.

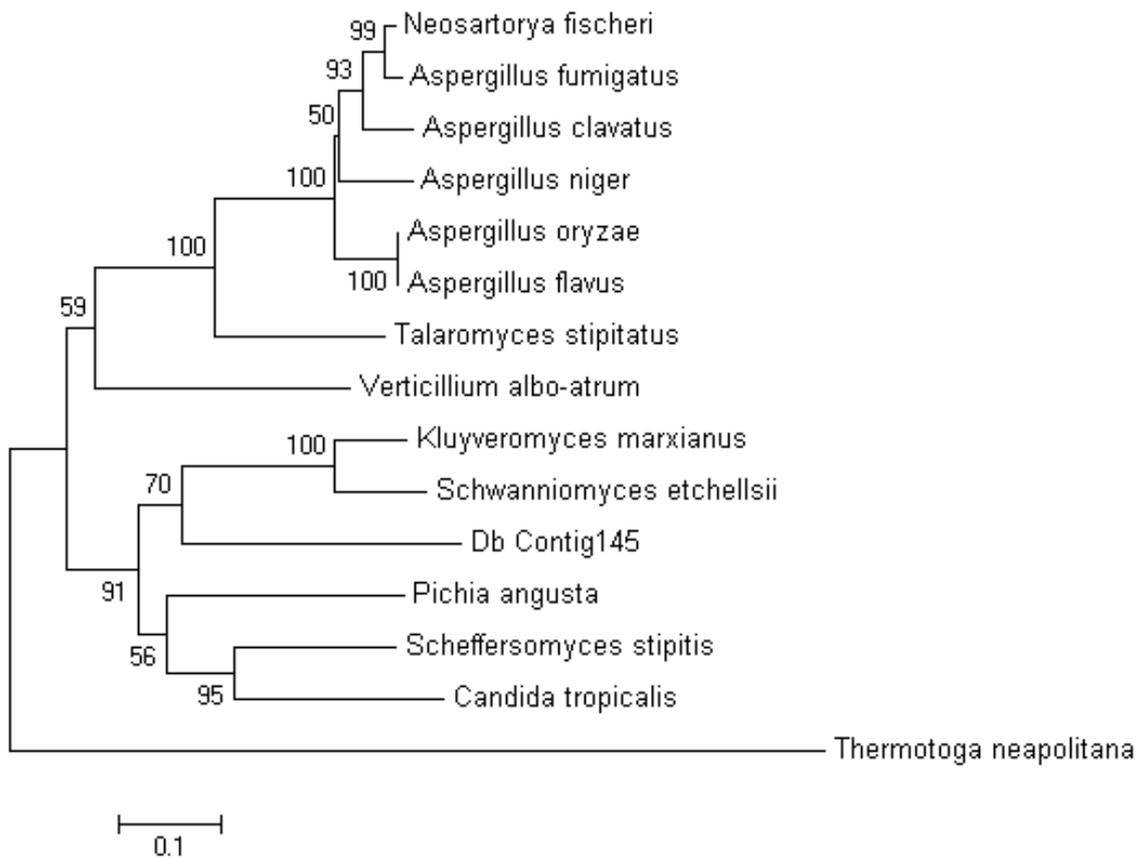


Figura 11. Árvore filogenética construída a partir das seqüências do gene da beta-glucosidase usando o agrupamento Neighbor-Joining do programa Mega 5.

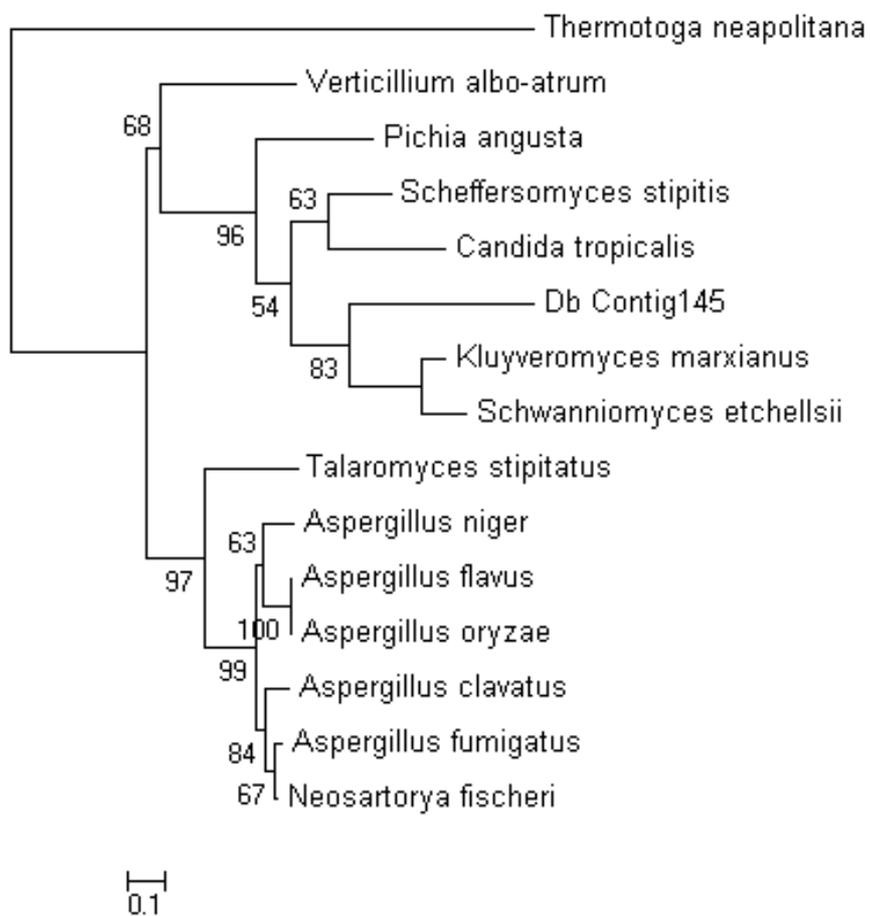


Figura 12. Árvore filogenética construída a partir das seqüências do gene da beta-glucosidase usando o agrupamento Maximo Verossimilhança do programa Phylml.

6. DISCUSSÃO

Fermentação industrial de lignocelulose em etanol requer microrganismos capazes de hidrolisar compostos orgânicos alternativos como fontes de carbono, além de glicose, mantendo uma alta produtividade e rendimento (Galafassi *et al*, 2010). A levedura *Saccharomyces cerevisiae* industrialmente utilizada em processos fermentativos na produção de álcool a partir de cana-de-açúcar é naturalmente incapaz de hidrolisar açúcares como pentoses ou polissacarídeos como amido (Moraes *et al*, 1995), entretanto, outras poucas leveduras têm sido caracterizadas pela capacidade fermentativa em açúcares pentoses como *Pichia stipitis* (Slininger *et al*, 2011; Agbogbo *et al*, 2008) . A levedura *Dekkera bruxellensis* tem sido freqüentemente apontada como sendo um microrganismo contaminante nas indústrias de vinho, cerveja e recentemente na produção contínua de álcool combustível (Woolfit *et al*, 2007; Liberal *et al*, 2005; Liberal *et al*, 2007). Devido à capacidade dessa levedura de tolerar altas concentrações de pH e etanol, o foco em torno dessa espécie tem sido maior, não mais como uma praga a ser erradicada mais como uma espécie de grande potencial econômico e industrial. Estudos com um complexo de cinco espécies *Brettanomyces/Dekkera* têm sido reportados mostrando capacidade de hidrolisar xilose, um açúcar composto por cinco carbonos (Galafassi *et al*, 2010 *apud* Parekh *et al*, 1988). Uma parceria entre pesquisadores da Itália e Suécia (Galafassi *et al*, 2010) estudou o potencial do grupo *Brettanomyces/Dekkera* na produção de etanol em fontes renováveis visando os processos industriais, como condições de baixo pH e oxigênio limitante. Eles avaliaram o crescimento de mais de 50 linhagens nas mais

variadas fontes de carbonos, incluindo sacarose, maltose, celobiose, lactose, amido, xilose, galactose e arabinose, além de testar a produção de etanol em hidrolisados de lignocelulose sintéticos. A maioria das linhagens de *D. bruxellensis* cresceu nos meios contendo galactose, maltose, manose, celobiose, amido e sucrose (Tabela 3). Os pesquisadores concluíram que em condições de baixo pH e oxigênio limitante, *D. bruxellensis* produziu etanol com rendimento comparável ao de *S. cerevisiae* (0.44 g g⁻¹ glicose), entretanto a produtividade de *D. bruxellensis* só é compatível com *S. cerevisiae* no sistema contínuo de fermentação e em condições sintéticas de hidrolisados lignocelulose, a espécie testada *B. naardenensis* foi capaz de produzir etanol nas condições limitantes de oxigênio.

Análises estruturais e funcionais do cristal da enzima β -glucosidase da bactéria *Th. neapolitana* foram determinados numa resolução de 2.3 Å. De acordo com o estudo, a enzima é composta por 721 aminoácidos arranjados em três domínios. O primeiro domínio abrange do resíduo 1 ao 306 e é similar ao domínio da isomerase triose fosfato composta por oito estruturas α/β . O segundo domínio começa do resíduo 321 e termina no resíduo 536, o qual é formado por uma estrutura de cinco filamentos sanduíches α/β . O terceiro domínio estende do resíduo 600 ao 721 e possui um dobramento fibronectina tipo III, entretanto sua função é desconhecida. Os dois primeiros domínios têm funções diretas na arquitetura do sítio ativo da proteína (Pozzo *et al*, 2010). Apesar do contig 145 de *D. bruxellensis* apresentar apenas 279 resíduos de aminoácidos, esses resíduos se apresentaram à assinatura dos dois primeiros domínios da proteína, os quais estão relacionados com arquitetura do sítio ativo da enzima (Figura 5).

Tabela 3. Crescimento das linhagens do grupo *Brettanomyces/Dekkera* em meio de cultura contendo variadas fontes de carbono. (Galafassi *et al*, 2010)

CBS strains	Xylose	Xylitol	Ara	Glu	Gal	Man	Sucr	Mal	Cellob	Starch	Agar
<i>B. custersianus</i>											
4805	-	-	-	+	+	+	+	+	+	-	+
4806	-	-	-	+	+	+	+	-	+	+	-
5207, 8347	-	-	-	+	+	+	+	+	+	-	-
5208	-	-	-	+	+	+	+	-	+	-	-
<i>B. naardenensis</i>											
6040	+	+	+	+	+	+	+	+	+	+	-
6108	-	-	-	+	-	+	+	+	+	+	-
6115	+	+	-	+	+	+	+	+	+	+	-
7540	+	+	-	+	+	-	+	-	+	+	-
<i>B. nanus</i>											
1945	-	-	-	+	+	+	+	+	+	+	-
1947	-	-	-	+	+	+	+	+	+	+	+
<i>D. anomala</i>											
76, 3026	-	-	-	+	+	+	+	+	+	+	-
77, 4460, 7250	-	-	-	+	+	+	+	+	+	-	-
1938, 4211, 4212, 7654, 8138	-	-	-	+	+	+	+	+	+	+	+
4210, 8139	-	-	-	+	+	+	+	+	+	-	+
4608	-	-	-	+	+	+	+	-	+	-	-
4711	-	-	-	+	+	-	+	+	+	-	-
4712	-	-	-	+	+	+	+	-	+	+	+
<i>D. bruxellensis</i>											
72, 73, 75, 96, 98, 1940, 1941, 1942, 1943, 2499, 2547, 2796, 3025, 4482, 5206, 5512, 5513	-	-	-	+	+	+	+	+	+	+	-
74, 78	-	-	-	+	+	+	+	+	-	-	-
2336, 2797, 4459, 6055, 4481, 8027	-	-	-	+	+	+	+	+	+	-	-
4602	-	-	-	+	-	+	+	-	-	+	-
4601	-	-	-	+	+	+	+	-	-	-	-
4914	-	-	-	+	+	+	+	-	+	+	-

ara arabinose, *glu* glucose, *gal* galactose, *man* mannose, *sucr* sucrose, *mal* maltose, *cellob* cellobiose. Approximately 200 cells were spotted onto the plate: “-” means growth, and “+” means no growth after 6 weeks of incubation

Pozzo *et al* (2010) identificaram sete pontes de hidrogênio envolvendo a glicose e os resíduos do sitio ativos, D58, R130, K163 H164, R174, Y210 e E458. Interações de Van der Waals também ocorriam entre os átomos de carbono C4, C5 e C6 da molécula de glicose e os resíduos W243 e M207. Cinco dos sete resíduos do sitio ativo da seqüência codificante de β -glucosidase da bactéria *Th. neapolitana* foram encontrados no contig 145 de

Dekkera bruxellensis (Figura 9). Os dois resíduos restantes D58 e E 458 não foram identificados devido à seqüência de *D. bruxellensis* não está completa (Figura 10). Apesar de não conter dois os resíduos do sítio ativo podemos observar que os sítios ativo e a região ao seu entorno são bastante conservados tanto em resíduos quanto na distância entre os resíduos para todas as espécies analisadas, com exceção da bactéria *Th. Neapolitana*, o que nos faz inferir que os aminoácidos ácidos aspáticos e glutâmico devem ser também encontrados na seqüência completa da levedura *D. bruxellensis*. As interações de Van der Waals, W243 e M207, também foram identificados presentes no contig 145 nas posições W176 e M140, respectivamente.

O alinhamento múltiplo com espécies do grupo de ascomicetos e a bactéria *Th. neapolitana* mostrou sendo bastante conservado o sítio ativo das enzimas β -glucosidase (Figura 3), com a exceção da espécie *S. etchellsii* que apresentou uma leucina ao invés de uma tirosina na posição 210. Apesar do resíduo E458 encontrado na bactéria *Th. neapolitana*, não ter sido encontrado no contig 145 de *D. bruxellensis*, devido a seqüência não estar completa, Pozzo *et al* (2010) identificaram como sendo um resíduo não conservado dentro da família hidrolase glicosidase 3 (GH3). Baseando-se na estrutura tridimensional, os autores construíram mutantes para quatro resíduos, D242A, E458G, D58A e W243F. O mutante W243F apresentou um valor de Km próximo ao selvagem indicando que a afinidade ao substrato não foi afetada pela mutação, entretanto a acomodação diminuiu 34%. O mutante D242A perdeu cerca de 96% da atividade em relação ao selvagem em pH 5.6 o que indicou grande importância na atividade enzimática, efeitos similares ocorreram com os mutantes E458G e D58A.

A restauração da atividade dos sítios ativos dos mutantes ao adicionar nucleófilos externos tem sido usado em β -glucosidase das diferentes famílias ajudando a explicar o mecanismo enzimático e a identificar os resíduos catalíticos, desse modo, a atividade catalítica de D242 foi restaurada na presença de 0.5M NaN₃ confirmando o papel de nucleófilo catalítico, mesmo foi feito com o mutante D58, indicando ter um papel direto na catálise. Os resíduos D242 e W243 foi encontrados no contig 145 de *D. bruxellensis* correspondendo às posições D175 e W176.

A enzima β -glucosidase da levedura *K. marxianus* (Figura 2) apresenta 845 resíduos de aminoácidos distribuídos em três domínios: domínio GH3 N-terminal (resíduos 30-245), domínio GH3 C-terminal (resíduos 321-408, 560-700) e o domínio PA14 (resíduos 409-559). O domínio PA14 tem sido encontrado na estrutura de várias proteínas inserido no domínio GH3 C-terminal, entretanto, na levedura *K. marxianus* esse domínio se encontra inserido no núcleo catalítico sugerindo está envolvido na ligação do carboidrato no sítio ativo ou na estrutura da proteína (Yoshida *et al*, 2009).

O domínio PA14 é encontrado numa variedade de proteínas com glicosidases, glicosiltransferases, proteases e moléculas de sinalização, as quais indicam que esse domínio tem papel importante na ligação do carboidrato e em alguns casos, é fundamental nos eventos biológicos (Yoshida *et al*, 2010). O domínio PA14 da *K. marxianus* se inicia no resíduo 392 e se estende até o resíduo 559 (Yoshida *et al*, 2010) comparando essa região no alinhamento múltiplo das seqüências de β -glucosidase do grupo dos ascomicetos e da bactéria *Th. neapolitana* percebemos que a levedura *D. bruxellensis* não possui essa região devido sua seqüência está incompleta,

entretanto, sua seqüência tem uma similaridade de aproximadamente 61% com a seqüência de *k. marxianus*. Outro fato que notamos foi à ausência dessa região pela inserção de gaps na seqüência da bactéria *Th. neapolitana*. Apesar desse domínio não ser uma região fortemente conservada e ainda não ter uma função consenso, para a família GH3 esse domínio pode formar um sub-sítio e afetar a especificidade do substrato devido à enzima KmBg1I apresentar maior atividade por substratos menores através da inserção do PA14 no sanduíche α/β (Yoshida *et al*, 2010).

Woolfit *et al* (2007) seqüenciaram aproximadamente 40% do genoma da levedura *Dekkera bruxellensis*, e em ordem de determinar a sua posição filogenética no grupo dos hemiascomicetos, foram construídas 396 árvores filogenéticas baseado nas seqüências protéicas ortólogas de *D. bruxellensis* e nove outras espécies de fungos, *S. cerevisiae*, *Candida glabrata*, *Kluyveromyces lactis*, *Ashbya gossypii*, *Candida albicans*, *Debaryomyces hansenii*, *Yarrowia lipolytica*, *Schizosaccharomyces pombe* e *Aspergillus nidulans*. Os autores inferiram três topologias 1, 2 e 3 (Figura 13) que ocorreram em 34%, 27% e 26% dos casos, respectivamente. Para os resultados gerados, o teste Shimodaira-Hasegawa apontou a topologia 2 sendo a mais adequada, apesar de não ser significativamente melhor de que a topologia 1. A incerteza da posição filogenética de *D. bruxellensis* se mantém, análises separadas baseadas nas seqüências 18S suportam as topologias 1 e 3, e outras análises utilizando 400 seqüências codificantes de proteínas rejeitam significativamente a topologia 3 permanecendo incapaz de determinar se a espécie *D. bruxellensis* diverge de uma linhagem próxima a *C. albicans* e *D. hansenii* ou se essas três espécies divergiram de um único ancestral.

Uma investigação sistemática molecular dos membros do grupo dos ascomicetos dos gêneros *Brettanomyces*, *Debaryomyces*, *Dekkera* e *Kluyveromyces* foi realizada utilizando seqüências gênicas de 18S rRNA onde revelaram que *Brettanomyces* e *Dekkera* têm relações filogenéticas muito próximas formando um grupo filogenético distinto e estável. Foi evidenciado também que o gênero *Kluyveromyces* não forma um grupo monofilético, por apresentar várias espécies inseridas entre os outros gêneros, entretanto a espécie *Kluyveromyces marxianus* formou um clado com outros quatro *Kluyveromyces* apresentando uma similaridade de 99%. (Cai *et al*, 1996).

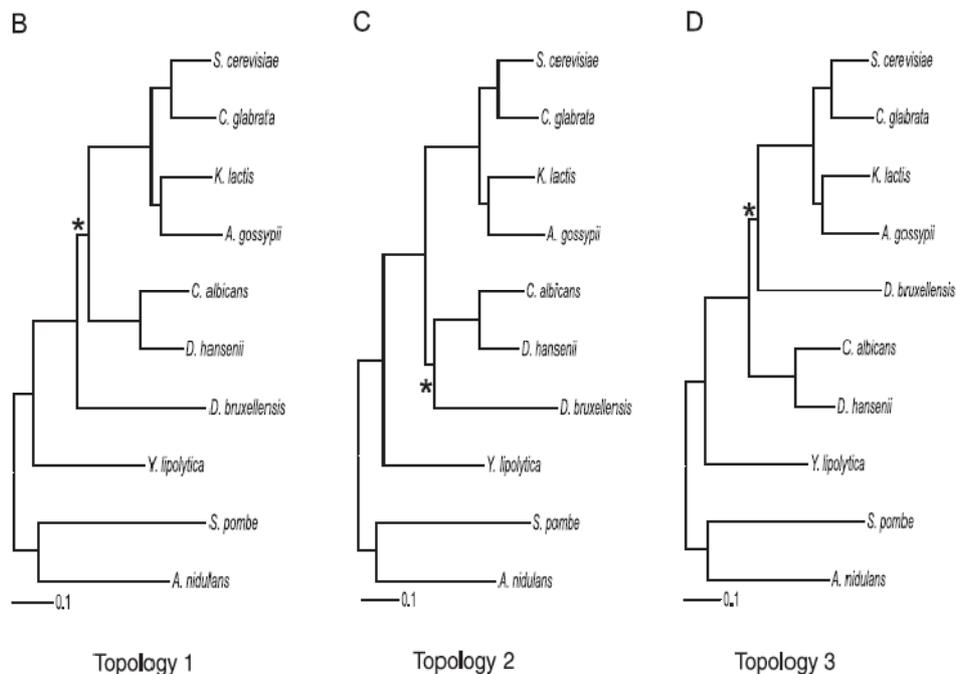


Figura 13. Topologias mais freqüentes encontradas a partir de seqüências protéicas de genes ortólogos de *D. bruxellensis*. (B) Topologia 1 apresentou uma freqüência de 34%. (C) Topologia 2 apresentou uma freqüência de 27%. (D) Topologia apresentou uma freqüência de 26%. Woolfit *et al*, 2007.

Nossos resultados mostraram que para esse gene *D. bruxellensis* divergiu de uma linhagem próxima de *K. marxianus* e *S. etchellsii*, sendo pertencentes às famílias Saccharomycetaceae e Debaryomycetaceae, respectivamente. Esses resultados corroboram com os resultados encontrados por Woolfit *et al* (2007) considerando ao nível de família (Figura 14). Em relação às outras espécies, como se era de se esperar, *A. niger*, *A. flavus*, *A. oryzae*, *A. fumigatus*, *A. clavatus*, *Neosartorya fischeri* e *Talaromyces stipitatus* formaram um grande clado sugerindo que para o gene da β -glucosidase essas espécies evoluíram de um ancestral comum seguindo sua classificação baseada nas características fenotípicas que os agrupam na família Trichocomaceae por serem fungos filamentosos de corpos frutíferos visíveis (Alexopoulos *et al*, 1996). Todas as leveduras *P. angusta*, *K. marxianus*, *S. etchellsii*, *S. stipitis*, *Candida tropicalis* também foram agrupadas em um grande clado, entretanto era de se esperar que as espécies *K. marxianus* e *P. angusta* formassem um grupo monofilético distinto por serem classificadas morfológicamente como pertencendo a família Saccharomycetaceae. O mesmo era de se esperar com as espécies *S. etchellsii* e *S. stipitis* que pertencem à família Debaryomycetaceae. A espécie *V. albo-atrum*, que é um fungo filamentoso, patógeno de plantas como o algodão e formador de conídio (DeSavigny, 2011), foi inserida no clado formado por leveduras. Já a bactéria *Th. neapolitana* se apresentou como um grupo externo, o que era de se esperar por ser geneticamente distantes das demais espécies.

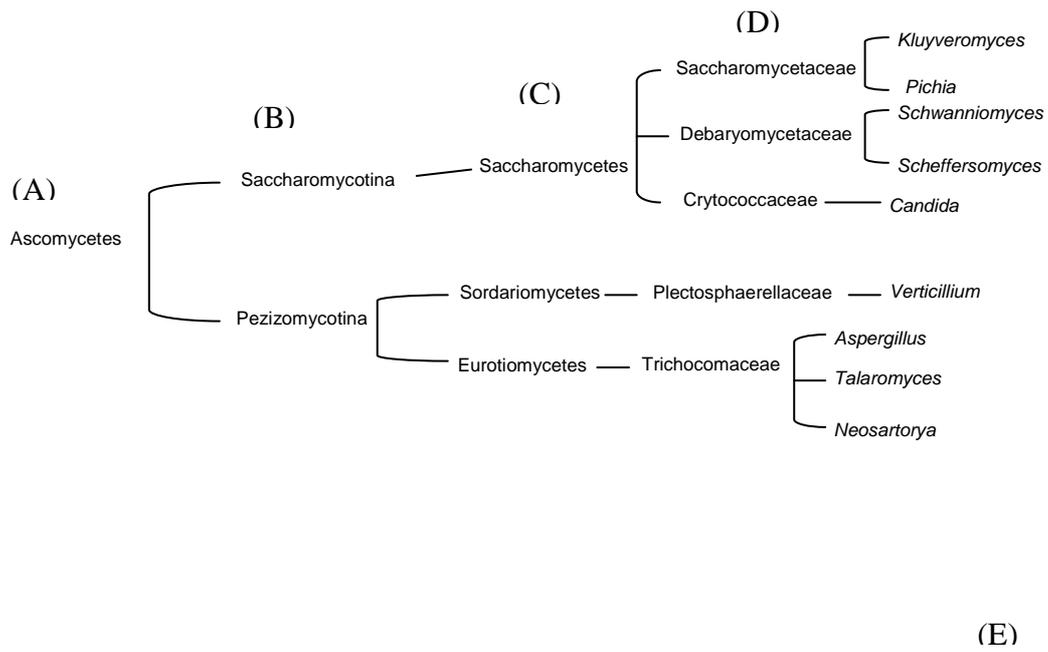


Figura 14. Esquema taxonômico do grupo dos ascomicetos. (A) Divisão. (B) Sub-divisão. (C) Classe. (D) Família. (E) Gênero.

7. CONCLUSÕES

- O alinhamento múltiplo permitiu visualizar as regiões conservadas e identificar os resíduos do sítio ativo R130 (Arginina), K163 (Lisina), H164 (Histidina), R174 (Arginina), Y210 (Tirosina) e E458 (Ácido glutâmico) a partir de dados encontrados na literatura.

- O gene incompleto de *Dekkera bruxellensis* apresentou os mesmos resíduos do sítio ativo da proteína beta-glucosidase encontrado na espécie *Thermotoga neapolitana*, com exceção dos resíduos R130 e E458.

- O gene incompleto de *Dekkera bruxellensis* apresentou dois dos três domínios encontrados na enzima da beta-glucosidase, o domínio N-terminal e o domínio C-terminal da família Glicolase Hidrogenase.

- O gene incompleto da levedura *Dekkera bruxellensis* apresentou uma similaridade de aproximadamente 61% com a seqüência codificante da enzima beta-glucosidase da levedura *Kluyveromyces marxianus*.

- Podemos inferir que o gene incompleto do contig 145 da levedura *Dekkera bruxellensis* é o gene codificante da enzima beta-glucosidase.

8. REFERÊNCIAS

ABBOTT, D.A; HYNES, S.H; INGLEDEW, W.M. Growth rates of *Dekkera/Brettanomyces* yeast hinder their ability to compete with *Saccharomyces cerevisiae* in batch corn mash fermentations. **Applied Microbiology Biotechnol.** v.66, n.6 . p. 641-647, 2005.

AGBOGBO, F.K; COWARD- KELLY, G. Cellulosic ethanol production using the naturally occurring xylose-fermenting yeast, *Pichia stipitis*. **Biotechnology Letters.** v, 30, n. 9, p. 1515-1524, 2008.

ALEXOPOULOS, C.J.; MIMS, C.W.; BLACKWELL, W. **Introductory mycology.** 4 ed. New York: John Wiley, 1996.

ALVARENGA, D. Produção de etanol só atenderá 45% dos carros flex em 2011, diz Única. Sem novos investimentos, proporção pode passar para 37% em 2020. Entidade prevê um déficit de 200 milhões de toneladas de cana em 2015.**G1**, São Paulo. www.globo.com/g1. Março, 2011.

ARAÚJO, K.M; OLIVEIRA, A.K.C; COSTA, B.G; QUEIROGA, R.N.G; PANNIR SELVAM, P.V. Estudo comparativo técnico e econômico de diferentes óleos vegetais brasileiros para produção de biocombustível. www.ufrngpec.hpg.com.br , 2002.

ARAÚJO, P.R.L.; BASÍLIO, A.C.M.; SIMÕES, D.A.; MORAIS JR, M.A.; MORAIS, J.O.F. Informação sobre algumas leveduras contaminantes da fermentação alcoólica

industrial isoladas no nordeste de Brasil. *In: SINAFERM*, 15., 2005, Recife. **Anais do XV Simpósio de Nacional de Bioprocessos**. Recife, Pernambuco, 2005.

AZEREDO, L.A.; GOMES, E.A.; MENDONÇA-HAGLER, L.C.; HAGLER, A.N. Yeast communities associated with sugarcane in Campos, Rio de Janeiro, Brazil. **International Microbiology**. v. 1, p. 205-208, 1998.

BALDRIAN, P; VALASKOVA, V. Degradation of cellulose by basidiomycetous fungi. **FEMS Microbiology**. v.32, p. 501-521, 2008.

BARNETT, J.A.; PAYNE, W.; YARROW, D. **Yeast: characteristics and identification**. 2. ed. Cambridge: University Press, 1990.

BASÍLIO, A.C.; PINHEIRO, W.; BRASILEIRO, B.; MORAIS JR, M.A.; SIMÕES, D.A. Utilização do padrão de amplificação com o marcador (GTG)₅ para identificação rotineira de leveduras contaminantes da fermentação alcoólica industrial. *In: SINAFERM*, 15., 2005, Recife. **Anais do XV Simpósio Nacional de Bioprocessos**. Recife, Pernambuco, 2005.

BASÍLIO, A.C; ARAÚJO, P.R.L; MORAIS, J.O.F; SILVA-FILHO, E.A; MORAIS JR, M.A; SIMÕES, D.A. Detection and identification of wild yeast contaminants of the industrial fuel ethanol fermentation process. **Current Microbiology**. v. 56, p. 322-326, 2008.

BENSON, D.A; KARSCH-MIZRACHI, I; LIPMAN, D.J; OSTELL, J; WEELER, D.L.
GenBank. **Nucleic Acids Research**. v. 31, n1, p. 23-27, 2003

BIOEDIT. Disponível em:< <http://iubio.bio.indiana.edu/molbio/seqpup/>>. Acessado em: 5 out 2007.

CABRINI, K.T. & GALLO, C.R. Identificação da leveduras contaminantes no processo de fermentação alcoólica na Usina Santa Elisa. **STAB**, v. 17, n. 4, p. 48-50, 1999.

CAI, J; ROBERTS, I.N; COLLINS, M.D. Phylogenetic Relationships among Members of the Ascomycetous Yeast Genera *Brettanomyces*, *Debaryomyces*, *Dekkera*, and *Kluyveromyces* Deduced by Small-Subunit rRNA Gene Sequences, **International Journal Of Systematic Bacteriology**, v. 46, n. 2, p. 542-549, 1996.

CARAZZOLLE, M; PERIRA, G. Disponível em:< <http://www.lge.ibi.unicamp.br/manuais/ManualGP.htm>>. Acessado em: 20 set 2007.

CECCATO-ANTONINI, S.R. & SILVA, D.F. Eficiência de meios diferenciais no isolamento de cepas de leveduras de processo industriais de fermentação alcoólica. **STAB**. v.18, n. 4, p. 40-46, mar/abr, 2000.

CHEAVEGATTI-GIANOTTO, A; ABREU, H.M.C; ARRUDA, P; BESPALHOK-FILHO, J.C; BURNQUIST, W.L; CRESTE, S; CIERO, L; FERRO, J.A; FIGUEIRA, A.V.O; FILGUEIRAS, T.S; GROSSI-DE-SÁ, M.F; GUZZO, E.C; HOFFMANN, H.P; LANDELL, M.G.A; MACEDO, N; MATSUOKA, S; REINACH, F.C; ROMANO, E;

SILVA, W.J; SILVA-FILHO, M,C; ULIAN, E.C. Sugarcane (*Saccharum x officinarum*): A reference study for the regulation of genetically modified cultivars in Brazil. **Tropical Plant Biology**. v.4, p. 62-89, 2011.

CLUSTAL W. Disponível em:< <http://www.ebi.ac.uk/>>. Acessado em: 15 set 2007.

COCOLIN, L.; RANTSOU, K.; IACUMIN, L.; ZIRONI, R.; COMI, G. Molecular detection and identification of *Brettanomyces/Dekkera bruxellensis* and *Brettanomyces/Dekkera anomalus* in spoiled wines. **Applied and Environmental Microbiology**. v.70, n.3, p. 1347-1355. mar, 2004.

COMITINI, F; DE INGENIIS, J; PEPE, L; MANNAZZU, I; CIANI, M. *Pichia anomola* and *Kluyveromyces wickerhamii* killer toxins as new tools against *Dekkera/Brettanomyces* spoilage yeasts. **FEMS Microbiology Letter**. v.239, p. 235-240, 2004.

COSTA-NETO, P.R; ROSSI, L.F.S; ZAGONEL, G.F; RAMOS, L.P. Produção de biocombustível alternativo ao óleo diesel através da transesterificação de óleo de soja usado em frituras. **Química Nova**. v. 23, n. 4, p. 531-537, 2000.

DAL POGGETTO, P. Metade da frota nacional em 2012 será de bicombustíveis, diz Única. Etanol já substituiu 50% do consumo de gasolina no Brasil. Brasil possui mais 7 milhões de carros 'flex' atualmente. **G1**. São Paulo. <http://g1.globo.com/Noticias/Carros/0,,MUL1090989-9658,00->

METADE+DA+FROTA+NACIONAL+EM+SERA+DE+BICOMBUSTIVEIS+DIZ+UNICA.html. Julho. 2011.

DESAVIGNY, T. *Verticillium albo-atrum*. North Carolina State University. http://www.cals.ncsu.edu/course/pp728/alboatrum/Verticillium_albo-atrum.html.

Acessado em: Novembro de 2011.

DIAS, L; DIAS, S; SANCHO, T; STENDER, H; QUEROL, A; MALFEITO-FERREIRA, M; LOUREIRO,V. Identification of yeast isolated from wine-related environments and capable of producing 4-ethylphenol. **Food Microbiology**. v.20, p.567-574, 2003.

GALAFASSI, S; MERICO, A; PIZZA, F; HELLBORG, L; MOLINARI, F; PISKUR, J; COMPAGNO, C. *Dekkera/Brettanomyces* yeast for ethanol production from renewable sources under oxygen-limited and low-pH conditions. **Journal of Industrial Microbiology and Biotechnology**. 2010.

GRANCHI, L; BOSCO, M; MESSINI, A; VINCENZINI, M. Rapid detection and quantification of yeast species during spontaneous wine fermentation by PCR-RFLP analysis of the rDNA ITS region. **Journal of Applied Microbiology**. v.87, p.949-956, 1999.

GUERRA, E.J. & ANGELIS, D.F. Flocculação da levedura induzida por bactéria na fermentação etanólica: I método de detecção preventiva e estudos para o controle. **STAB**, v. 16, n. 6, p. 25-27. jul/ago,1998.

HAHN-HAGERDAL, B; GALBE, M; GORWA-GRAUSLUND, M.F; LIDÉN, G; ZACCHI, G. Bio-ethanol – The fuel of tomorrow from the residues of today. **TRENDS in Biotechnology**. v.24, n. 12, p.549-556, 2006.

ISHIKAWA, E; SAKAI, T; IKEMURA, H; MATSUMOTO, K; ABE, H. Identification, cloning and characterization of a *Sporobolomyces singular* in beta-galactosidase-like enzyme involved in galacto-oligosaccharide production. **Journal Bioscience Bioengineering**. v, 99, p. 331-339, 2005.

LEMEY, P; SALEMI, M; VANDAMME, A.M. **The phylogenetic handbook: A practical approach to phylogenetic analysis and hypothesis testing**. 2 ed. New York: Cambridge University Press, 2009.

LIBERAL, A.T.S. Identificação molecular da levedura *Dekkera bruxellensis* como principal contaminante do processo de fermentação alcoólica industrial. 2005.76f. Dissertação (Mestrado em Genética) Universidade Federal de Pernambuco, Recife, 2005.

LIBERAL, A.T.S; BASÍLIO, A.C.M; RESENDE, A.M; BRASILEIRO, B.T.V; SILVA-FILHO, E.A; MORAIS, J.O.F; SIMÕES, D.A; MORAIS JR, M.A. Identification of *Dekkera bruxellensis* as a major contaminant yeast in continuous fuel ethanol fermentation. **Journal of Applied Microbiology**. v. 102, p. 538-547, 2007.

LYND, L.R; WEIMER, P.J; VAN ZYL, W.H; PRETORIUS, I.S. Microbial cellulose utilization: Fundamentals and biotechnology. **Microbiology and Molecular Biology Reviews.** v. 66, p. 506-577, 2002.

MARQUES, F. O alvo é o bagaço: Subproduto abundante da indústria da cana dá vantagem competitiva ao Brasil na busca do etanol de segunda geração. **Pesquisa FAPESP 163.** Setembro, 2009.

MORAES, L.M.P; ASTOLFI-FILHO, S; OLIVER, S.G. Development of yeast strains for the efficient utilisation of starch: evaluation of constructs that express α -amylase and glucoamylase separately or as bifunctional fusion proteins. **Applied Microbiology and Biotechnology.** v.43, p. 1067-1076, 1995.

MITRAKUL, C.M; HENICK-KLING, T; EGLI, C.M. Discrimination of *Brettanomyces /Dekkera* yeast isolates from wine by using various DNA finger-printing methods. **Food Microbiology.** v.16, p.3-14, 1999.

ONISHI, N; TANAKA, T. Purification and properties of a galacto-and-gluco-oligosaccharide producing beta-glycosidase from *Rhodotorula minuta* IFO879. **Journal of Fermentation and Bioengineering.** v. 82, p. 439-443, 1996.

PARAZZI, C.; SILVA, A.R.da; ARAUJO, R.T.de. Avaliação de diferentes linhagens de leveduras selvagens na fermentação alcoólica. **STAB.** V. 20, n. 3, p. 36-40, jan/fev, 2002.

POZZO, T; PASTEN, J.L; KARLSSON, E.N; LOGAN, D.T. Structural and functional analyses of β -glucosidase 3B from *Thermotoga neapolitana*: A thermostable three-domain representative of glycoside hydrolase 3. **Journal of Molecular Biology**. v.397, p. 724-739, 2010.

PRODOCINI, F; FILHO, F.C; CERQUEIRA, G. C; KESSEDJIAN, J. L; BINNECK, E; PETRETSKI, J. H; SILVA, A. F; CAMARGO, L. P; REIS, A. N; FERREIRA, R. G. M; JUNQUEIRA, A. C. M; LIMA, R. O; SANTOS, A. C F; PEREIRA, R. M; JUNIOR, A. N; JARDIM, S; WUST, C. I; SAMPAIO, V. S; FOLGUERAS-FLATSCHART, A. V. Bioinformática: Manual do Usuário. **Biotecnologia Ciência & Desenvolvimento**. v. 12, n.12, p. 12-25. 2002

ROUZÉ, P; PAVY N. and ROMBAUTS, S. Genome annotation: which tool do we have for it?. **Current Opinion in Plant Biology**. v. 2, p. 90-95. 1999.

RODRIGUES, N; GONÇALVES, G; PEREIRA-DA-SILVA, S; MALFEITO-FERREIRA, M; LOUREIRO, V. Development and use of a new médium to detect yeast of the genera *Dekkera*/*Brettanomyces*. **Journal of Applied Microbiology**. v. 90. n. 4, p. 588-599, 2001.

SAITOU, N; NEI, M.S. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**. v. 4, n. 4, p. 406-425, 1987.

SHERATA, A.M. Yeasts isolated from sugar cane and its juice during the production of aguardente de cana. **Journal of Applied Microbiology**. v. 8, p. 73-75, 1960.

SILVA-FILHO, E.A.; SANTOS, S.K.B.; RESENDE, A.M.; MORAIS, J.O.F.; MORAIS JR, M.A.; SIMÕES, D.A. Yeast population dynamics of industrial fuel-ethanol fermentation process assessed by PCR-fingerprinting. **Antonie van Leeuwenhoek**. v. 88, p. 13-23, 2005a.

SILVA-FILHO, E.A.; MELO, H.F.; ANTUNES, F.A.; SANTOS, S.K.B.; RESENDE, A.M.; SIMÕES, D.A.; MORAIS JR, M.A. Isolation by genetic and physiological characteristics of a fuel-ethanol fermentative *Saccharomyces cerevisiae* strain with potential for genetic manipulation. **Journal of Industrial Microbiology and Biotechnology**. v. 32, p. 481-486, 2005b.

SINDAÇÚCAR. [São Paulo: Sn., Sd.] Disponível em: <<http://www.sindacucar.com.br>>. Acesso em: 17 Julho 2011.

SLININGER, P.J; THOMPSON, S.R; WEBER, S; LIU, Z.L; MOON, J. Repression of xylose-specific enzymes by ethanol in *Scheffersomyces (Pichia) stipitis* and utility of repitching xylose-grown populations to eliminate diauxic lag. **Biotechnology and Bioengineering**. v. 108, n. 8, p. 1801-1815, 2011.

SOCCOL, R.C; VANDENBERGHE, L.P.S; MADEIROS, A.B.P; KARP, S.G; BUCKERIDGE, M; RAMOS, L.P; PITARELO, A.P; FERREIRA-LEITÃO, V; GOTTSCHALK, L.M.F; FERRARA, M.A; BON, E.P.S; MORAES, L.M.P; ARAÚJO, J.A; TORRES, F.A.G. Bioethanol from lognocelluloses: Status and perspectives in Brazil. **Bioresource Technology**. v.101, p. 4820-48-25, 2010.

STROPPA, C.T.; ANDRIETTA, S.R.; ANDRIETTA, M.G.S.; SERRA, G.E. Dinâmica populacional de levedura em processo de fermentação alcoólica. *In: SINAFERM*, 13., 2000, Teresópolis. **Anais do X Simpósio Nacional de Bioprocessos**. Teresópolis, Rio de Janeiro, 2000. p. 1-5.

TATENO, Y; IMANISHI, T; MIYAZAKI, S; FUKAMI-KOBAYASHI, K; SAITOU, N; SUGAWARA, H; GOJOBORI, T. DNA Data Bank of Japan (DDBJ) for genome scale research in life science. **Nucleic Acids Research**. v. 30, n.1, p. 27-30, 2002.

THOMPSON, J.D; GIBSON, T.J; PLEWNIAK, F; JEANMOUGIN, F; HIGGINS, D.G. The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Research**. v. 25, n. 24, p. 4876-4882, 1997.

TAMURA, K; DUDLEY, J; NEI, M; KUMAR S. Molecular Evolutionary Genetics Analys (MEGA) Software Version 4.0. **Molecular Biology and Evolution**. v.24, n.8, p. 1596-1599, 2007.

ÚNICA. O álcool combustível. [São Paulo: Sn., Sd.] Disponível em: <http://www.unica.com.br/pages/alcool_combustivel.asp>. Acesso em: 11 maio 2005.

ÚNICA. O álcool combustível. [São Paulo: Sn., Sd.] Disponível em: <http://www.unica.com.br/pages/alcool_combustivel.asp>. Acesso em: 17 Julho 2011.

VIDAL, M.F.; SANTOS, J.A.N.; SANTOS, M.A. Setor sucroalcooleiro no nordeste brasileiro: Estruturação da cadeia produtiva, produção e mercado. *In: XLIV Congresso da Sociedade Brasileira de Economia e Sociologia Rural*. Fortaleza-CE. 2006. Disponível em: <http://www.bnb.gov.br/content/aplicacao/ETENE/Artigos/doc/sober_sucroalcooleiro.pdf>. Acessado em: 10 jan 2007.

WHEALS, A.E.; BASSO, L.C.; ALVES, D.M.G.; AMORIM, H. Fuel ethanol after 25 years. **Tibtech**. v. 17, p. 482-486, 1999.

WOOLFIT, M; ROZPEDOWSKA, E; PISKUR, J; WOLFE, K.H. Genome survey sequencing of the wine spoilage yeast *Dekkera (Brettanomyces) bruxellensis*. **Eukaryotic Cell**. v.6, n.4, p. 721-733, 2007.

YOSHIDA, E; HIDAKA, M; FUSHINOBU, S; KOYANAGI, T; MINAMI, H; TAMAKIS, H; KITAOKA, M; KATAYAMA, T; KUMAGAI, H. Purification, crystallization and preliminary X-ray analysis of β -glucosidase from *Kluyveromyces marxianus* NBRC1777. **Acta Cryst**. v.65, p. 1190-1192, 2009.

YOSHIDA, E; HIDAKA, M; FUSHINOBU, S; KOYANAGI, T; MINAMI, H; TAMAKIS, H; KITAOKA, M; KATAYAMA, T; KUMAGAI, H. Role of a PA14 domain in determining substrate specificity of a glycoside hydrolase family 3 β -glucosidase from *Kluyveromyces marxianus*. **Biochemical Journal**. v. 431, p. 39-49, 2010.

9. ANEXOS

>Seqüência de Nucleotídeos do contig 145 da levedura *Dekkera bruxellensis*.

AAGGATATGGCATTGGTCCAATGACGAGAACCGGGGTTTCCTTTAAAAGT
GGTAGAACTTTCCTCTCATTTTTAAGTAGTACAATACCTTCAGCAGCAGCTC
TCCGTAATTTTTCACCTTGTTTCGGGCTTATCGTTTAGCGAATCCTCTACCTT
GCTAAAGAGTGTTGTTCCATTGGATTCAATCAATGACTTAATAACTTTTAAA
ATGTGTCGACAATGATTGTCAATATCTGTTTGTTCAGTTATCCGCTTTAG
ATTGTAAAAGATGCTTGATTGTATCCCACCTCCTAAATTGTGAAGGTCCTG
GAAATTCAATATCAATTCCATGCTTCAAGGAGTTCAACGTAGAATATGTTCC
AAACCAATCGCTCATAAGTAAACCATCCAAATTCCATTCTTTTCTTAAAATG
TTTTCAATTAATTGATAGTTTTCTGAACAGTGTATTCCATTTACTTTATTGTA
AGCGGTCATAATACAATTTGGATTAGCAAGTTGCACAGCGATTCTGAACGG
TTCTAAGTATATTTCCCTTAATGCCCTTTCACCTTACAACACTGTTTCGAGCTC
AGTCGTTTCATGTTCTAAATCATTACAAACAAAATGCTTAACTGTAGCACAAA
TCCGACCATCATTTTGTATTCCCTTTCACGATTGCCGCCGTGCATAATCCACT
TAATACTGGATCCTCGCTAAAGCTCTCAAATCCCCGTCCACCCAAGGGTCC
CCTCAGAATGTTTGTGTTGGACCAAGAATTATTTGGGCGCCCTTATGTTTT
GCCTCAACTGCCATCAACTCTCCAATTTCTTCTAACAAGTCGCAGTCAAAT
GTCGATGCCAACCCGTGCCATTTGGAAAGCACGCTGATGGTACACCGTT
AAAAAATTTTGTTC

>Seqüência de Aminoácido do contig 145 da levedura *Dekkera bruxellensis*.

GTKFFNGVPSACFPNGTGLASTFDCDLLEEIGELMAVEAKHKGAQIILGPTTNI
LRGPLGGRGFESFSEDPVLSGLCTAAIVKGIQNDGRICATVKHFVCNDLEHER
LSSNSVVSERALREIYLEPFRIAVQLANPNCIMTAYNKVNGIHCSENYQLIENIL
RKEWNLDGLLMSDWFGTYSTLNSLKHGIDIEFPGPSQFRRWDTIKHLLQSKAD
NLKQTDIDNHCRHILKVIKSLIESNGTTLFSKVEDSLNDKPETSEKLRRAAEGI
VLLKNERKVLPLLKETPVLVIGPNA

10. MEMORIAL

Rochane Regina Neves Baptista Torres [rochanebaotista@hotmail.com], nascida em 08 de fevereiro 1982, Bacharel em Ciências Biológicas, título obtido em 16 de outubro de 2005, pela Universidade Federal de Pernambuco, com o trabalho de conclusão de curso: Análise da microbiota de levedura do processo fermentativo da usina Santa Teresa na safra 2004/2005. Especialista em Biologia Molecular pela UPE, em 9 de novembro de 2007, com monografia Identificação e análise estrutural do gene codificante da enzima piruvato descarboxilase de *Saccharomyces cerevisiae* em *Dekkera bruxellensis*. Já como resumo em anais de congresso publicou como primeira autora os seguintes trabalhos: Identificação e análise estrutural do gene da levedura *Dekkera bruxellensis* ortólogos a família dos genes que codificam a enzima piruvato descarboxilase em *Saccharomyces cerevisiae* (54^o Congresso Brasileiro de Genética, 2008), Tipagem genética da microbiota do processo fermentativo alcoólico da usina Santa Teresa e seleção da levedura dominante (53^o Congresso Brasileiro de Genética, 2007) e os publicou como segunda autora os seguintes trabalhos: Análise da expressão gênica através de PCR em tempo real da enzima piruvato descarboxilase da levedura *Dekkera bruxellensis* (56^o Congresso Brasileiro de Genética, 2010), Identificação e análise estrutural do gene codificante da enzima álcool desidrogenase de *Saccharomyces cerevisiae* em *Dekkera bruxellensis* (54^o Congresso Brasileiro de Genética, 2008). Como pesquisadora, realizou estágio no Laboratório de Biologia Molecular da UnB, de 05 de março a 12 de abril de 2006 com o seguinte objetivo: Clonagem molecular de cassete de expressão alfa-amilase e glicoamilase em células de *Escherichia coli* (DH5 α) para a construção de um vetor de expressão para ser integrado ao genoma da levedura *Saccharomyces cerevisiae*. Como parte da formação técnica, participei Applied Biosystem: Muito além do genoma, dentre outros cursos de curta duração.